

HIGHWAYS AND GLOBALIZATION

Taylor Jaworski  
Carl Kitchens  
Sergey Nigai

WORKING PAPER 27938

NBER WORKING PAPER SERIES

HIGHWAYS AND GLOBALIZATION

Taylor Jaworski  
Carl Kitchens  
Sergey Nigai

Working Paper 27938  
<http://www.nber.org/papers/w27938>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
October 2020, Revised April 2022

We thank Johannes Uhl for assistance with ArcGIS. Comments from seminar participants at Auburn University, Clemson University, University of Colorado (Boulder), Florida State University, Hitotsubashi University, University College Dublin, Wake Forest University, West Virginia University, and several conferences are greatly appreciated. Any remaining errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2020 by Taylor Jaworski, Carl Kitchens, and Sergey Nigai. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Highways and Globalization

Taylor Jaworski, Carl Kitchens, and Sergey Nigai

NBER Working Paper No. 27938

October 2020, Revised April 2022

JEL No. F11,F14,H54,R13,R42

**ABSTRACT**

This paper quantifies the value of US highways. We develop a multisector general equilibrium model with many locations in the United States (i.e., counties) and many countries. In the model, producers choose shipping routes subject to domestic and international trade costs, endogenous congestion, and port efficiency at international transshipment points. We find that removing the Interstate Highway System reduces real GDP by \$601.6 billion (or 3.8 percent). We also show how to quantify the value of individual segments using our framework. The results highlight the role of domestic transportation infrastructure in shaping regional comparative advantage as well as the gains from intersectoral and international trade.

Taylor Jaworski  
Department of Economics, UCB 256  
University of Colorado at Boulder  
Boulder, CO 80309  
and NBER  
taylor.jaworski@colorado.edu

Sergey Nigai  
Department of Economics  
University of Colorado, Boulder  
Boulder, CO 80309  
sergey.nigai@colorado.edu

Carl Kitchens  
Department of Economics  
Florida State University  
239 Bellamy Building  
Tallahassee, FL 32306  
and NBER  
ckitchens@fsu.edu

# 1 Introduction

Countries around the world devote substantial resources to new road infrastructure and maintenance. Given the size of these expenditures, it is crucial to understand the mechanisms behind the aggregate and distributional effects of transportation infrastructure as a guide for ongoing and future investments. Quantifying these benefits is challenging since the quality of domestic infrastructure shapes the pattern of specialization vis-à-vis domestic and foreign trading partners, determines congestion, and generates spillovers across industries and locations. In addition, calculating benefits at a refined geographic scale, e.g., for counties in the United States, is subject to constraints due to data limitations.

In this paper, we develop a novel methodology to address these challenges and ultimately quantify the value of transportation infrastructure in the United States. In particular, focusing on highways, our approach captures the role of the transportation system in mediating the intensity of spatial and sectoral linkages along two margins. First, producers choose optimal routes to domestic and international markets such that the domestic portion of industry-specific trade costs reflect travel time via the US highway network or using alternative modes (e.g., rail, water, or air). Second, producers' decisions take into account congestion endogenously generated by industry-specific trade.

Our theoretical framework integrates each US county with all other counties and all foreign trading partners. To do this, we use a two-tier spatial structure that combines state- and county-level data with measures of industry-specific domestic and international trade costs. The framework also incorporates more standard features of economic geography models, including input-output linkages, imperfect labor mobility, and agglomeration. We apply the model to data for the entire contiguous United States made of more than 3,000 counties, 22 sectors, and 36 international trading patterns (as well as the rest of the world).

In counterfactual exercises, we use the model to produce three main results. First, we quantify the value of the Interstate Highway System (IHS). Removing the entire IHS decreases real GDP by \$601.6 billion in 2012 dollars (or 3.8 percent). Both intersectoral trade via input-output linkages and international trade play an important role in determining the aggregate effects. Regionally, losses are concentrated in the Northeast and West of the United States and more remote counties experience the largest relative losses.

Second, underlying the intuition for these results is that the IHS allows remote regions to exploit their comparative advantage and concentrate production in a few sectors with relatively high productivity. We confirm this by showing how removing the IHS affects a measure of revealed comparative advantage (see [Balassa, 1965](#)). In particular, we show

that removing the IHS leads alters the sectoral composition of output across US states. In addition, states consume more of their own production and export less to other states and foreign countries. This suggests that the reduction in trade costs from the IHS plays an important role in shaping the location of production, the pattern of specialization, and the distribution of the gains from trade across US regions.

Finally, we quantify the value of ten of the largest IHS interstates (I-5, I-10, I-95, etc). From this exercise we find that I-10, I-75, and I-70 are the most valuable. Aggregate losses from removing these ten segments range between \$8.8 and \$51.7 billion and losses per mile range between \$4.7 and \$22.7 million. We also quantify the contribution of each segment to international trade costs and find this to be approximately half of the total effect. These results are useful for allocating funds for maintenance and repair of existing highways or evaluating proposed changes.

In the end, we provide a framework that highlights the interaction between domestic transportation infrastructure and international trade while accounting for congestion, input-output linkages, and other salient features of economic geography environments. This paper builds on the literature in two ways. First, we contribute to research on the impact and value of the IHS in the United States. One strand of this literature estimates the effects of transportation infrastructure on economic activity (Isserman and Rephann, 1994; Fernald, 1999; Chandra and Thompson, 2000; Michaels, 2008) and household location decisions (Baum-Snow, 2007).<sup>1</sup> For example, Herzog (2021) examines the impact of domestic market access due to the construction of the Interstate Highway System on employment and wages at the county level.

Another strand of this literature uses quantitative models to value the Interstate Highway System. Most closely related to this paper is work by Allen and Arkolakis (2014, 2019). These authors quantify the value of the Interstate Highway System focusing on aggregate domestic trade and the welfare gains associated with improving shorter sections.<sup>2</sup> Our contribution is to consider the impact of the IHS in the presence of input-output linkages and international as well as domestic trade. As in Allen and Arkolakis (2019), we account for the role of endogenous congestion levels and route choice.<sup>3</sup> Our results suggest that accounting input-

---

<sup>1</sup>In addition, there is a growing literature that estimates the effects of transportation infrastructure in the context of developing countries (e.g., Faber, 2014; Baum-Snow, Brandt, Henderson, Turner and Zhang, 2017; Cosar, Demir, Ghose and Young, 2020).

<sup>2</sup>Also related is work on historical railroads (Fogel, 1964; Fishlow, 1965; Donaldson and Hornbeck, 2016; Donaldson, 2018; Nagy, 2020) and highways more recently (Alder, 2017; Bartelme, 2018; Jaworski and Kitchens, 2019) as well research on optimal infrastructure investment in general equilibrium settings (Fajgelbaum and Schaal, 2020).

<sup>3</sup>Our specific approach to incorporating congestion differs from Allen and Arkolakis (2019) and may be useful

output linkages and international trade are crucially important for the quantifying the value of infrastructure: the estimated value of the IHS increases threefold relative to previous estimates in the literature.<sup>4</sup>

Second, we contribute to recent work on the role of domestic trade costs in shaping trade and welfare (Agnosteva, Anderson and Yotov, 2019; Atkin and Donaldson, 2015; Coşar and Demir, 2016; Coşar and Fajgelbaum, 2016; Fajgelbaum and Redding, 2018; Redding, 2016; Ramondo, Rodríguez-Clare and Saborío-Rodríguez, 2016; Bartelme, 2018; Ramondo, Rodríguez-Clare and Saborío-Rodríguez, 2019). We show that improvements in domestic transportation allows remote regions to concentrate output and exports in comparatively advantaged sectors. This leads to large welfare gains and has important distributional consequences within and between countries. In addition, our approach to measuring trade costs (and the related congestion) is novel. We use detailed information on travel time as a function of distance, speed, and traffic on all county-to-county and county-to-port routes. This provides tractable approach for incorporating congestion and yields a straightforward way to decompose the contribution of highways (or any other portion of the transportation system) to domestic versus international market access.

The remainder of this paper is organized as follows. In the next section, we describe the key components of the US highway network and describe how we use data on distance, speed, and traffic to calculate travel time to incorporate congestion. Section 3 presents the model of interregional and international trade, including the role of domestic and international trade costs, and the solution method for carrying out counterfactuals. In Section 4, we provide an overview of the data used to calibrate the model, which includes the construction of sector-specific trade costs and the estimation of sector-specific model parameters. Section 5 presents our main results and Section 6 concludes.

## 2 The US Highway Network with Congestion

There are over four million miles of paved road in the United States. The Interstate Highway System (IHS) comprises nearly 50,000 miles with posted speeds typically set at 70 miles per hour. Although it accounts for roughly 1 percent of paved road mileage in the United States, the IHS facilitates one quarter of vehicle miles traveled annually, while most of the remainder

---

in other settings.

<sup>4</sup>This result is consistent with Costinot and Rodríguez-Clare (2014), Caliendo and Parro (2015), and Ossa (2015) that shows including input-output linkages magnifies the gains from trade. To the best of our knowledge, our paper is the first to emphasize this mechanism in the context of domestic transportation infrastructure.

of the system is made up of a combination of federal-aid and state highways. The US highway network used in this paper is shown in Panel A of Figure 1 and includes the Interstate, other federal-aid, and state highways. The remaining roads (not shown) are primarily used for local travel including county roads as well as city and neighborhood streets.

The highway network shown in Panel A of Figure 1 includes the major roads used for the movement of goods within the United States and constitutes the main focus of our analysis.<sup>5</sup> In 2010, trucking accounted for almost half of ton-miles nationally. The fraction of the value of domestic trade moved by truck was nearly 70 percent relative to 10 percent by rail and 5 percent by water. The highway network also provides important links for international trade. In general, 30 percent of all imports by weight use trucks exclusively to deliver goods domestically, while 34 percent use trucks for at least part of their journey. For exports, more than half of shipments by weight use trucks as a single mode to deliver goods to ports and 60 percent used trucks partially to ship goods to ports. For trade with Canada and Mexico, highways play a larger role: 70 percent of the value of trade between the United States, Canada, and Mexico is transported on US highways (Bureau of Transportation, 2017).

While highways in the United States are vital for facilitating domestic and international trade, movement of goods via highways is subject to congestion. For example, recent surveys suggest that congestion costs are as high as 100 hours per driver each year (INRIX Research, 2019). To measure the severity of congestion, the Federal Highway Administration collects average annual daily traffic conditions on highways from state-level agencies. The average annual daily traffic is then combined with the road capacity to create a measure that reflects how congestion affects travel speeds, this measure is known as the level of service (LOS). We illustrate observed values of the LOS on different segments of US highways in Panel B of Figure 1. Using LOS suggests that nearly 18,000 miles (or 40 percent) of the IHS experiences reduced travel speeds due to congestion. The figure also suggests that the levels of congestions are highly heterogeneous across segments and concentrated in the East and West of the United States. For this paper, this suggests that incorporating congestion costs is important for quantifying the impact of highways on regional specialization and trade outcomes.

As mentioned above, Figure 1 includes the IHS, US Highways, and state highways. Shapefiles of this network that we use in this paper divide these key components into roughly 330,000 individual sections. We observe LOS, distance, and speed, which together determine the travel time on each section of the highway network, which we denote with subscript  $\mathcal{S}$ .

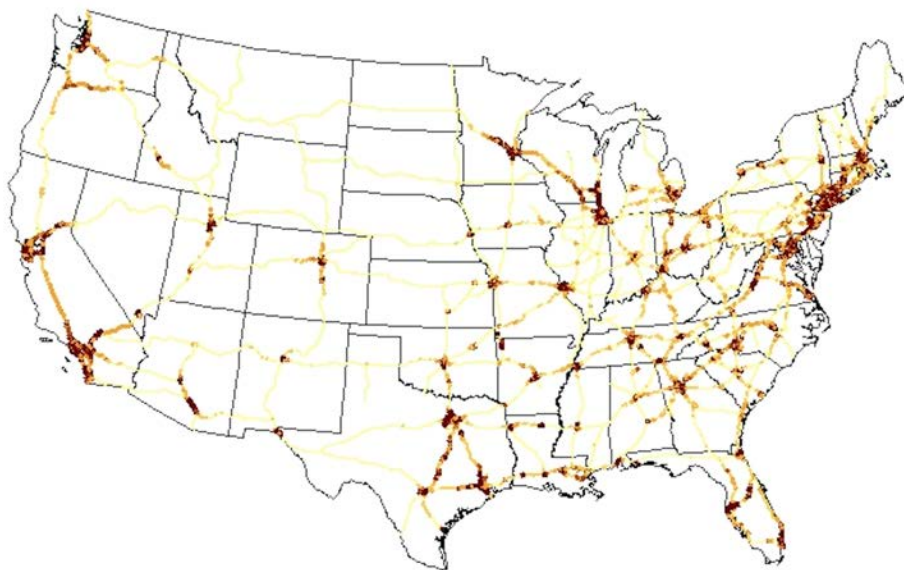
---

<sup>5</sup>While portions of domestic trade are conducted via waterways, railroads and air transportation, trucking by far remains the most important mode of transportation.

**Figure 1: US Highways and Congestion**



A. Components of the US Highway Network



**Level of Service (LOS) on Interstates**  
— 0<LOS<.55 — .55<LOS<.77 — .77<LOS<.92 — .92<LOS<1 — 1<LOS

B. Congestion on US Highways

*Notes:* The figure depicts the US highway network and congestion in 2010. Panel A shows Interstate Highway System in black and the remaining national and state highways in gray. Panel B shows congestion on the Interstate Highway System based on the level of service.



We then use information for each  $\mathcal{S}$  to calculate travel time according to a piecewise function derived from the Highway Capacity Manual (see [National Research Council, 2000](#), Exhibit 23-2):

$$\mathcal{T}_{\mathcal{S}} = \frac{\text{distance}_{\mathcal{S}}}{\mathcal{A}_{\mathcal{S}} \cdot \text{posted speed}_{\mathcal{S}}} \quad , \text{ where } \mathcal{A}_{\mathcal{S}} = \begin{cases} 1.000 & \text{for } 0 \leq \text{LOS}_{\mathcal{S}} < 0.55 \\ 0.950 & \text{for } 0.55 \leq \text{LOS}_{\mathcal{S}} < 0.77 \\ 0.825 & \text{for } 0.77 \leq \text{LOS}_{\mathcal{S}} < 0.92 \\ 0.708 & \text{for } 0.92 \leq \text{LOS}_{\mathcal{S}} < 1.00 \\ 0.600 & \text{for } 1.00 < \text{LOS}_{\mathcal{S}} \end{cases} \quad (1)$$

The travel time function in equation (1) captures positive relationship between travel time and LOS. Whenever a segment,  $\mathcal{S}$ , is characterized by high congestion (or low  $\mathcal{A}_{\mathcal{S}}$ ), it takes longer to complete  $\mathcal{S}$ .

We then can calculate total travel time between any pair of locations  $i$  and  $j$  as follows:

$$\mathbb{T}_{ij} = \sum_{\mathcal{S} \in \mathcal{S}} \mathbb{1}_{\mathcal{S} \in \mathcal{S}_{ij}} \mathcal{T}_{\mathcal{S}} \quad (2)$$

where  $\mathcal{S}$  is a set of all road segments in the network and  $\mathcal{S}_{ij} \subseteq \mathcal{S}$  is a subset that consists of segments used to transport goods from  $i$  to  $j$ ;  $\mathbb{1}_{\mathcal{S} \in \mathcal{S}_{ij}}$  is an indicator function equal to one if  $\mathcal{S}$  is in subset  $\mathcal{S}_{ij}$  or zero otherwise. Note that  $\mathbb{T}_{ij}$  is endogenous to the route choice  $\mathcal{S}_{ij}$  between  $i$  and  $j$  as well as to  $\mathcal{T}_{\mathcal{S}}$ , which is a function of traffic-including trade-generated congestion-on segment  $\mathcal{S}$ . The theoretical model presented in the next section incorporates both of these sources of endogeneity.

### 3 Theoretical Framework

In this section we present a theoretical model of interregional and international trade. The model accommodates multiple countries with potentially many regions. The spatial structure of the model together with the assumption on the movement of goods allows us to use state-level economic outcomes together with intrastate measures of trade costs to calculate the impact on county-level outcomes. This is important for overcoming data limitations. For ease of exposition we present the model using US states (and the District of Columbia) each with multiple counties and all other countries in the world with a single county each. The geographic structure of the model (i.e., counties specifically nested within states) allows us to accommodate rich internal geography within and across US states and reflects the constraints

of available data described in Section 4. Each US county is integrated with all other counties and all countries using county-to-county, county-to-state, and county-to-country travel via US highways, ports, and international shipping lanes. We first present the model in levels and then show how to express the model in relative changes to conduct counterfactuals.

The model in this section extends the standard multisector Ricardian model in two ways. First, the model features two location tiers such that production, consumption, and trade of counties (first tier) can be consistently aggregated to the corresponding state-level variables (second tier). Ultimately, we formulate all county-level variables as functions of their state-level counterparts and trade costs. This allows us to examine economic outcomes at the county level, while keeping the solution of the model computationally feasible and matching the most detailed level of aggregation in the available data. Second, as in [Allen and Arkolakis \(2019\)](#), the model accounts for two sources of endogeneity in trade costs: the choice of transportation routes (including modes) and the effect of trade-generated traffic on congestion.

#### COUNTY-LEVEL PRODUCTION, CONSUMPTION, AND TRADE

We start by describing the supply and demand side in each county  $c$  in state  $i$ . Consumers in county  $c \in i$  allocate their total income across goods from sectors  $s \in S$  to maximize the following utility function:

$$U_i^c = \prod_{s \in S} Q_i^c(s)^{\alpha_i(s)} \quad \text{s.t.} \quad \sum_{s \in S} \alpha_i(s) = 1, \quad (3)$$

where  $\alpha_i(s)$  is Cobb-Douglas consumption share and  $Q_i^c(s)$  is the total quantity consumed of goods from sector  $s$ . Equation (3) leads to the following indirect utility function:

$$V_i^c(s) = \frac{I_i^c}{P_i^c}, \quad \text{where} \quad P_i^c = \prod_{s \in S} \left( \frac{P_i^c(s)}{\alpha_i(s)} \right)^{\alpha_i(s)},$$

where  $I_i^c$  denotes total nominal income of consumers in  $c \in i$  and  $P_i^c(s)$  is the price of one unit of  $Q_i^c(s)$ .

The model features two distinct production levels in each sector  $s$  as in [Costinot, Donaldson and Komunjer \(2012\)](#) and [Caliendo and Parro \(2015\)](#). Varieties in sector  $s$ ,  $z(s)$ , are produced by individual producers and are aggregated into sectoral goods  $Q(s)$  via a CES function prior to intermediate and final consumption. Varieties can be traded across counties (and states) and countries subject to trade costs.

We start with specifying cost minimization outcome of county-level producers of varieties in county  $c \in i$  and sector  $s$ . The cost of an input bundle for those producers is as follows:

$$\kappa_i^c(s) = B_i(s) w_i^{\gamma_i(s)} \left( \prod_{\dot{s} \in S} P_i^c(\dot{s})^{\eta_i(\dot{s}s)} \right)^{1-\gamma_i(s)}, \quad (4)$$

where  $B_i(s)$  is a constant,  $\gamma_i(s)$  is the share of value added, and  $\eta_i^{\dot{s}s}$  is the share of inputs that producers in sector  $s$  source from sector  $\dot{s}$  that reflect input-output linkages across sectors. There is a continuum of varieties  $z(s) \in [0, 1]$  produced in each sector  $s$ . Given the cost of the input bundle in equation (4), a producer located in county  $c \in i$  offers variety  $z(s)$  to state  $j$  at the following price:

$$p_{ij}^{cm^*}(z(s)) = \frac{\kappa_i^c(s) \tau_{ij}^{cm^*}(s)}{z_i^c(s)},$$

where  $\tau_{ij}^{cm^*}(s)$  are total iceberg trade costs between county  $c \in i$  and the aggregation location  $m^*$  in state  $j$  and  $z_i^c(s)$  is the efficiency parameter drawn from a county-sector-specific productivity distribution. As in [Eaton and Kortum \(2002\)](#), efficiency is distributed Fréchet with a location and shape parameters  $T_i(s)K_i^c(s)$  and  $\theta(s)$ , respectively. The specification of the location parameter implies that the average productivity of producers located in county  $c \in i$  can be specified as a product of the relative county productivity within a state,  $K_i^c(s)$ , and a state-level shifter  $T_i(s)$ . Without loss of generality, the mean of  $K_i^c(s)$  across counties within a state and sector is normalized to one.

In each state  $j$  there is a single and unique aggregation county  $m^* \in j$ , which searches for the minimum price for each variety across all potential supplier counties  $c \in i$ . We explain how we identify  $m^*$  for each state  $j$  below. The minimum price of variety  $z(s)$  in that aggregation location is as follows:

$$p_j^{m^*}(z(s)) = \min_{\{i,c\}} \left\{ \frac{\kappa_i^c(s) \tau_{ij}^{cm^*}(s)}{z_i^c(s)} \right\}. \quad (5)$$

As in [Eaton and Kortum \(2002\)](#), varieties  $z(s)$  are aggregated into  $Q(s)$  according to a standard CES aggregator prior to consumption. The probabilistic representation of technologies allows us to specify the price index of  $Q(s)$  in the aggregation county  $m^* \in j$  as:

$$P_j(s) = B_p(s) \left( \sum_i \sum_c T_i(s) K_i^c(s) (\kappa_i^c(s) \tau_{ij}^{cm^*}(s))^{-\theta(s)} \right)^{-\frac{1}{\theta(s)}}, \quad (6)$$

where  $B_p(s)$  is a constant. Since there is a unique  $m^*$  in each state, the CES price index in equation (6) reflects state prices that are available to consumers in  $m^* \in j$ . This also implies that the share of  $m^* \in j$  expenditure in sector  $s$  that it spends on varieties from  $c \in i$  also reflects aggregate trade share of state  $j$ :

$$\pi_{ij}^c(s) = \frac{T_i(s)K_i^c(s) (\kappa_i^c(s)\tau_{ij}^{cm^*}(s))^{-\theta(s)}}{\sum_n \sum_k T_n(s)K_n^k(s) (\kappa_n^k(s)\tau_{nj}^{km^*}(s))^{-\theta(s)}}. \quad (7)$$

Representing county-level outcomes as functions of state-level variables is possible via a consistent aggregation procedure which relies on two assumptions. First, there is a single aggregation county  $m^* \in i$  in each state that aggregates  $z(s)$  into  $Q(s)$ . This CES aggregate can then be consumed in all other counties  $m \in i$  subject to intrastate trade costs. Second, total trade costs between  $c \in i$  and  $m^* \in j$  can be specified as a multiplicative function:

$$\tau_{ij}^{cm^*}(s) = \varepsilon_i^{cc(ij)}(s) \cdot \bar{\tau}_{ij}(s) \cdot \varepsilon_j^{m(ij)m^*}(s), \quad (8)$$

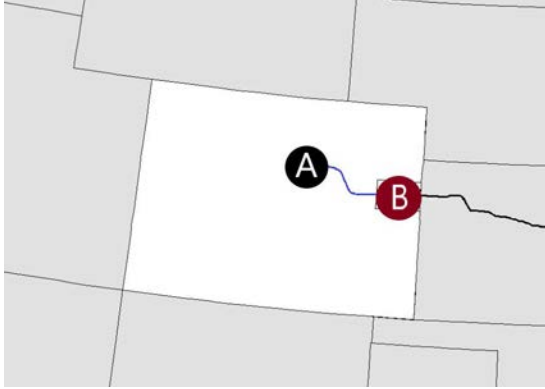
where  $\varepsilon_i^{cc(ij)}(s)$  is exporter intrastate trade cost of transporting goods from county  $c$  to the “exporting” county  $c(ij)$ , which is  $ij$ -pair-specific;  $\varepsilon_j^{m(ij)m^*}(s)$  is importer intrastate trade cost of transporting goods from the “importing” county  $m(ij)$  to the aggregation county  $m^*$ ;  $\bar{\tau}_{ij}(s)$  denotes average interstate trade cost of transporting goods from  $i$  to  $j$ . Note that since every  $ij$  pair is characterized by unique exporting and importing counties,  $\bar{\tau}_{ij}(s)$  is unique and doesn’t require county indices.

Within each state, we assume goods can only be transported via roads. For each  $ij$  state pair, there are unique exporting and importing counties  $c(ij) \in i$  and  $m(ij) \in j$  that are determined endogenously by producers who choose optimal routes to transport goods from  $i$  to  $j$ . We identify  $c(ij) \in i$  and  $m(ij) \in j$  by allowing producers choose the least cost route between the core area  $c^* \in i$  and core area  $m^* \in j$ . These core counties are largest population centers.<sup>6</sup> We illustrate core counties in each state in Figure A1 in the Appendix. When producers in  $c^* \in i$  choose the optimal route to  $m^* \in j$ , they automatically determine  $c(ij) \in i$  and  $m(ij) \in j$  that then apply to all counties in  $i$  and  $j$ . Hence, we do not impose any assumptions about trade routes between the largest economic areas.

For example, consider transporting goods from Colorado to Florida illustrated in Figure 2. The core counties in these two states are Denver County, CO (point A) and Duval

<sup>6</sup>These core counties are typically part of the largest agglomerations within each state, which account for the vast majority of production and trade. For example, Core-Based Statistical Areas accounted for 57% and 64% of total domestic export and imports transported by trucks in 2012.

**Figure 2: Illustration of Intrastate and Interstate Routes**



A. Intrastate route Colorado



B. Interstate route Colorado-Florida

County, FL (point D), respectively. First, the least cost route between A and D is chosen without imposing any restrictions. Second, the optimal route determines the relevant export county (point B) in Colorado and import county (point C) in Florida. Then according to trade costs in equation (8), the intrastate trade cost component for an arbitrary county  $c \in$  Colorado exporting to Florida is  $\varepsilon_i^{cc(ij)}$ . On the other hand, intrastate trade costs in Florida for the core county are  $\varepsilon_j^{m(ij)m^*}$ .<sup>7</sup> We demonstrate how these assumptions match the available county-level data in Section 4.2.

While goods are transported via roads within states, there are several transportation modes available for moving goods between states. In equation (8),  $\bar{\tau}_{ij}(s)$  denotes the average interstate trade cost between states  $i$  and  $j$  across different transportation modes. Let us use  $\tau_{ij}(s)$  to denote trade costs between  $i$  and  $j$  when goods are transported via roads and  $v_{ij}^\ell(s)$  to denote trade costs for mode  $\ell = \{\text{rail, water, air}\}$ . We follow Allen and Arkolakis (2014) and use a discrete choice model across modes to specify average trade costs between  $i$  and  $j$ :

$$\bar{\tau}_{ij}(s) = B_t(s) \left( \tau_{ij}(s)^{-\sigma} + \sum_{\ell} v_{ij}^\ell(s)^{-\sigma} \right)^{-\frac{1}{\sigma}} \quad \text{for } \ell = \{\text{rail, water, air}\}, \quad (9)$$

<sup>7</sup>For concreteness, in the figure, trade between Denver County, CO and Duval County, FL begins by following I-70 East, exiting CO through Kit Carson County (B) and continuing to I-57 South to I-24 East to I-75 South. The route enters Florida in Hamilton County (C) via I-75 South and terminated in Duval County, FL via I-10 East.

where  $\sigma$  governs the elasticity of substitution between transportation modes and  $B_t(s)$  is a constant. The share of total trade in sector  $s$  shipped from  $i$  to  $j$  by road is given by:

$$\varsigma_{ij}(s) = \frac{\tau_{ij}(s)^{-\sigma}}{\bar{\tau}_{ij}(s)^{-\sigma}}. \quad (10)$$

Since our main interest is the implications of changes in highway infrastructure, we focus on changes in  $\bar{\tau}_{ij}$  due to changes in  $\tau_{ij}$ . Hence, while we do account for the possibility of transporting goods between  $i$  and  $j$  via multiple modes, we treat  $v_{ij}^{\text{rail}}$ ,  $v_{ij}^{\text{water}}$ , and  $v_{ij}^{\text{air}}$  as constant.

Next, we aggregate county-level exports to state-to-state and state-to-country trade flows so that county-level variables are expressed as functions of their state-level counterparts and intrastate trade costs. We start by defining two variables:

$$\mu_{ij}^c(s) = \frac{K_i^c(s) \left( \kappa_i^c(s) \varepsilon_i^{cc(ij)}(s) \right)^{-\theta(s)}}{\sum_{c'} K_i^{c'}(s) \left( \kappa_i^{c'}(s) \varepsilon_i^{c'c(ij)}(s) \right)^{-\theta(s)}} \quad \text{and} \quad \kappa_{ij}(s) = \left( \sum_{c'} K_i^{c'}(s) \left( \kappa_i^{c'}(s) \varepsilon_i^{c'c(ij)}(s) \right)^{-\theta} \right)^{-\frac{1}{\theta(s)}},$$

where  $\mu_{ij}^c(s)$  is the share of total exports in sector  $s$  from state  $i$  to  $j$  that comes from county  $c \in i$  and  $\kappa_{ij}(s)$  is the average relative cost of the input bundle faced by producers exporting to state  $j$ . We use the expression for trade shares in equation (7) together with the expressions for  $\mu_{ij}^c(s)$  and  $\kappa_{ij}(s)$  to express total nominal exports from  $c \in i$  to  $j$  as:

$$X_{ij}^c(s) = \mu_{ij}^c(s) \cdot \frac{T_i(s) \left( \kappa_{ij}(s) \bar{\tau}_{ij}(s) \varepsilon_j^{m(ij)m^*}(s) \right)^{-\theta(s)}}{\sum_{i'} T_{i'}(s) \left( \kappa_{i'j}(s) \bar{\tau}_{i'j}(s) \varepsilon_j^{m(i'j)m^*}(s) \right)^{-\theta(s)}} \cdot Y_j(s). \quad (11)$$

Summing  $X_{ij}^c(s)$  over all exporting counties  $c$  and dividing by total absorption of state  $j$  in sector  $s$  allows us to get the expression for state-to-state trade shares as:

$$\pi_{ij}(s) = \frac{T_i(s) \left( \kappa_{ij}(s) \bar{\tau}_{ij}(s) \varepsilon_j^{m(ij)m^*}(s) \right)^{-\theta(s)}}{\sum_{i'} T_{i'}(s) \left( \kappa_{i'j}(s) \bar{\tau}_{i'j}(s) \varepsilon_j^{m(i'j)m^*}(s) \right)^{-\theta(s)}}. \quad (12)$$

Finally, we derive the expression for the CES price indices in county  $m \in j$  taking into account that all varieties are first aggregated in  $m^*$  and then transported to  $m$  subject to

intrastate trade costs  $\varepsilon_j^{m^*m}(s)$ :

$$P_j^m(s) = \varepsilon_j^{m^*m}(s)B(s) \left( \sum_i T_i(s) \left( \kappa_{ij}(s)\bar{\tau}_{ij}(s)\varepsilon_j^{m(ij)m^*}(s) \right)^{-\theta(s)} \right)^{-\frac{1}{\theta(s)}}.$$

This aggregation allows us to express all county-level variables as functions of state-level counterparts and intrastate trade costs. Importantly, this means we can examine county-level outcomes while using state-level data to overcome the absence of county-level data for trade flows and prices. We next turn to describing trade between states and foreign countries.

### INTERNATIONAL TRADE

To describe trade shares between states and countries, we introduce international trade costs. To do this, we account for the fact that US exports and imports can be transported via ports and international sea freight.<sup>8</sup> In addition, we allow trade between US states, Mexico and Canada to be conducted via inland ports along the US-Canada and US-Mexico borders.

Consider international trade between county  $c \in i$  with a foreign country  $n$  via port  $r$  that is located in state  $j$ . With a slight abuse of notation let  $r$  denote the port as well as state and county where it is located. The deterministic part of total trade costs between county  $c \in i$  and country  $n$  via port  $r$  is as follows:

$$\tau_{in}^{c,r}(s) = \underbrace{\varepsilon_i^{cc(ir)}(s)\bar{\tau}_{ir}(s)\varepsilon_r^{m(ir)r}(s)}_{\text{Domestic Component}} \cdot \underbrace{\xi_r^r(s)}_{\text{Port Component}} \cdot \underbrace{t_{rn}^r(s)}_{\text{International Component}},$$

where the domestic component measures the cost of transporting goods via US domestic infrastructure from the production county  $c \in i$  to port  $r$ . The second component,  $\xi_r^r(s)$ , measures port  $r$  efficiency in transporting goods in sector  $s$ . Lastly,  $t_{rn}^r(s)$  measures the cost of transporting goods from port  $r$  to country  $n$ . For tractability, we also assume that there is also a random component of trade costs drawn from an extreme value distribution such that the choice of ports can be modeled as a discrete choice. Then, the share of goods in sector  $s$  that  $n$  consumes from county  $c \in i$  transported via port  $r$  is:

$$\pi_{in}^{c,r}(s) = \frac{T_i(s)K_i^c(s) \left( \kappa_i^c(s)\varepsilon_i^{cc(ir)}(s)\bar{\tau}_{ir}(s)\varepsilon_r^{m(ir)r}(s)\xi_r^r(s)t_{rn}^r(s) \right)^{-\theta(s)}}{\sum_{i'} T_{i'}(s) \left( \kappa_{i'n}(s)\bar{\tau}_{i'n}(s)\varepsilon_n^{m(i'n)m^*}(s) \right)^{-\theta(s)}}.$$

<sup>8</sup>For example, 75 percent of all international freight tons weight traveled by water (US Department of Transportation, 2013).

Summing across all counties  $c \in i$  allows us to derive the share of varieties in sector  $s$  that  $n$  imports from state  $i$  via port  $r$  in  $n$ 's total expenditure on  $s$ -goods:

$$\pi_{in}^r(s) = \lambda_{in}^r(s) \cdot \frac{T_i(s) (\kappa_{in}(s) \bar{\tau}_{in}(s))^{-\theta(s)}}{\sum_{i'} T_{i'}(s) \left( \kappa_{i'n}(s) \bar{\tau}_{i'n}(s) \varepsilon_n^{m(i'n)m^*}(s) \right)^{-\theta(s)}}, \quad (13)$$

where  $(\kappa_{in}(s) \bar{\tau}_{in}(s))^{-\theta(s)} = \sum_r (\kappa_{ir}(s) \bar{\tau}_{ir}(s) \varepsilon_r^{m(ir)r}(s) \xi_r^r(s) t_{rn}^r(s))^{-\theta(s)}$  and  $\lambda_{in}^r(s)$  is the share of goods exported from  $i$  to  $n$  via port  $r$ , which can be specified as follows:

$$\lambda_{in}^r(s) = \frac{\left( \kappa_{ir}(s) \bar{\tau}_{ir}(s) \varepsilon_r^{m(ir)r}(s) \xi_r^r(s) t_{rn}^r(s) \right)^{-\theta(s)}}{\sum_{r'} \left( \kappa_{ir'}(s) \bar{\tau}_{ir'}(s) \varepsilon_{r'}^{m(ir')r'}(s) \xi_{r'}^{r'}(s) t_{r'n}^{r'}(s) \right)^{-\theta(s)}}.$$

Next, we derive port shares for imports from foreign countries to state  $j$ . Consider exports from a foreign country  $n$  to state  $j$  such that the share of goods transported via port  $r$  can be specified as:

$$\lambda_{nj}^r(s) = \frac{\left( t_{nr}^r(s) \xi_r^r(s) \varepsilon_r^{rc(rj)}(s) \bar{\tau}_{rj}(s) \varepsilon_j^{m(rj)m^*}(s) \right)^{-\theta(s)}}{\sum_{r'} \left( t_{nr'}^{r'}(s) \xi_{r'}^{r'}(s) \varepsilon_{r'}^{r'c(r'j)}(s) \bar{\tau}_{r'j}(s) \varepsilon_j^{m(r'j)m^*}(s) \right)^{-\theta(s)}}.$$

Then the share of  $j$ 's total expenditure on  $s$  that it imports from  $n$  via port  $r$  can be specified as:

$$\pi_{nj}^r(s) = \lambda_{nj}^r(s) \cdot \frac{T_n(s) (\kappa_{nj}(s) \bar{\tau}_{nj}(s))^{-\theta(s)}}{\sum_{i'} T_{i'}(s) \left( \kappa_{i'n}(s) \bar{\tau}_{i'n}(s) \varepsilon_n^{m(i'n)m^*}(s) \right)^{-\theta(s)}}, \quad (14)$$

where  $(\kappa_{nj}(s) \bar{\tau}_{nj}(s))^{-\theta(s)} = \sum_r (\kappa_n(s) \varepsilon_j^{m(rj)m^*} \bar{\tau}_{rj}(s) \varepsilon_r^{rc(rj)}(s) \xi_r^r(s) t_{nr}^r(s))^{-\theta(s)}$ . Hence, state-to-country and country-to-state trade shares can be expressed in the same way as state-to-state shares in equation (12).

## CONGESTION AND ENDOGENOUS ROUTE CHOICE

Transporting goods between states in the United States as well as ports for international transshipment by road involves using segments of the available highway network. In the model, producers choose optimal highway routes to domestic and international destinations such that the domestic portion of trade costs reflects travel time via the highway network and takes into account congestion on each road segment. Congestion on each segment, in



turn, depends on trade-generated traffic. Thus, we allow trade costs to endogenously depend on traffic generated by domestic and international trade.

Let  $\mathcal{C}(s)$  denote a single shipment of goods in sector  $s$  so that trade flows between  $i$  and  $j$ —including trade when  $i$  or  $j$  act as a port hub—generate the following interstate traffic:

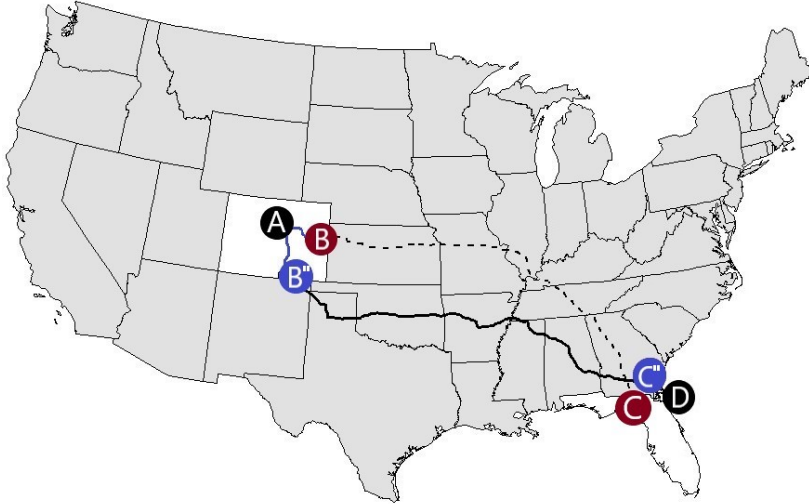
$$M_{ij}(s) = \frac{\pi_{ij}(s)Y_j(s)}{\kappa_{ij}(s)\mathcal{C}(s)} + \sum_n \mathbb{1}_{r=i} \lambda_{nj}^r(s) \frac{\pi_{nj}(s)Y_j(s)}{\kappa_{nj}(s)\mathcal{C}(s)} \quad \text{and} \quad M_{ij} = \sum_s M_{ij}(s), \quad (15)$$

where  $\mathbb{1}_{r \in j}$  is an indicator function that takes value of one whenever port  $r$  is located in state  $i$  and zero otherwise.

Not every segment of the highway network will be affected by total trade generated traffic  $M_{ij}$  but only those actually used when transporting goods between  $i$  and  $j$ . This includes interstate trade as well as transportation of goods between states and ports for international trade. Recall that  $\mathcal{S}$  denotes the set of all available segments  $\mathcal{S} \in \mathcal{S}$  in the road network. Further, let  $\mathcal{S}_{ij}^{\setminus}$  and  $\mathcal{S}_{ij}^{\setminus\setminus}$  denote the best and the second best routes between  $i$  and  $j$ , respectively. The first best route is chosen without any restrictions as long as  $\mathcal{S}_{ij}^{\setminus} \subseteq \mathcal{S}$ . The second best route, however, must be chosen such that  $\mathcal{S}_{ij}^{\setminus\setminus} \cap \mathcal{S}_{ij}^{\setminus} = \emptyset$ , which guarantees that  $\mathcal{S}_{ij}^{\setminus\setminus}$  is unique. In principle, this set up can be extended to more than two routes; however, in our counterfactuals, trade-generated traffic does not create congestion levels that would require using the third best route. This is because in practice the difference in travel time between the first best route, second best route, and the rest is large such that it suffices to examine endogenous traffic allocation between the two best routes for each  $ij$  pair. Hence, in our framework endogenous route choice is a problem of traffic allocation between  $\mathcal{S}_{ij}^{\setminus}$  and  $\mathcal{S}_{ij}^{\setminus\setminus}$ .

To illustrate the difference between  $\mathcal{S}_{ij}^{\setminus}$  and  $\mathcal{S}_{ij}^{\setminus\setminus}$  consider the previous example from Figure 2 that illustrated the first best route,  $\mathcal{S}_{ij}^{\setminus}$ , between Colorado and Florida. As before, the first best route is characterized by A-B-C-D. The second best route between A and D is chosen such that the A-D travel time is minimized subject to the restriction  $\mathcal{S}_{ij}^{\setminus\setminus} \cap \mathcal{S}_{ij}^{\setminus} = \emptyset$ , which states that segments that belong to B-C interstate route cannot be used in the second best route. This restriction is necessary to ensure that B<sup>\setminus\setminus</sup>-C<sup>\setminus\setminus</sup> is unique. The second best route (solid black) between Colorado and Florida is depicted in Figure 3 relative to the first best route (dashed black). The best route,  $\mathcal{S}_{ij}^{\setminus}$ , primarily follows I-70 East exiting Colorado through Kit Carson County and entered Florida in Hamilton County via I-75 South. The second best route,  $\mathcal{S}_{ij}^{\setminus\setminus}$ , begins by heading south on I-25, exiting Colorado via Las Animas County, then follows US-87 South to I-40 East to US-72 East, South on I-65 to US-80 East,

**Figure 3: Example of First Best vs. Second Best Route**



to US-231 South, to US-82 East, to US-23 South, where the route enters Florida in Nassau County, terminating in Duval County. The second best route in this case takes approximately 30 hours to complete compared to just under 27 hours for the best route.

As we have illustrated above, total trade-generated traffic  $M_{ij}$  can potentially go through different routes. Let  $h_{ij}$  denote the share of total traffic that uses segments in  $\mathcal{S}_{ij}^{\setminus}$ , then total traffic generated by trade in all sectors and across all locations that is relevant for segment  $\mathcal{S}$  can be expressed as:

$$N_{\mathcal{S}} = \begin{cases} \sum_{\mathcal{S}} \mathbf{1}_{\mathcal{S} \in \mathcal{S}_{ij}^{\setminus}} h_{ij} M_{ij} & \text{for } \mathcal{S} \in \mathcal{S}_{ij}^{\setminus} \\ \sum_{\mathcal{S}} \mathbf{1}_{\mathcal{S} \in \mathcal{S}_{ij}^{\setminus}} (1 - h_{ij}) M_{ij} & \text{for } \mathcal{S} \in \mathcal{S}_{ij}^{\setminus} \end{cases} \quad (16)$$

Further, we parameterize travel time,  $\mathcal{T}_{\mathcal{S}}$ , introduced in Section 2 as  $\mathcal{T}(N_{\mathcal{S}}, F_{\mathcal{S}})$  which measures travel time on segment  $\mathcal{S}$  and depends on trade generated congestion,  $N_{\mathcal{S}}$ , conditional on the segment fundamentals,  $F_{\mathcal{S}}$ . The latter includes distance, speed, and traffic generated by sources other than trade. Let us start by defining two benchmarks that pin down minimum travel time between  $i$  and  $j$  for each of the two routes. The minimum travel time is achieved when there is zero trade generated congestion conditional on other fundamentals and non-trade-related traffic:

$$\min(\mathcal{T}_{ij}^{\setminus}) = \sum_{\mathcal{S} \in \mathcal{S}} \mathbf{1}_{\mathcal{S} \in \mathcal{S}_{ij}^{\setminus}} \mathcal{T}(0, F_{\mathcal{S}}) \quad \text{and} \quad \min(\mathcal{T}_{ij}^{\setminus}) = \sum_{\mathcal{S} \in \mathcal{S}} \mathbf{1}_{\mathcal{S} \in \mathcal{S}_{ij}^{\setminus}} \mathcal{T}(0, F_{\mathcal{S}}), \quad (17)$$

By definition conditional on zero traffic from trade, it is always less costly to travel via  $\mathbf{S}_{ij}^{\setminus}$  than  $\mathbf{S}_{ij}^{\setminus\setminus}$ . However, when  $M_{ij} > 0$  allocation of traffic between the two routes is endogenous. At first trade-generated traffic uses the first best route. However, the capacity of the first best route is capped at  $\min(\mathbf{T}_{ij}^{\setminus\setminus})$ . If and when trade-generated traffic creates enough additional congestion on the first best route such that the travel time  $\mathbf{T}_{ij}^{\setminus}$  reaches the capacity limit  $\min(\mathbf{T}_{ij}^{\setminus\setminus})$ , the marginal truck becomes indifferent between staying on the first best route and switching to the second best route. All trade traffic beyond that level will be using the second best route. Hence, the share of traffic  $h_{ij}$  is determined as:

$$h_{ij} = \min(1, h_{ij}^*) \quad \text{subject to} \quad \sum_{\mathcal{S} \in \mathcal{S}} \mathbf{1}_{\mathcal{S} \in \mathbf{S}_{ij}^{\setminus}} \mathcal{T} \left( \sum_{ij} \mathbf{1}_{\mathcal{S} \in \mathbf{S}_{ij}^{\setminus}} h_{ij}^* M_{ij}, F_{\mathcal{S}} \right) = \min(\mathbf{T}_{ij}^{\setminus\setminus}). \quad (18)$$

The intuition behind equation (18) is straightforward. All trucks will use route  $\mathbf{S}_{ij}^{\setminus}$  such that  $h_{ij} = 1$  if the travel time of the last truck on this route is below the minimum travel time on the next best route. In this case, congestion would have an effect on trade costs between  $i$  and  $j$  solely by increasing travel time on the first best route. However, the share  $h_{ij} < 1$  if the capacity of route  $\mathbf{S}_{ij}^{\setminus}$  is reached. In this case,  $h_{ij}^*$  is implicitly pinned down by the second equality in equation (18). Then the travel time between  $i$  and  $j$  for an average truck can be expressed as:

$$\mathbf{T}_{ij} = h_{ij} \sum_{\mathcal{S} \in \mathcal{S}} \mathbf{1}_{\mathcal{S} \in \mathbf{S}_{ij}^{\setminus}} \mathcal{T} \left( \sum_{ij} \mathbf{1}_{\mathcal{S} \in \mathbf{S}_{ij}^{\setminus}} h_{ij} M_{ij}, F_{\mathcal{S}} \right) + (1 - h_{ij}) \sum_{\mathcal{S} \in \mathcal{S}} \mathbf{1}_{\mathcal{S} \in \mathbf{S}_{ij}^{\setminus\setminus}} \mathcal{T} \left( \sum_{ij} \mathbf{1}_{\mathcal{S} \in \mathbf{S}_{ij}^{\setminus\setminus}} (1 - h_{ij}) M_{ij}, F_{\mathcal{S}} \right) \quad (19)$$

The expression for travel time in equation (19) accounts for two channels of how trade-generated traffic affects trade costs between  $i$  and  $j$ . First,  $M_{ij}$  increases congestion levels  $N_{\mathcal{S}}$  on relevant segments. Second, depending on the traffic patterns related to trade and road capacity, a positive share of trucks may choose to travel via the second best route. As we have emphasized before, this framework can be extended to potentially more routes but it is sufficient to consider  $\mathbf{S}_{ij}^{\setminus}$  and  $\mathbf{S}_{ij}^{\setminus\setminus}$  for the purposes of this paper.

## LABOR MOBILITY

Our specification of labor mobility in the model is standard and largely follows [Anderson \(2011\)](#).<sup>9</sup> Labor is mobile across counties in the United States subject to migration costs. Workers choose where to live to maximize indirect utility,  $V_i^c$ , across all possible counties

<sup>9</sup>See [Allen and Arkolakis \(2014\)](#) for an alternative approach to modeling trade in the presence of labor or factor mobility.

subject to migration costs. Each county has an initial stock of labor  $L_i^c$  and workers choose to migrate to  $m \in j$  if the following holds:

$$(V_j^m \delta_{ij}^{cm}) \epsilon > V_i^c,$$

where  $\delta_{ij}^{cm} \in (0, 1)$  is the deterministic component of migration costs and  $\epsilon$  is a random component drawn from an extreme value distribution. The share of workers that migrate from  $c \in i$  to  $m \in j$  can then be written as follows:

$$\omega_{ij}^{cm} = \frac{(V_j^m \delta_{ij}^{cm})^{\frac{1}{\varrho}}}{\sum_{k,n} (V_n^k \delta_{in}^{ck})^{\frac{1}{\varrho}}}. \quad (20)$$

Note that  $\varrho > 0$  governs the degree of labor mobility across counties. In our quantitative exercise, we consider a range of values for  $\varrho$  above 1 consistent with the literature. For example, [Caliendo et al. \(2019\)](#) estimate  $\varrho = 2$ . Given migration flows, total labor in each county and state is given by:

$$L_i^c = \sum_{k,n} \omega_{ni}^{kc} L_n^k \text{ and } L_i = \sum_c L_i^c. \quad (21)$$

### TRADE BALANCE AND EQUILIBRIUM

Total expenditures of state  $i$  on goods produced in sector  $s$  is the combination of demand for final and intermediate goods. Nominal wages are determined at the state level and are equal across all counties  $c \in i$  such that the total expenditure can be expressed as follows:

$$Y_i(s) = \sum_{\dot{s}} (1 - \gamma_i(s)) \eta_i(s\dot{s}) \sum_j \pi_{ij}(s) Y_j(s) + \alpha_i(s) (I_i + D_i), \quad (22)$$

where  $I_i = \sum_c I_i^c \equiv \sum_c L_i^c w_i$  and  $D_i$  is an exogenous deficit constant. Given  $Y_i(s)$ , we specify the trade balance condition:

$$\sum_s \sum_n \pi_{ni}(s) Y_i(s) - D_i = \sum_s \sum_n \pi_{in}(s) Y_n(s), \quad (23)$$

which given a numeraire determines wages in all states and countries. This completes the description of the model and allows us to formally define the equilibrium conditions. Let us use  $\mathbb{V}$  to denote the following parameters  $\{\alpha_i(s), \gamma_i(s), \eta_i(s\dot{s}), \theta(s), \sigma, \varrho\}$ .

**Definition 1:** Given primitives  $\mathbb{V}$ ,  $K_i^c$ ,  $T_i(s)$ ,  $\xi_i^r(s)$ ,  $L_i^c$ ,  $D_i$ ,  $\delta_{ij}^{cm}$  and trade costs structure  $\varepsilon_{ij}^{cc(ij)}(s)$ ,  $\mathbf{S}_{ij}^\lambda$ ,  $\mathbf{S}_{ij}^\omega$ ,  $F_{\mathcal{S}}$ ,  $\tau_{ij}(s)$ , and  $v_{ij}^\ell(s)$ , an equilibrium is a vector of wages,  $\mathbf{w} \in \mathbb{R}_+$ , and prices,  $\{P_i^c(s)\}$ , such that the conditions in (4), (6), (7), (9), (11), (12), (13), (14), (15), (16), (17), (18), (19), (20), (21), (22), (23) are satisfied for all  $c$ ,  $i$ ,  $s$  and  $\mathcal{S}$ .

### 3.1 Counterfactual Equilibrium in Relative Changes

In our counterfactual exercises, we examine the effects of changes in domestic and international trade costs. To do this, it is useful to express the model in relative changes. For convenience, we define the following identity for an arbitrary variable  $a$ :

$$\hat{a} = \frac{a'}{a},$$

where  $a'$  and  $\hat{a}$  denote the counterfactual value of  $a$  and the change relative to its benchmark value, respectively. To calculate counterfactual outcomes we use the *hat algebra* approach. This approach has been used for counterfactual analysis in the context of international trade, e.g., Dekle, Eaton and Kortum (2007) and Caliendo and Parro (2015). More recently, Allen and Arkolakis (2019) show how to apply this approach in models featuring endogenous congestion and route choice. Our solution method is consistent with their approach and uses the observed allocations of domestic and international trade shares, labor and traffic as a way to sidestep the challenge of solving for the unobservable fundamentals of the economies represented with the large number of interacting locations and sectors. This is particularly important for our setting in which we focus on the impact of changes in domestic trade costs for all US counties and all foreign trading partners. As it turns out, we can examine county-level outcomes using the observed state-level allocations together with the data on intrastate trade costs.

We start by calculating the counterfactual changes in trade costs relative to the benchmark equilibrium. In particular, we remove certain parts of the highway system in the United States, e.g., the entire Interstate Highway System or individual segments (I-5, I-10, etc) so that producers and consumers are presented with a subset of segments available in the benchmark,  $\mathbf{S}' \subseteq \mathbf{S}$ . We then calculate counterfactual outcomes with a specific focus on the county-specific and aggregate welfare changes generated by the exogenous changes in the available road network. Note that even after removing the entire IHS, the remaining federal-aid (approximately 133,00 miles) and state (approximately 213,00 miles) highways are available to move between any  $i$  and  $j$ , although subject to higher trade costs. Hence, our counterfactuals will calculate the marginal value of the IHS or individual segments con-

ditional on the remaining highway network (and other modes).

Given the new set  $\mathbf{S}'$  and fundamental characteristics of each segment, producers choose optimal routes to minimize trade costs between states and between states and ports such that we observe counterfactual  $\mathbf{S}_{ij}' \subseteq \mathbf{S}'$  and  $\mathbf{S}_{ij}'' \subseteq \mathbf{S}'$ . Conditional on the chosen routes and counterfactual level of trade-generated congestion (discussed below), we calculate counterfactual changes in interstate highway trade costs as:

$$(i) \quad \widehat{\tau}_{ij}(s) = \frac{\rho_s(\mathbf{T}'_{ij})}{\rho_s(\mathbf{T}_{ij})},$$

where  $\rho_s(\cdot)$  is a sector-specific function that translates travel time from  $i$  to  $j$  in sector  $s$  trade costs. We specify the exact functional form of  $\rho_s(\cdot)$  in the next section. Next, conditional on costs of transporting goods via rail, water and air staying constant, we can specify changes in the average trade costs between  $i$  and  $j$  as follows:

$$(ii) \quad \widehat{\tau}_{ij}(s) = \left( \varsigma_{ij}(s) \widehat{\tau}_{ij}(s)^{-\sigma} + [1 - \varsigma_{ij}(s)] \right)^{-\frac{1}{\sigma}},$$

where  $\varsigma_{ij}(s)$  are observed in the data. We keep sea trade costs from each port  $r$  to each country  $n$  as well as port efficiency constant such that  $\widehat{t}_{jn}(s) = 1$  and  $\widehat{\xi}_j^r(s) = 1$ . This allows us to calculate counterfactual changes in sectoral average production costs gross of trade costs between states and foreign countries and vice versa as:

$$(iii) \quad \left( \widehat{\kappa}_{in}(s) \widehat{\tau}_{in}(s) \right)^{-\theta(s)} = \sum_r \lambda_{in}^r(s) \left( \widehat{\kappa}_{ir}(s) \widehat{\tau}_{ir}(s) \widehat{\varepsilon}_r^{m(ir)r}(s) \right)^{-\theta(s)},$$

$$(iv) \quad \left( \widehat{\kappa}_{nj}(s) \widehat{\tau}_{nj}(s) \right)^{-\theta(s)} = \sum_r \lambda_{nj}^r(s) \left( \widehat{\kappa}_n(s) \widehat{\varepsilon}_j^{m(rj)m^*} \widehat{\tau}_{rj}(s) \widehat{\varepsilon}_r^{c(rj)}(s) \right)^{-\theta(s)}.$$

Hence, given the counterfactual road network  $\mathbf{S}' \subseteq \mathbf{S}$ , the corresponding counterfactual changes in intrastate, inter-state and international trade costs are specified in (i) – (iv). Changes in trade costs together with the observations on  $\{\mu_{ij}^c(s), \pi_{ij}(s), \lambda_{ij}^r(s), \omega_{ij}^{cm}, L_i^c, M_{ij}(s)\}$  in the initial equilibrium allow us to characterize counterfactual equilibrium by the following conditions:

$$(v) \quad \text{Changes in county-level costs: } \widehat{\kappa}_i^c(s) = \widehat{w}_i^{\gamma_i(s)} \left( \prod_{\dot{s} \in S} \widehat{P}_i^c(\dot{s}) \eta_i(\dot{s}s) \right)^{1-\gamma_i(s)}.$$

$$(vi) \quad \text{Changes in state-level costs: } \widehat{\kappa}_{ij}(s) = \left( \sum_c \mu_{ij}^c \left( \widehat{\kappa}_i^c(s) \widehat{\varepsilon}_i^{cc(ij)}(s) \right)^{-\theta(s)} \right)^{-\frac{1}{\theta(s)}}.$$

$$(vii) \text{ Counterfactual trade shares: } \pi_{ij}(s)' = \frac{\pi_{ij}(s) \left( \widehat{\kappa}_{ij}(s) \widehat{\tau}_{ij}(s) \widehat{\varepsilon}_j^{m(ij)m^*}(s) \right)^{-\theta(s)}}{\sum_{i'} \pi_{i'j}(s) \left( \widehat{\kappa}_{i'j}(s) \widehat{\tau}_{i'j}(s) \widehat{\varepsilon}_j^{m(i'j)m^*}(s) \right)^{-\theta(s)}}.$$

$$(viii) \text{ Changes in state-level prices: } \widehat{P}_j(s) = \left( \sum_{i'} \pi_{i'j}(s) \left( \widehat{\kappa}_{i'j}(s) \widehat{\tau}_{i'j}(s) \widehat{\varepsilon}_j^{m(i'j)m^*}(s) \right)^{-\theta(s)} \right)^{-\frac{1}{\theta(s)}}.$$

$$(ix) \text{ Changes in county-level prices and real wages: } \widehat{P}_i^c(s) = \widehat{\varepsilon}_i^{m^*c}(s) \widehat{P}_i(s) \text{ and } \widehat{V}_i^c = \widehat{w}_i / \widehat{P}_i^c.$$

$$(x) \text{ Counterfactual migration shares: } \omega_{ij}^{cm'} = \frac{\omega_{ij}^{cm} (\widehat{V}_j^m)^{\frac{1}{\theta}}}{\sum_{k,n} \omega_{in}^{ck} (\widehat{V}_n^k)^{\frac{1}{\theta}}}.$$

$$(xi) \text{ Counterfactual labor force: } L_i^{c'} = \sum_{k,n} \omega_{ni}^{kc'} L_n^{k'} \text{ and } L_i' = \sum_c L_i^{c'}.$$

$$(xii) \text{ Counterfactual state nominal income: } I_i' = L_i \widehat{L}_i w_i \widehat{w}_i.$$

$$(xiii) \text{ Counterfactual absorption: } Y_i(s)' = \sum_s (1 - \gamma_i(s)) \eta_i(ss) \sum_j \pi_{ij}(s)' Y_j(s)' + \alpha_i(s) (I_i' + D_i).$$

$$(xiv) \text{ Counterfactual port shares: } \lambda_{in}^r(s)' = \frac{\lambda_{in}^r(s) \left( \widehat{\kappa}_{ir}(s) \widehat{\tau}_{ir}(s) \widehat{\varepsilon}_r^{m(ir)r}(s) \right)^{-\theta(s)}}{\sum_{r'} \lambda_{in}^{r'}(s) \left( \widehat{\kappa}_{ir'}(s) \widehat{\tau}_{ir'}(s) \widehat{\varepsilon}_{r'}^{m(ir')r'}(s) \right)^{-\theta(s)}}.$$

$$(xv) \text{ Counterfactual state wages: } \sum_s \sum_j \pi_{ji}(s)' Y_i(s)' - D_i = \sum_s \sum_j \pi_{ij}(s)' Y_j(s)'.$$

$$(xvi) \text{ Counterfactual traffic: } M_{ij}(s)' = \frac{\widehat{\pi}_{ij}(s) \widehat{Y}_j(s)}{\widehat{\kappa}_{ij}(s)} M_{ij}(s) + \sum_n \mathbf{1}_{r=i} \lambda_{nj}^r(s)' \frac{\widehat{\pi}_{nj}(s) \widehat{Y}_j(s)}{\widehat{\kappa}_{nj}(s)} M_{nj}(s).$$

Lastly, we determine traffic allocation shares:

$$(xvii) \sum_{S \in S'} \mathbf{1}_{S \in S_{ij}^v} \mathcal{T} \left( \sum_{ij} \mathbf{1}_{S \in S_{ij}^v} h_{ij}^* M_{ij}', F_S \right) = \min(T_{ij}^{\setminus'}) \text{ and } h_{ij}' = \min(1, h_{ij}^*),$$

and counterfactual travel time that accounts for endogenous route choice:

$$(xviii) T_{ij}' = h_{ij}' \sum_{S \in S'} \mathbf{1}_{S \in S_{ij}^v} \mathcal{T} \left( \sum_{ij} \mathbf{1}_{S \in S_{ij}^v} h_{ij}' M_{ij}', F_S \right) + (1 - h_{ij}') \sum_{S \in S'} \mathbf{1}_{S \in S_{ij}^w} \mathcal{T} \left( \sum_{ij} \mathbf{1}_{S \in S_{ij}^w} (1 - h_{ij}') M_{ij}', F_S \right)$$

Given the interstate traffic shares we calculate counterfactual changes in the intrastate trade

costs:

$$(xix) \quad \tilde{\varepsilon}_i^{cc(ij)}(s) = \left( \frac{\varepsilon_i^{cc'(ij)}(s)'}{\varepsilon_i^{cc'(ij)}(s)} \right)^{h_{ij}} \left( \frac{\varepsilon_i^{cc''(ij)}(s)'}{\varepsilon_i^{cc''(ij)}(s)} \right)^{1-h_{ij}},$$

where all  $c'(ij)$  and  $c''(ij)$  are export counties for the first best and second best routes, respectively. The functional form for the condition above is the result of trade costs parameterization as a function of travel time, which we discuss in the next section.

Hence, given the structure of counterfactual trade costs, the counterfactual equilibrium is a vector of counterfactual wages and prices such that the system defined by the conditions above is satisfied for all  $c$ ,  $i$ ,  $s$  and  $\mathcal{S}$ .

## 4 Data and Estimation

Solving the model and conducting counterfactuals requires information on trade flows, value-added, employment, migration, consumption shares, and input-output linkages in the benchmark equilibrium. Crucially, we also need information on trade costs among counties in the United States as well as between US counties and foreign countries. This section describes the underlying data and estimation. We provide additional information on data construction and sources in the Data Appendix. The benchmark year for all variables is 2012 unless noted otherwise.

We calibrate the model with data on 3,106 counties in the United States including the District of Columbia, but excluding Alaska and Hawaii. For computational purposes, we group counties into states according to conventional US states boundaries except for California and Texas. Due to their size and geographic shapes, we group counties in California into North and South agglomerations and counties in Texas into East, West, and North agglomerations. This means that 3,106 counties are grouped in 52 state areas. Our results are robust to not dividing California and Texas into smaller areas.

We also include 35 other countries and an aggregate that combines data for the rest of the world.<sup>10</sup> In terms of the sectoral coverage, we consider 22 sectors including 12 manu-

<sup>10</sup>The countries included are Australia, Austria, Belgium, Brazil, Canada, China, Cyprus, Czech Republic, Germany, Denmark, Estonia, Finland, France, Great Britain, Greece, Hungary, Indonesia, India, Ireland, Italy, Japan, Korea, Lithuania, Mexico, the Netherlands, Poland, Portugal, Romania, Russia, Slovakia, Slovenia, Spain, Sweden, Turkey, and Taiwan.



facturing sectors, 8 service sectors, construction, and combined wholesale and retail trade.<sup>11</sup> The list of sectors is available in the Appendix. To solve for counterfactual equilibria, we need parameters in  $\mathbb{V}$  and the values of  $\{\varsigma_{ij}(s), \mu_{ij}^c(s), \pi_{ij}(s), \lambda_{ij}^r(s), \omega_{ij}^{cm}, L_i^c, M_{ij}(s)\}$  in the benchmark equilibrium. We next describe how we obtain the benchmark values of these parameters and variables.

### PRODUCTION AND CONSUMPTION SHARES

To construct the value added shares, intermediate input shares, and Cobb-Douglas consumption shares, we use data from the County Business Patterns and World Input-Output Database in 2012. We calculate the value-added share in sector  $s$  as the ratio of value-added to output, which corresponds to  $\gamma_i(s)$ ; we calculate the consumption share as the fraction of final consumption in sector  $s$ , which gives  $\alpha_i(s)$ ; and we calculate the intermediate input shares as the fraction of the intermediate input usage of sector  $\dot{s}$  sourced from sector  $s$ , which is  $\eta_i(s\dot{s})$ . In each case, the parameters are specific to a state or country  $i$ . Due to data availability,  $\gamma_i(s)$  vary by state but  $\alpha_i(s)$  and  $\eta_i(s\dot{s})$  are homogeneous for all states in the United States.

### TRANSPORTATION MODE SHARES

To account for substitution across different transportation modes in counterfactual equilibria, we need initial transportation modes shares,  $\varsigma_{ij}(s)$ , and the elasticity parameter  $\sigma$ . We adopt the value of  $\sigma = 14.2$  from [Allen and Arkolakis \(2014\)](#). To calculate transportation modes shares across states, we use data from the 2012 Commodity Flow Survey. We aggregate sectoral trade flows to state level by summing flows transported via a single mode in four categories: truck (*for-hire truck* and *private truck*), rail, water (*inland water*, *Grate Lakes*, *deep sea*, and *multiple waterways*), and air, which together allow us to calculate  $\varsigma_{ij}(s)$ .

### DOMESTIC AND INTERNATIONAL TRADE FLOWS

We require data on domestic trade flows between US states and international trade flows between US states and foreign countries. Domestic trade flows are taken from the 2012 Commodity Flow Survey and state-to-country and country-to-state trade flows for 2012 are downloaded from the Census Bureau’s USA Trade online portal. Domestic trade flows are observed at the level of Core Based Statistical Areas, which allows us to calculate domestic

---

<sup>11</sup>The Data Appendix describes the mapping from the sectors listed in the North American Industry Classification System (NAICS) or World Input-Output Database (WIOD) to the 22 sectors we consider in this paper.

export and import flows for two agglomeration areas in California and three agglomeration areas in Texas. In addition, we combine foreign export and import flows from USA Trade at the district level with the data on international trade flows from the World Input Output Database to calculate foreign trade for California and Texas. In addition, we use information on trade flows between US states and foreign countries through US ports, which are also drawn from the Commodity Flow Survey and USA Trade for 2012. Data on international trade flows between foreign countries comes from the World Input Output Database.

## COUNTY EMPLOYMENT, OUTPUT, AND MIGRATION

Data on employment and payroll at the county level are drawn from the County Business Patterns in 2012. We match the initial value of the county-level employment to  $L_i^c$ . To calculate the initial values of  $\mu_{ij}^c$ , we combine Commodity Flow Survey data which reports sectoral trade flows at the level of Core Based Statistical Areas and the data on annual payroll at the county level as follows:

$$\mu_{ij}^c(s) = \mu_{ij}^{CBS}(s) \cdot \mu^{c,CBS}(s),$$

where  $\mu_{ij}^{CBS}(s)$  is the share of CBS area in state  $i$ 's total exports to  $j$  and  $\mu^{c,CBS}(s)$  is the share of county  $c$  in CBS area's total value added in sector  $s$ .<sup>12</sup> Migration flows between US counties are constructed from Internal Revenue Service data for 2011-2012. In particular, this data aggregates information on the county of residence in 2011 and 2012 from individual tax returns, which we use to calculate  $\omega_{ij}^{cm}$  from the model. Hence, the stock of labor in  $c \in i$  is the sum of all workers across all destinations  $m \in j$  that resided in  $c$  prior to 2012.

### 4.1 Constructing Trade Costs

The starting point for constructing trade costs is detailed information on the US highway network shown in Figure 1. We use the highway network, domestic navigable waterways, international shipping lanes, and trade flows to construct the domestic and international trade cost components used to calibrate the model.

<sup>12</sup>The concordance between Core Based Statistical Areas and counties is available at <https://www.nber.org/research/data/census-core-based-statistical-area-cbsa-federal-information-processing-series-fips-county-crosswalk>.

## DATA FOR DOMESTIC AND INTERNATIONAL TRADE COSTS

The key inputs into the domestic trade cost components are the travel time and distance between US county pairs as well as the travel time and distance between US counties and US ports. To construct these inputs we represent each location as the geographic centroid or centroid of the county in which a port is located. Each county or port centroid is connected to the US highway network via an access road network. Each of the 331,074 pieces of the highway network is assigned a speed based on its classification, specifically, we assign 70, 55, and 45 miles per hour to the components of the Interstate Highway System, US highways, and state highways, respectively, and 10 miles per hour to the access road network.<sup>13</sup> Next we use the highway network to identify the routes and corresponding travel time underlying the domestic trade costs by road, including interstate,  $\tau_{ij}(s)$ , and intrastate,  $\varepsilon_{ij}^{cm}$  trade costs.

To calculate the international trade costs we use information on the location of 21 US ports, domestic navigable waterways, and international shipping lanes between US ports and 35 foreign trading partners. US ports are shown in Panel A of Figure 4 and shipping lanes are presented in Panel B of Figure 4. Using this data we calculate the minimum distance route between each port and country. This corresponds to the international trade cost component  $t_{ik}$ . The combined domestic and international trade cost components can be used to construct the trade costs between any pair of locations in the model.

## ESTIMATION OF TRADE ELASTICITY PARAMETER

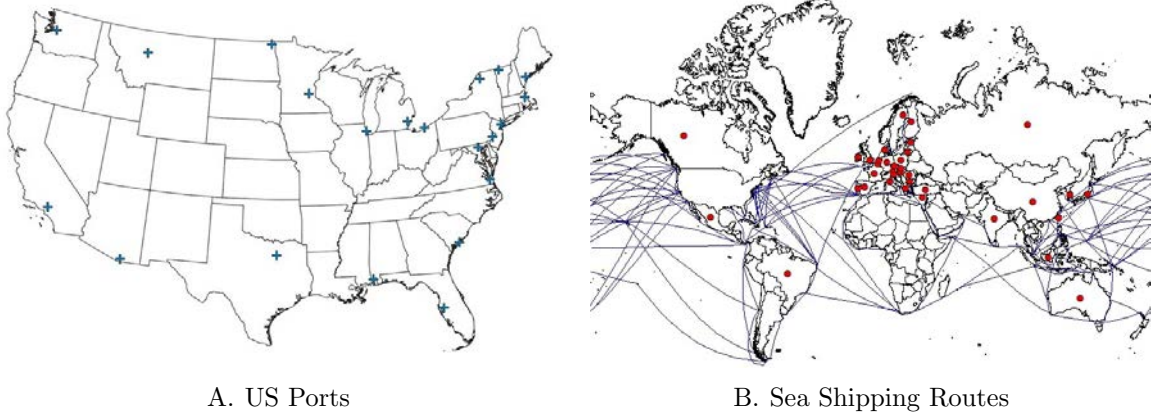
A key input for constructing trade costs and performing quantitative analysis is the set of parameters governing the dispersion of productivity within sectors,  $\theta(s)$ . Importantly, values for these parameters determine the elasticity of trade flows with respect to trade costs. Following [Head and Ries \(2001\)](#), let us establish the following identify for an arbitrary variable  $a_{ij}$  that varies by exporter  $i$  and importer  $j$ :

$$\tilde{a}_{ij} = \frac{a_{ij} a_{ji}}{a_{ii} a_{jj}}.$$

---

<sup>13</sup>To find the county centroids we overlay shapefiles for county boundaries in 2012 using shapefiles from the [US Census Bureau \(2012\)](#) and identify the geographic centroid. [Jaworski and Kitchens \(2019\)](#) find that using population-weighted county centroids or assigning alternate speeds to the components of the highway network does not lead to substantially different results.

**Figure 4: US Ports and International Sea Shipping Routes**



*Notes:* This figure shows the portions of the transportation network that contribute to international trade costs. Panel A shows the location of US ports. Panel B shows the international shipping lanes and country centroids.

This Head-Ries decomposition allows us to eliminate  $i$ -specific and  $j$ -specific components in  $a_{ij}$  such that the structural gravity equation in equation (12) can be reformulated as:

$$\ln(\tilde{\pi}_{ij}(s)) = -\theta(s) \ln(\tilde{\tau}\varepsilon_{ij}(s)) + \tilde{Z}_{ij}(s)\Theta^s + \epsilon_{ij}, \quad (24)$$

where  $\tilde{\tau}\varepsilon_{ij}(s)$  is a measure of total trade costs between  $i$  and  $j$ , which we describe below, and  $\tilde{Z}_{ij}(s)$  includes  $\tilde{\kappa}_{ij}(s)$  and  $\tilde{\zeta}_{ij}(s)$ ; <sup>14</sup>  $\epsilon_{ij}$  is a stochastic error term. To estimate  $\theta(s)$  in equation (24), we need observable proxy measures of trade costs. We construct such measures of  $\tau\varepsilon_{ij}(s)$  for the purpose of estimating  $\theta(s)$  following specification suggested in [Combes and Lafourcade \(2005\)](#). We combine information on the time and distance of moving goods between state  $i$  and the core county in state  $j$ .<sup>15</sup> In particular, we use the labor cost determined by the hourly wage of a truck driver averaged between origin  $i$  and destination  $j$  and the fuel cost based on the price of fuel per gallon together with fuel usage per mile to calculate:

$$\tau\varepsilon_{ij}(s) = 1 + \frac{\text{hours}_{ij} \cdot \text{wage per hour}_{ij} + \text{miles}_{ij} \cdot \text{cost per mile}}{\text{average value of shipment in sector } s} \quad (25)$$

<sup>14</sup>We obtain the estimates of  $\kappa_{ij}(s)^{\theta(s)}$  as fitted values of  $ij$ -specific fixed effects,  $f_{ij}(s)$ , in:

$$\ln(\mu_{ij}^c(s)) = f_c(s) + f_{ij}(s) + \text{error}_{ij},$$

which is a stochastic version of the structural equation for  $\mu_{ij}^c(s)$ .

<sup>15</sup>For estimation of  $\theta(s)$  and  $\varphi(s)$  (below), we do not impute trade flows for separate agglomerations in CA and TX but rather rely on aggregate flows for these two states.

**Table 1: Estimates of  $\theta(s)$** 

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\theta(s)$	7.8	4.4	5.5	14.2	10.5	7.8	3.3	7.4	6.6	6.3	11.5	3.6
s.e.	(0.77)	(0.72)	(0.44)	(2.15)	(1.46)	(0.59)	(0.43)	(0.75)	(1.31)	(1.36)	(2.26)	(0.44)
Observations	1,818	1,502	1,898	768	1,912	1,870	1,248	1,980	1,806	1,708	1,340	1,878

*Notes:* The table shows estimates of  $\theta(s)$  for each sector  $s$ . The dependent variable is the [Head and Ries \(2001\)](#) transformation of trade shares referred to in the text. Standard errors clustered on origin and destination state are reported in parentheses. Column 1 is Food, Beverage, and Tobacco, Column 2 is Textiles and Leather, Column 3 is Wood, Paper, and Printing, Column 4 is Petroleum and Coal, Column 5 is Chemicals, Column 6 is Plastics and Rubber, Column 7 is Nonmetallic Minerals, Column 8 is Primary and Fabricated Metals, Column 9 is Machinery, Column 10 is Computers, Electronics, and Electrical, Column 11 is Transportation Equipment, and Column 12 is Furniture and Miscellaneous. The number of observations in each column reflect the number of state origin-destination pairs with non-zero trade flows.

where the denominator is the average value of a shipment in sector  $s$  taken from the Commodity Flow Survey in 2012. The hourly wages for truck drivers are from the Bureau of Labor Statistics and the data on fuel cost per mile are calculated from the decennial census ([Ruggles, Alexander, Genadek, Goeken, Schroeder, Sobek et al., 2010](#)) and [US Census Bureau \(2010\)](#).

Table 1 shows the results of estimating  $\theta(s)$  for each of the 12 manufacturing sectors. Two-way clustered standard errors on states  $i$  and  $j$  are reported in parentheses. The results reveal substantial variation across sectors: the estimates of  $\theta(s)$  range between 3.3 for 14.2 and are statistically significant at the 1 percent level. The magnitudes are consistent with the existing estimates in the literature, e.g., [Eaton and Kortum \(2002\)](#) and [Caliendo and Parro \(2015\)](#) who estimate average manufacturing  $\theta$  to be 8.3 and 6.5, respectively. Our approach is complementary to approaches used in the international trade literature. For example, [Caliendo and Parro \(2015\)](#) use data on international trade flows and exploit variation in tariffs to estimate  $\theta(s)$ .<sup>16</sup> Finally, we assign the average value of these estimates to the sectors where trade flow data is not available.

## PARAMETERIZING DOMESTIC TRADE COSTS AND CONGESTION

We parameterize intrastate and interstate domestic trade costs as a function of shipping time via the available highway network. This parameterization is consistent with [Allen and Arkolakis \(2019\)](#) and [Hummels and Schaur \(2013\)](#) and is consistent with the multiplicative specification of total trade costs in equation (8) as a function of travel time in equation (19).

<sup>16</sup>[Caliendo and Parro \(2015\)](#) report estimated values of  $\theta(s)$  that range between 0.37 and 51.08 for manufacturing sectors.

In particular, we parameterize trade costs as a function of  $T_{ij}$ :

$$\tau_{ij}(s) \equiv \rho_s(T_{ij}) = \exp\left(-\frac{\varphi(s)}{\theta(s)} T_{ij}\right),$$

such that the following equation is satisfied,

$$\tau_{ij}(s) = \tau_{ik}(s)\tau_{kj}(s) = \exp\left(-\frac{\varphi(s)}{\theta(s)}(T_{ik} + T_{kj})\right).$$

Given this specification of trade costs, we estimate the gravity equation in (12) using Poisson Pseudo Maximum Likelihood for each sector  $s$ :

$$\pi_{ij}(s) = \exp[\varphi(s)T_{ij} + Z_{ij}(s)\Xi^s + exp_i(s) + imp_j(s)] + \epsilon_{ij}, \quad (26)$$

where  $T_{ij}$  is the travel time on the first best route (in hours) between state  $i$  and state  $j$ ,<sup>17</sup>  $exp_i(s)$  and  $imp_j(s)$  are exporter-sector- and importer-sector-specific fixed effects, and  $Z_{ij}(s)$  includes  $\kappa_{ij}(s)$  and  $\varsigma_{ij}(s)$ . The results of estimating  $\varphi(s)$  for each sector  $s$  are shown in Table 2, where standard errors clustered on states  $i$  and  $j$  are reported in parentheses.

**Table 2: Estimates of  $\varphi(s)$**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\varphi(s)$	-0.177	-0.071	-0.184	-0.327	-0.113	-0.111	-0.267	-0.137	-0.076	-0.055	-0.101	-0.082
s.e.	(0.016)	(0.015)	(0.009)	(0.027)	(0.014)	(0.015)	(0.019)	(0.013)	(0.008)	(0.011)	(0.016)	(0.008)
Observations	2,343	2,343	2,333	2,226	2,333	2,349	2,311	2,341	2,337	2,295	2,303	2,337

*Notes:* The table shows estimates of  $\varphi(s)$  using equation (26) for each sector  $s$ . The dependent variable is the trade shares between state  $i$  and state  $j$ . All specifications include state importer and exporter fixed effects. Standard errors clustered on origin and destination state are reported in parentheses. Column 1 is Food, Beverage, and Tobacco, Column 2 is Textiles and Leather, Column 3 is Wood, Paper, and Printing, Column 4 is Petroleum and Coal, Column 5 is Chemicals, Column 6 is Plastics and Rubber, Column 7 is Nonmetallic Minerals, Column 8 is Primary and Fabricated Metals, Column 9 is Machinery, Column 10 is Computers, Electronics, and Electrical, Column 11 is Transportation Equipment, and Column 12 is Furniture and Miscellaneous.

The results indicate substantial heterogeneity across sectors, with estimates of  $\varphi(s)$  ranging from -0.055 for *Computer, Electronic Products, and Electrical Equipment* to -0.327 for *Petroleum and Coal Products*. This is consistent with intuition that trade flows for relatively high value and light weight goods will be less responsive to shipping time, while cheaper and heavier goods are more sensitive to shipping time.

We can now characterize the exact functional form for the relationship between trade-

<sup>17</sup>Note that by construction  $h_{ij} = 1$  in the benchmark equilibrium.

generated congestion and trade costs. Intuitively, higher trade-generated traffic increases the level of service (LOS), which decreases speed and increases travel time. To estimate this relationship, we calculate  $N_{\mathcal{S}}$  following equations (15) and (16), where we measure the benchmark value of  $M_{ij}$  using the data on total value of trade flows and average shipment size from the Commodity Flow Survey in 2012.<sup>18</sup> We then estimate the following regression:

$$\ln \text{LOS}_{\mathcal{S}} = \zeta \ln N_{\mathcal{S}} + \nu \ln \text{distance}_{\mathcal{S}} + \text{pair}_{\mathcal{S}}^{ij} + \epsilon_{\mathcal{S}}, \quad (27)$$

where  $\text{distance}_{\mathcal{S}}$  is the length  $\mathcal{S}$  in miles, which we include to control for the fact that shorter segments may be subject to higher per-mile traffic, and  $\text{pair}_{\mathcal{S}}^{ij}$  is a set of  $ij$ -symmetric dummy variables that equal one if  $\mathcal{S}$  is used for transporting goods to or from state  $i$  to  $j$  and zero otherwise.

In addition to estimating equation (27) via OLS, we also control for potential confounding and measurement errors by using an instrumental variable approach. For that, we develop an instrument for  $N_{\mathcal{S}}$  based on the distribution of population across US states in year 1900. First, we calculate a bilateral population variable  $B_{ij} = L_i + L_j$  between states and then the instrument for  $N_{\mathcal{S}}$  as:

$$B_{\mathcal{S}} = \sum_{\mathcal{S}} \mathbf{1}_{\mathcal{S} \in \mathcal{S}_{ij}} B_{ij}, \quad (28)$$

such that the first stage of the IV approach is as follows:

$$\ln N_{\mathcal{S}} = \varpi \ln B_{\mathcal{S}} + \nu^b \ln \text{distance}_{\mathcal{S}} + \text{pair}_{\mathcal{S}}^{ij} + \epsilon_{\mathcal{S}}^b.$$

Hence, the identification assumption behind the IV approach and the instrument for  $\ln N_{\mathcal{S}}$  is that the distribution of population across US states in 1900 is correlated with bilateral trade patterns but not congestion levels on an individual highway segment. It is plausible that the population distribution in 1900 is largely unrelated to segment commuting traffic and other factors affecting segment congestion in 2012. We report the OLS and IV results in Table 3.

The estimated OLS and IV coefficients on  $\ln N_{\mathcal{S}}$  are precisely estimated and equal 0.021 and 0.07, respectively. For our calibration and counterfactual experiments, we rely on the IV coefficient, which implies that a 1 percent increase in trade-generated traffic increase the

---

<sup>18</sup>Note that trade-generated traffic is not the only determinant of  $\text{LOS}_{i\mathcal{S}}$ . Other factors such as commuting may play an important role in determining congestion on each highway segment. While disentangling different factors that shape LOS is beyond the scope of this paper, note that we use measures of calculated traffic on a state-to-state level, which is influenced by local commuting traffic to a lower extent.

**Table 3: The Relationship Between LOS and Traffic**

	OLS	IV
$\ln N_{\mathcal{S}}$	0.021 (0.005)	0.070 (0.007)
$\ln \text{distance}_{\mathcal{S}}$	-0.016 (0.002)	-0.016 (0.002)
First-stage F-statistic	20,259	
Observations	34,236	

*Notes:* The table shows the results of estimating (27) on all highway segments with positive traffic. The dependent variable is the level of service (LOS) on segment  $\mathcal{S}$  in state  $i$ . All specifications include a set of  $ij$ -symmetric dummy variables that equal one if  $\mathcal{S}$  is used for transporting goods to or from state  $i$  to  $j$  and zero otherwise. The log of  $N_{\mathcal{S}}$  reflects the amount of trade-related traffic on  $\mathcal{S}$  and the log of  $\text{distance}_{\mathcal{S}}$  is the distance (in miles) of segment  $\mathcal{S}$ . Eicker–Huber–White standard errors are reported. The first-stage F-statistic is the Kleibergen–Paap statistic.

level of service by 0.07 percent. This allows us to construct the counterfactual level of service for each  $\mathcal{S}$  as:

$$\text{LOS}'_{\mathcal{S}} = (\widehat{N}_{\mathcal{S}})^{0.07} \text{LOS}_{\mathcal{S}}.$$

Using  $\text{LOS}'_{\mathcal{S}}$  we can calculate the counterfactual congestion speed coefficient  $\mathcal{A}'_{\mathcal{S}}$  following the step function in equation (1).

### CALIBRATING STATE-PORT TRADE SHARES

The remaining component required for the counterfactuals is the state-port-country share,  $\lambda_{in}^r(s)$ , which are not directly observed in the data. To overcome this limitation, we use the predictions of the theoretical model together with a parameterization of international trade costs to calibrate  $\lambda_{in}^s(s)$  from the data at a higher aggregation level. In particular, we combine data on total sectoral shipments from each state to each of the 21 US ports with data on total sectoral shipments from each port to each foreign country.

The first data set describes shipments from each US state  $i$  to each US port  $r \in j$  and corresponds to the following equation in the context of our theoretical framework:

$$\Lambda_i^r(s) = \sum_{n \neq i} \pi_{in}^r(s) Y_n(s) = \sum_{n \neq i} T_i(s) (\kappa_{ir}(s) \bar{\tau}_{ir}(s) \varepsilon_r^{m(ir)r}(s) \xi_r^r(s) t_{rn}^r(s))^{-\theta(s)} P_n(s)^{\theta(s)} Y_n(s).$$

The second data set describes shipments from each US port to each foreign trading partner. In the context of our theoretical model, the following equation describes shipments between



port  $r \in j$  and country  $n$  (for  $n \neq i$ ):

$$V_n^r(s) = \sum_{i \neq n} \pi_{in}^r(s) Y_n(s) = \sum_{i \neq j} T_i(s) (\kappa_{ir}(s) \bar{\tau}_{ir}(s) \varepsilon_r^{m(ir)r}(s) \xi_r^r(s) t_{rn}^r(s))^{-\theta(s)} P_n(s)^{\theta(s)} Y_n(s)$$

To isolate the parameters of interest, we first pre-multiply  $\Lambda_i^r(s)$  by  $(\kappa_{ir}(s) \bar{\tau}_{ir}(s) \varepsilon_r^{m(ir)r}(s))^{\theta(s)}$  and then use Poisson Pseudo Maximum Likelihood to estimate the following equation using data on trade flows from each US state  $i$  to each US port  $j$  separately for each sector  $s$ :

$$\Lambda_i^r (\kappa_{ir}(s) \bar{\tau}_{ir}(s) \varepsilon_r^{m(ir)r}(s))^{\theta(s)} = \exp(\text{state}_i(s) + \text{iport}_r(s)) + \epsilon_{ir}(s), \quad (29)$$

where  $\text{state}_i(s)$  and  $\text{iport}_r(s)$  are sector-state-specific and sector-port-specific fixed effects. We then estimate the following equation (again, with Poisson Pseudo Maximum Likelihood) using data on trade flows between each US port  $r$  and each foreign trading partner  $j$ :

$$V_n^r(s) = \exp\left(\text{xport}_r(s) + \sum_{q=1}^5 \psi_q(s) \mathcal{Q}_q \ln(\text{distance}_{rn}) + \text{country}_n(s)\right) + \epsilon_{rj}(s) \quad (30)$$

where  $\text{xport}_r(s)$  and  $\text{country}_n(s)$  are sector-port-specific and sector-country-specific fixed effects. We parameterize international trade costs via water using the (log) distance in miles between port  $r$  and country  $n$  and allow this to vary with indicators for each quintile  $\mathcal{Q}_q$ . The estimated coefficients on the indicator variables are reported in Figure C1 in the Appendix.

The estimates from equations (29) and (30) allow us to calibrate international trade costs,  $t_{rn}(s)^{-\theta(s)}$ , and the sector-port-specific productivity level,  $\xi_r^r(s)$ , as follows:

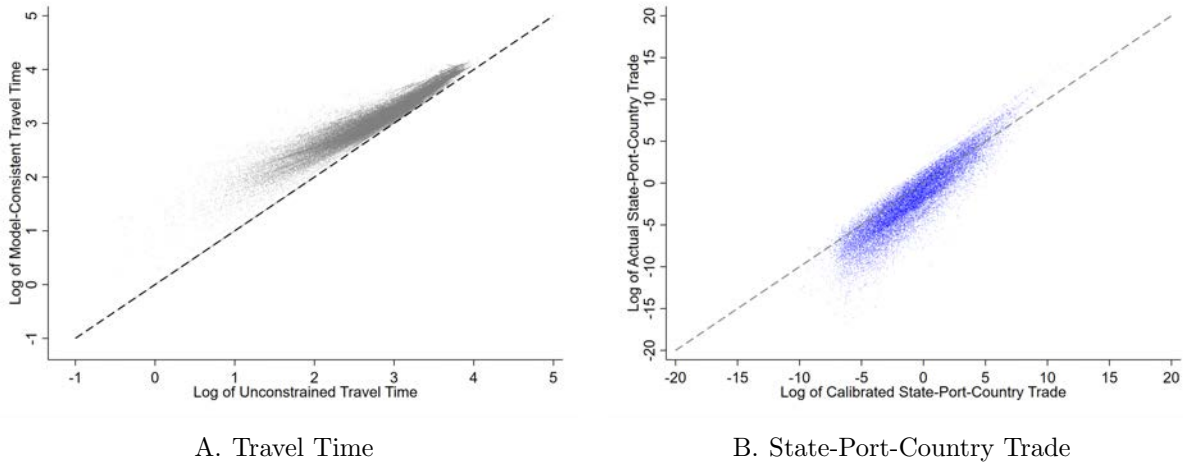
$$t_{rn}(s)^{-\theta(s)} = \sum_{q=1}^5 \hat{\psi}_q(s) \mathcal{Q}_q \ln(\text{distance}_{rn}) \quad \text{and} \quad \xi_r^r(s)^{-\theta(s)} = \frac{\exp(\widehat{\text{xport}}_r(s))}{\sum_i \exp(\widehat{\text{state}}_i(s)) (\kappa_{ir}(s) \bar{\tau}_{ir}(s) \varepsilon_r^{m(ir)r}(s))^{-\theta(s)}}.$$

Once we have recovered  $t_{rn}(s)^{-\theta(s)}$  and  $\xi_r^r(s)$ , we can calculate exporter-port-importer shares  $\lambda_{in}^r(s)$  and  $\lambda_{nj}^r(s)$ .

## 4.2 Calibrated Model versus Data

In this section, we quantitatively evaluate two assumptions of the model against the available data. First, our procedure of aggregating counties into states as well as using state-level outcomes to inform county-level economic outcomes depends on the specification of trade costs in equation (8), where we assume that there is a common export county  $c(ij)$  used by all counties in  $i$  to export to  $j$ . We compare how accurately this specification reflects travel

**Figure 5: Calibrated Model versus Data**



*Notes:* Panel A shows the relationship between unconstrained average county-to-county trade costs and their model consistent counterparts. Panel B shows the relationship between calibrated ( $x$ -axis) and actual ( $y$ -axis) exporter-port-importer trade shares together with the 45 degree line. The calibrated shares are constructed by summing  $\lambda_{in}^r(s)$  from the model over  $s$ . The actual shares are aggregated data from USA Trade Online that were not used in calibration.

time when the common-hub assumption is not imposed. In Panel A of Figure 5, we plot travel time between all  $c \in i$  and  $m^* \in j$  calculated in an unconstrained way vs. under the assumptions of the model. Unconstrained travel times are lower than their model-consistent counterparts, which is not surprising as outcomes of any constrained optimization should deliver relatively longer travel times uniformly for all  $ij$ -pairs. However, for our purposes it is important to match the spatial variation in travel times across different origins and destinations. The figure suggests that our multiplicative specification of trade costs is able to do that as the spatial correlation between model-consistent  $ij$ -specific travel times and their unconstrained counterparts is 0.96.

Second, we evaluate how the calibrated exporter-port-importer shares square up with the available data. We compare the predictions of the model to other data available from USA Trade Online not used in the estimation or calibration. These data include information on total exports from each US state via each US port with each foreign country. We use the port share estimates,  $\lambda_{ij}^r(s)$ , multiplied by bilateral sectoral trade flows and aggregated over sectors to predict total exporter-port-importer trade flows. These predictions (in log) are then compared to the corresponding actual data (in log). The results in Panel B of Figure 5 suggest that our calibration of  $\lambda_{in}^r(s)$  matches the data well. The correlation between (log) predicted and (log) actual aggregate state-port-country trade flows is 0.92. The correlation between (log) model trade costs and (log) unconstrained trade costs is 0.96.

## 5 Counterfactual Results

In this section, we use the calibrated model to run counterfactual experiments to quantify the value of the IHS and its separate segments. The first set of experiments quantifies the losses from removing the entire Interstate Highway System. We eliminate segments that belong to the IHS such that the counterfactual available road network is  $\mathcal{S}' \subseteq \mathcal{S}$  and producers are forced to re-optimize by choosing different shipping routes and generally face higher trade costs relative to the benchmark equilibrium. When 50,000 miles which form the IHS become unavailable, producers are forced to choose optimal routes using the remaining 342,000 miles of non-IHS federal and state highways. In the second set of experiments, we evaluate the losses from removing individual segments of the IHS (I-5, I-10, etc). For these counterfactuals we focus on ten of the largest numbered interstate segments.

For each counterfactual scenario, we also decompose the aggregate effect into the contribution of domestic versus international trade costs components. To isolate the domestic component of total welfare costs, we assume that producers use counterfactual road network  $\mathcal{S}' \subseteq \mathcal{S}$  for interstate trade but not for international exports and imports. The value of the international component is then calculated as the difference between the total value and the value due to the domestic component. Hence, the interpretation of the international component is the marginal value due to better international market access conditional on using counterfactual highway network for domestic trade. Hence, by design we rule out a possibility that counterfactual domestic roads are used exclusively for international trade. Note that our counterfactual experiments quantify the value of the entire Interstate Highway System and its individual segments through the lens of domestic and international trade. Hence, our results should be interpreted as additional value that the IHS provides due to easier movement of goods across domestic and international locations while taking other potential benefits of the highway system, e.g. easier commuting, as given.

### 5.1 Removing the Entire IHS

The baseline counterfactual results are shown in the first row of Table 4. In the benchmark specification we assume complete labor mobility such that ( $\varrho \rightarrow \infty$ ). Column 1 reports that in the absence of the IHS—including nearly 50,000 miles of limited-access roads graded for high travel speeds—real GDP losses are equal to \$601.4 billion in 2012 dollars (or 3.8 percent).<sup>19</sup> The remaining columns decompose this aggregate effect from removing the IHS for

---

<sup>19</sup>These estimates are larger than existing estimates in the literature. For example, our estimates are about three times larger than the \$150 to \$200 billion reported by [Allen and Arkolakis \(2014\)](#). Given

routes associated with all US trade into the domestic (column 2) and international (column 3) components, respectively, \$433.8 and \$154.9 billion. It is noteworthy that the international component of trade costs accounts for roughly one quarter of the total losses from removing the IHS. The total and decomposed losses presented in row 1 indicate both that the value of the IHS is substantial and that the access it provides to international markets is quantitatively important.

To quantify the importance of different channels in driving the total welfare results, we conduct three additional experiments where we turn off trade-generated traffic congestion, migration ( $\varrho = 1$ ), and input-output linkages one at a time. In addition, in Appendix B, we report results from an extension of the model that includes agglomeration and local congestion externalities in the spirit of Allen and Arkolakis (2014). The second and third rows of Table 4 show alternative versions of the baseline model without congestion and no labor mobility. As noted above, approximately 18,000 miles of the Interstate Highway System are considered to be congested. In row 2 we remove congestion as a mechanism in the model that reduces the speed of travel on roads with more traffic and also eliminate the potential for the reallocation of traffic in response to changes in the highway network. The results suggest that the effects of trade related congestion are nearly \$128 billion or 21 percent of total effect in row 1.<sup>20</sup> The difference between our results and the implied congestion costs derived from surveys may reflect the difference between our focus on trade-related congestion and the congestion faced by urban commuters. Integrating these two sources of congestion in a consistent theoretical framework and quantifying their effects is an important avenue for future research.

The results in row 3 do not allow for migration in response to the changes in the highway network from removing the IHS. The total and decomposed losses are close to the baseline counterfactual results that allow for migration; the difference between the baseline and “no migration” scenarios is about \$1 billion with the latter being higher. This suggests that migration may be less important in the context of our counterfactuals where productivity and local amenities are fixed and migration is costly.<sup>21</sup>

---

upfront construction costs of \$560 billion in 2007 dollars and assuming a 5% annual cost of capital, annual upfront construction costs are \$28 Billion (see [Federal Highway Administration, 2012](#)). Including annual maintenance costs and capital improvements of \$3.5 billion and \$20.4 billion respectively, the annual payback of the IHS dwarfs the annual costs. For international context, our estimates are up to twice as large as the impact of India’s Golden Quadrilateral ([Alder, 2017](#); [Asturias, García-Santana and Ramos, 2019](#)).

<sup>20</sup>This means that trade-generated congestion costs are roughly \$400 per person. These costs are large, but smaller than those estimated from surveys (e.g., [INRIX Research, 2019](#)).

<sup>21</sup>If, alternatively, we allowed for changes of the fundamental characteristics of locations in response to counterfactual changes in the highway network we would expect the associated losses to be larger in the

**Table 4: Total Losses from Removing the IHS**

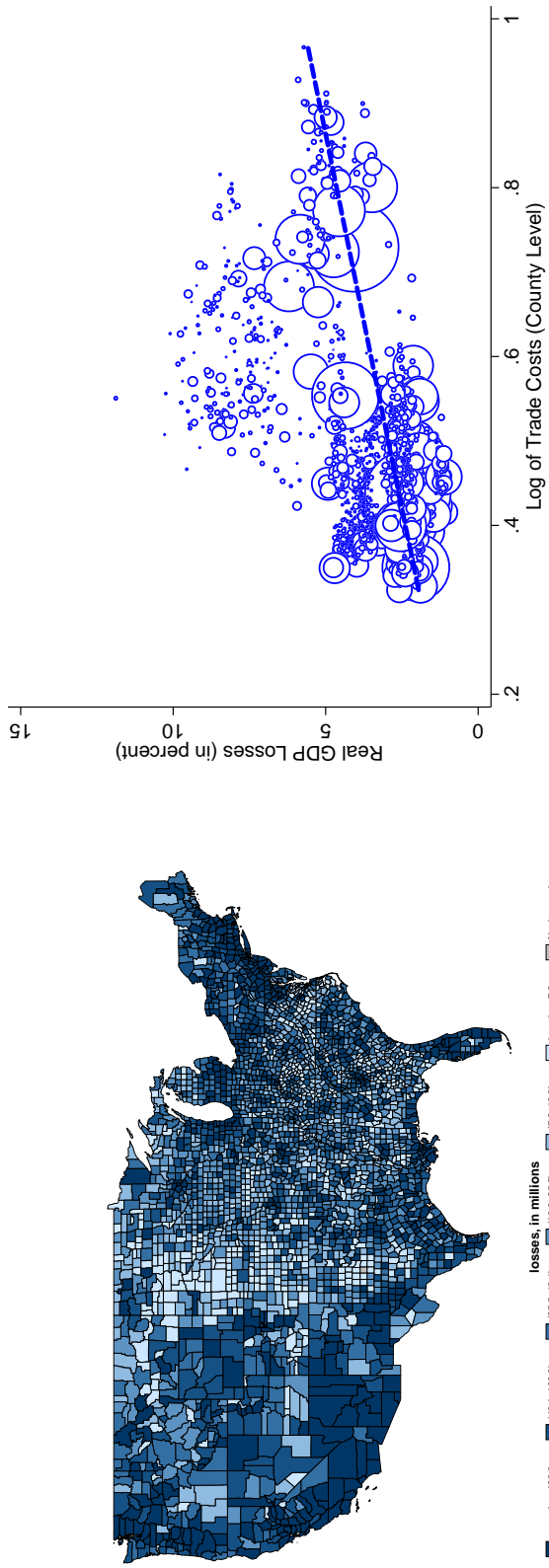
	Total (1)	Domestic (2)	International (3)
Baseline, No IHS	601.4	445.9	155.6
No Congestion	473.9	369.4	104.5
No Migration	602.1	447.0	155.1
No Input-Output	238.6	172.9	65.7

*Notes:* The table shows results from counterfactual exercises removing the Interstate Highway System. Column 1 shows the reduction in real GDP from removing the IHS for both the domestic and international components of trade costs. Column 2 shows the reduction in real GDP from removing the IHS domestic components of trade costs. Column 3 shows the difference between columns 1 and 2, which is the reduction in real GDP from removing the IHS foreign components of trade costs. The results in row 1 are for the baseline version of the model. The remaining rows show results for versions of the model with no congestion (row 2), no migration (row 3), and no input-output linkages (row 4).

An important feature of our model is the presence of many sectors and linkages across sectors and countries through input-output relationships. This may be particularly important in the context of transportation infrastructure as better road networks allow remote locations specialize in specific sectors, which improves overall efficiency. For example, [Hornbeck and Rotemberg \(2019\)](#) and [Asturias, García-Santana and Ramos \(2019\)](#), respectively, find substantial gains from allocative efficiency associated with improvements in railroads in the United States during the late nineteenth and early twentieth centuries and roads in India more recently. The fourth row of [Table 4](#) presents results of removing the IHS without the input-output structure linking sectors in the full model. In this case the results differ substantially from the baseline results. Instead of \$601.4 billion reported in row 1, the total losses in row 4 are \$238.6 billion and the decomposed losses are \$172.9 and \$64.7 billion, respectively, for the domestic and international components of trade costs. This highlights the importance of sectoral heterogeneity, spatial specialization patterns, and the input-output linkages for understanding the effects changes in the US highway network. In addition, the results in row 4 more closely match the results obtained by [Allen and Arkolakis \(2014\)](#) for a similar counterfactual exercise. These authors estimate losses from removing the IHS between \$150 and \$200 billion in 2007 dollars. We obtain results that fall in this range when we eliminate intermediate inputs from the model and focus only on the effect of the domestic trade cost components.

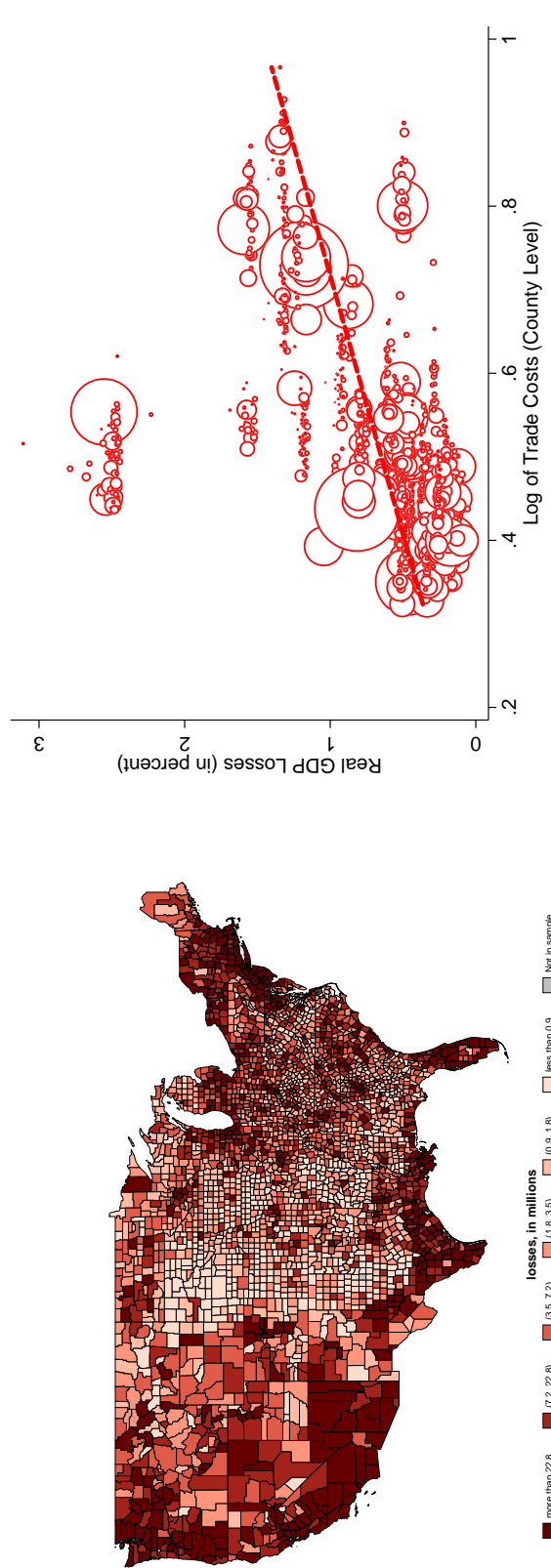
Panel A of [Figure 6](#) illustrates the geographic distribution of the total effects from removal of migration. See recent work by [Heblich, Redding and Sturm \(2020\)](#) and [Brinkman and Lin \(2019\)](#) on the effects of transportation infrastructure on the fundamentals within cities.

Figure 6: Losses from Removing the IHS by County



Notes: The figure shows the results for removing the Interstate Highway System at the county level. Panel A shows the geographic distribution of the reduction in real GDP (in percent) from removing the IHS for both the domestic and international components of trade costs. Panel B shows the relationship between the reduction in real GDP from removing the IHS for all trade cost components and the level of actual trade costs in 2010.

Figure 7: Losses from Removing the International Trade Cost Component of the IHS by County



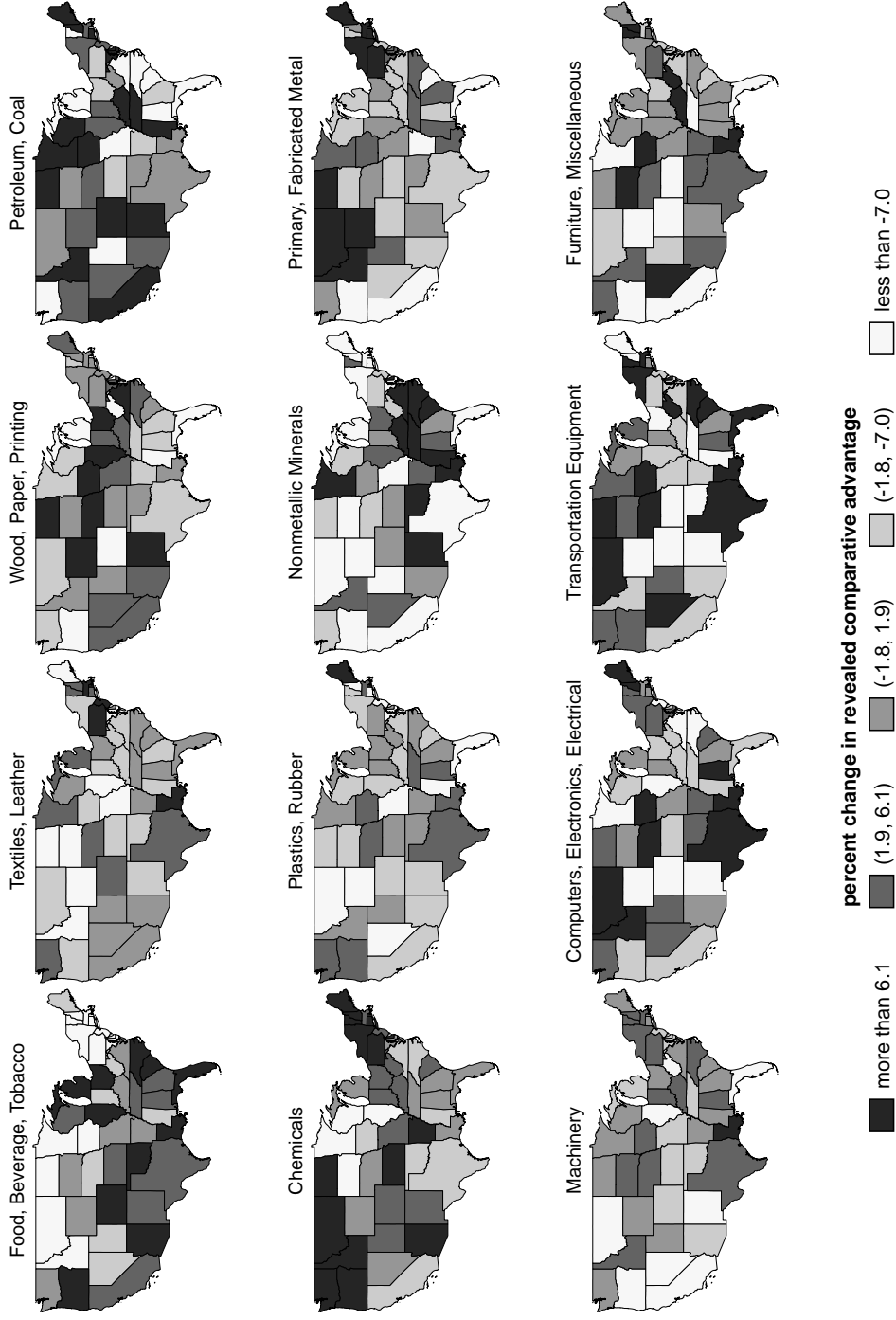
A. Contribution of Foreign Trade Costs to the Reduction Real GDP

B. Reduction of Real GDP and Foreign Trade Costs

Notes: The figure shows the decomposition results for removing the Interstate Highway System due to the contribution of the international trade cost components at the county level. Panel A shows the geographic distribution of the reduction in real GDP (in percent) from removing the IHS for the international components of trade costs. Panel B shows the relationship between the reduction in real GDP from removing the IHS for the international trade cost components and the level of actual trade costs in 2010.



**Figure 8: The Effect of Removing the IHS on Revealed Comparative Advantage by State**



*Notes:* The figure shows the change (in percent) in revealed comparative advantage from removing the IHS for both the domestic and international components of trade costs at the state level for twelve manufacturing sectors. Revealed comparative advantage is calculated as in Balassa (1965):  $(E_i(s) / \sum_{s'} E_i(s')) / (\sum_{i'} E_{i'}(s) / \sum_{i',s'} E_{i'}(s'))$ .



ing the IHS at the county level. The largest total losses are concentrated in the northeastern and western regions of the United States. Across the 3,106 counties in the sample, all experience at least some loss, while the average loss is \$193.7 million. Panel B of Figure 6 shows the relationship between the log of average trade costs in 2010 and the losses (in percent) from removing the IHS, which indicates that losses are concentrated in counties that are more remote from domestic and international markets. Figure 7 shows the losses attributed to removing the IHS for the international component of trade costs. These losses overlap in some counties, but other counties are affected differently by the changes in domestic versus international trade costs due to the IHS—the correlation between the domestic and foreign components is 0.80. This suggests that the IHS plays different roles in facilitating trade across US counties and states. For example, total losses of \$65 billion for Texas are split more evenly between domestic and international trade costs (\$38 and \$26) than in smaller state economies, e.g., Alabama, where losses from the change in the domestic trade cost component are substantially more important.

Next, we calculate how the IHS shaped specialization patterns across locations in the United States. For that we use the concept of revealed comparative advantage as in Balassa (1965) for the twelve manufacturing sectors across US states. Let  $E_i(s)$  denote total exports of state  $i$  of sector  $s$  goods, then the measure of comparative advantage is:

$$CA_i(s) = \frac{E_i(s)}{\sum_{s'} E_i(s')} / \frac{\sum_{i'} E_{i'}(s)}{\sum_{i',s'} E_{i'}(s')}$$

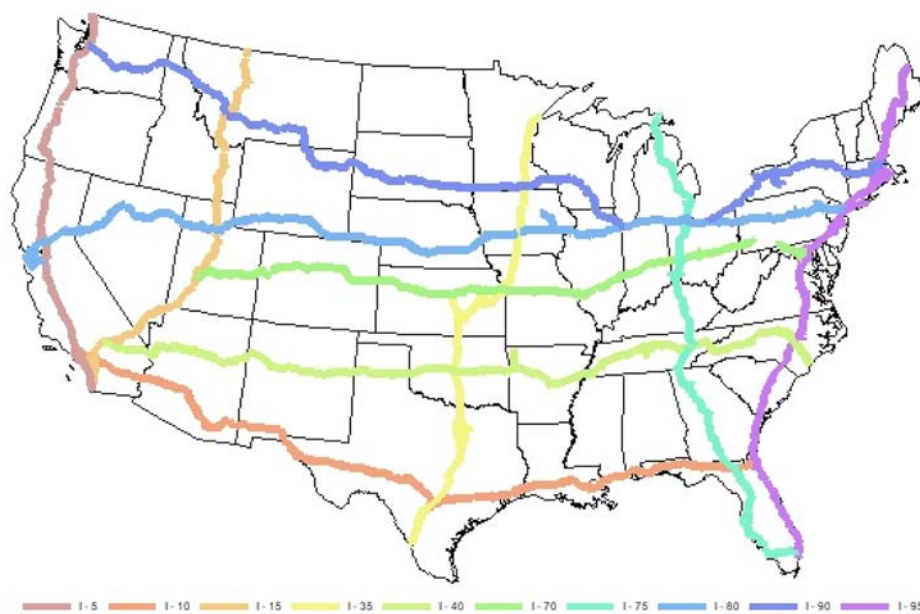
We calculate the change in  $CA_i(s)$  under two scenarios. First, we calculate how the measure of revealed comparative advantage changes in the absence of the IHS denoted by  $\Delta CA_i(s)'$ . However, these changes have two components. On the one hand, removal of the IHS changes market access for all states due to average reductions in trade costs. On the other hand, the IHS has a differential impact on specialization across locations. To isolate the latter, we also calculate changes in  $CA_i(s)$  under uniform 34% increase in travel time (average change across all  $ij$  pairs) for all state pairs denoted by  $\Delta CA_i(s)''$ . We then plot the differences between these two changes,  $\Delta CA_i(s)' - \Delta CA_i(s)''$  in Figure 8. Hence, the results in Figure 8 reflect the role of the IHS in shaping regional specialization patterns net of average changes in trade costs. In general, we expect to observe changes in revealed comparative advantage that reflect trade cost minimizing decisions on the part of producers that balance the access to input and output markets. For example, in some sectors (e.g., *Food, Beverage, and Tobacco*), states in the middle of the country export relatively more to other states and countries in the absence of the IHS. This change partially reflects proximity to final goods

consumers as well as proximity to suppliers. For other sectors (e.g., *Machinery*), states on the coast export relatively less without access to the IHS. These findings complement work that shows how industrial composition change in response to trade costs (Michaels, 2008; Duranton, Morrow and Turner, 2014).

## 5.2 Removing Individual IHS Segments

The results in the previous subsection focus on removing the entire IHS. From the perspective of policymaking, it is also useful to address smaller changes in the highway network that quantify the value of individual segments and can thus serve as a guide for the allocation of funding for new construction, improvements, and maintenance. To do this, in this subsection, we consider counterfactual exercises that remove ten longest highways of the IHS illustrated in Figure 9. This will shed light on the aggregate benefits of improving individual segments, the distribution of those gains across US states, and variation in the importance of the access provided to domestic versus international markets.

**Figure 9: Individual Segments of the Interstate Highway System**



In this exercise, we focus on the losses associated with ten of the largest Interstate Highway System segments. For these counterfactuals, we remove all sections of the corresponding numbered interstate, including loops and spurs, and allow traffic to adjust endogenously to changes in trade costs. The results are presented in Table 5. Column 1 gives the total length in miles of each segment. Columns 2 and 3 report the total and per mile reduction in real

**Table 5: Results for Removing IHS Segments**

Interstate Highway Segment:	Total Miles (1)	Total, in billions (2)	Total Per Mile, in millions (3)	Domestic, in billions (4)	International, in billions (5)
I-5	1386.2	26.3	19.0	16.0	10.3
I-10	2452.2	38.4	15.6	24.4	14.0
I-15	1437.7	8.8	6.1	0.8	7.9
I-35	1428.2	22.7	15.9	8.5	14.3
I-40	2528.3	11.9	4.7	1.4	10.6
I-70	2066.0	40.0	19.3	28.4	11.6
I-75	1752.2	39.9	22.7	24.7	15.2
I-80	2875.1	51.7	18.0	34.1	17.5
I-90	2796.6	37.8	13.5	15.8	22.0
I-95	1888.1	31.2	16.5	15.0	16.1

*Notes:* The table shows results from counterfactual exercises removing the ten longest individual segments (in miles) of the Interstate Highway System. Column 1 shows the total number of miles. Columns 2 and 3 show the total and per-mile reduction in real GDP, respectively. Columns 4 and 5 show the portion of the total reduction attributed to the domestic and international components of trade costs, respectively.

GDP from removing each IHS segment, while fixing the rest of the highway network. For these ten segments, the total losses range from \$51.7 billion for I-80 to \$8.8 billion for I-15. The two segments with the next largest losses after I-80 are I-70, and I-75.

A few details are noteworthy. First, both I-80, I-70, I-75, I-10, and I-90 stand out with losses that are substantial relative to the other IHS segments. This reflects a combination of the lack of available alternate routes along the West-East direction in the United States. Looking at the losses by numbered interstate, on per mile basis reveal that two primary east-west routes (I-70 and I-80) together with the coastal route (I-5) and north-south route (I-75) are the most valuable.<sup>22</sup> Coming back to the substantial losses associated with the removal of specific highways, it is clear from Columns 4 and 5 that these interstates generate a significant portion of their value by facilitating international trade. For example, the \$10.3 billion in international trade generated by I-5 is larger than the total value generated by I-15. These results suggest that international market access provides a substantial value of individual highway segments.

Finally, in addition to these aggregate results, it is useful to highlight variation in the losses across US states. For example, removing I-5 generates \$26.3 billion in total losses, but

<sup>22</sup>These numbered routes roughly correspond to the proposed system of interstate highways by Franklin D. Roosevelt in 1938 (Department of Transportation, 1967).

losses for California, Oregon, and Washington together are \$28.1 billion and gains accrue to some remaining states as trade and economic activity are reallocated to other locations. We can also see that even among states that are directly affected by the removal a highway, there can be substantial differences in losses. For example, removing I-95 reduces real GDP in Maine by 2.9 percent and Massachusetts by about 0.5 percent. These findings are important in the event that highway funding model in the United States is revised.

In general, the results in this subsection are useful for prioritizing spending on new construction, improvements, and maintenance, as well as understanding the distributional consequences of these decisions. For example, there are currently several high priority corridors designated as future interstates and portions of four highways that are planned to be integrated into the IHS. Our approach and the results from this paper can be used to evaluate these proposals, i.e., which have the largest gains, the source of those gains, and the distributional consequences across regions and sectors.

## 6 Conclusion

Domestic transportation infrastructure facilitates trade within countries and international trade with the rest of the world. This suggests that the value of domestic transportation infrastructure reflects its contribution to both types of market access. For the United States, a key part of the domestic transportation infrastructure is the nearly 50,000 limited-access high-grade road miles that make up the Interstate Highway System. Despite the vital role that these highways play in both domestic *and* international trade, there is limited research quantifying the aggregate and relative importance of the dual functions performed by the IHS in US domestic and international trade.

In this paper, we build a multisector model of interregional and international trade of the United States. Importantly, the model accounts for the rich internal geography of the United States by integrating each US county with all other counties and foreign countries via the US highway network, US ports, and international shipping. In addition, the model accounts for the potential congestion of the US highway network that affects trade costs and may alter the associated pattern of both internal and external trade. In the first set of results, we use the model to quantify the losses associated with removing the entire IHS. We find losses equal to \$601.4 billion with about one quarter due to higher trade costs for accessing foreign markets and over 20 percent due to congestion. In the second set of results we focus on the twenty longest IHS segments and find a range of losses between \$51.7 billion for I-80 and \$8.8 billion for I-15.

Our results contribute to a growing literature in international trade and economic geography on the role of transportation infrastructure. We provide a framework that can be used to quantify the value of existing or proposed infrastructure. In addition, our approach also highlights the interaction between domestic transportation infrastructure and international trade. This is particularly important for understanding the implications of the changing patterns of globalization for the value and distributional consequences of future infrastructure spending and trade policy.

## References

- Agnosteva, Delina E., James E. Anderson, and Yoto V. Yotov, “Intra-national Trade Costs: Measurement and Aggregation,” *European Economic Review*, 2019.
- Alder, Simon, “Chinese Roads in India: The Effect of Transport Infrastructure on Economic Development,” 2017.
- Allen, Treb and Costas Arkolakis, “Trade and Topography of the Spatial Economy,” *Quarterly Journal of Economics*, 2014, 129 (3), 1085–1140.
- and —, “The Welfare Effects of Transportation Infrastructure Improvements,” *Working Paper*, 2019.
- Anderson, James E., “The Gravity Model,” *Annual Review of Economics*, September 2011, 3 (1), 133–160.
- Asturias, Jose, Manuel García-Santana, and Roberto Ramos, “Competition and the Welfare Gains from Transportation Infrastructure: Evidence from the Golden Quadrilateral of India,” *Journal of the European Economic Association*, 2019, 17 (6), 1881–1940.
- Atkin, David and Dave Donaldson, “Who’s Getting Globalized? The Size and Implications of Intra-National Trade Costs,” Technical Report, National Bureau of Economic Research 2015.
- Balassa, Bela, “Trade Liberalisation and “Revealed” Comparative Advantage,” *The Manchester School*, 1965, 33 (2), 99–123.
- Bartelme, Dominick, “Trade Costs and Economic Geography: Evidence from the US,” *Working Paper*, 2018.
- Baum-Snow, Nathaniel, “Did Highways Cause Suburbanization?,” *Quarterly Journal of Economics*, 2007, 122 (2), 775–805.
- , Loren Brandt, J Vernon Henderson, Matthew A Turner, and Qinghua Zhang, “Roads, Railroads, and Decentralization of Chinese Cities,” *Review of Economics and Statistics*, 2017, 99 (3), 435–448.
- Brinkman, Jeffrey and Jeffrey Lin, “Freeway Revolts!,” 2019.
- Bureau of Transportation, “Border Crossing/Entry Data,” 2017.
- Caliendo, Lorenzo and Fernando Parro, “Estimates of the Trade and Welfare Effects of NAFTA,” *Review of Economic Studies*, 2015, 82 (1), 1–44.
- , Maximiliano Dvorkin, and Fernando Parro, “Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock,” *Econometrica*, May 2019, 87 (3), 741–835.

- Chandra, Amitabh and Eric Thompson, “Does Public Infrastructure Affect Economic Activity? Evidence from the Rural Interstate Highway System,” *Regional Science and Urban Economics*, 2000, 30 (4), 457–490.
- Combes, Pierre-Philippe and Miren Lafourcade, “Transport Costs: Measures, Determinants, and Regional Policy Implications for France,” *Journal of Economic Geography*, 2005, 5 (3), 319–349.
- Coşar, A. Kerem and Banu Demir, “Domestic Road Infrastructure and International Trade: Evidence from Turkey,” *Journal of Development Economics*, 2016, 118, 232–244.
- and Pablo D. Fajgelbaum, “Internal Geography, International Trade, and Regional Specialization,” *American Economic Journal: Microeconomics*, 2016, 8 (1), 24–56.
- Cosar, A. Kerem, Banu Demir, Devaki Ghose, and Nathaniel Young, “Road Capacity, Domestic Trade and Regional Outcomes,” 2020.
- Costinot, Arnaud and Andres Rodriguez-Clare, “Trade Theory with Numbers: Quantifying the Consequences of Globalization,” in G. Gopinath, . Helpman, and K. Rogoff, eds., *Handbook of International Economics*, Vol. 4 of *Handbook of International Economics*, Elsevier, 2014, chapter 0, pp. 197–261.
- , Dave Donaldson, and Ivana Komunjer, “What Goods Do Countries Trade? A Quantitative Exploration of Ricardo’s Ideas,” *Review of Economic Studies*, 2012, 79 (2), 581–608.
- Dekle, Robert, Jonathan Eaton, and Samuel Kortum, “Unbalanced Trade,” *American Economic Review*, 2007, 97 (2), 351–355.
- Department of Transportation, *United States System of Highways* 1967.
- Donaldson, Dave, “Railroads of the Raj: Estimating the Impact of Transportation Infrastructure,” *American Economic Review*, 2018, 108 (4-5), 899–934.
- and Richard Hornbeck, “Railroads and American Economic Growth: ‘A Market Access’ Approach,” *Quarterly Journal of Economics*, 2016, 131 (2), 799–858.
- Duranton, Gilles, Peter M Morrow, and Matthew A Turner, “Roads and Trade: Evidence from the US,” *Review of Economic Studies*, 2014, 81 (2), 681–724.
- Eaton, Jonathan and Samuel Kortum, “Technology, Geography, and Trade,” *Econometrica*, 2002, 70 (5), 1741–1779.
- Faber, Benjamin, “Trade Integration, Market Size, and Industrialization: Evidence from China’s National Trunk Highway System,” *Review of Economic Studies*, 2014, 81 (3), 1046–1070.
- Fajgelbaum, Pablo and Stephen J. Redding, “Trade, Structural Transformation and Development: Evidence from Argentina 1869-1914,” 2018.

- Fajgelbaum, Pablo D and Edouard Schaal, “Optimal Transport Networks in Spatial Equilibrium,” *Econometrica*, 2020, 88 (4), 1411–1452.
- Fernald, John G., “Roads to Prosperity? Assessing the Link between Public Capital and Productivity,” *American Economic Review*, 1999, 80 (3), 619–638.
- Fishlow, Albert, *American Railroads and the Transformation of the Ante-bellum Economy*, Cambridge, MA: Harvard University Press, 1965.
- Fogel, Robert W., *Railroads and American Economic Growth: Essays in Econometric History*, Baltimore, MD: Johns Hopkins University Press, 1964.
- Head, Keith and John Ries, “Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade,” *American Economic Review*, September 2001, 91 (4), 858–876.
- Heblich, Stephan, Stephen J. Redding, and Daniel M. Sturm, “The Making of the Modern Metropolis: Evidence from London,” *The Quarterly Journal of Economics*, 2020, 135 (4), 2059–2133.
- Herzog, Ian, “National transportation networks, market access, and regional economic growth,” *Journal of Urban Economics*, 2021, 122, 103316.
- Hornbeck, Richard and Martin Rotemberg, “Railroads, Reallocation, and the Rise of American Manufacturing,” Technical Report, National Bureau of Economic Research 2019.
- Hummels, David L. and Georg Schaur, “Time as a Trade Barrier,” *American Economic Review*, 2013, 103 (7), 2935–2959.
- INRIX Research, *Global Traffic Scorecard* 2019.
- Isserman, Andrew and Terance Rephann, “New Highways as Economic Development Tools: An Evaluation using Quasi-Experimental Matching Methods,” *Regional Science and Urban Economics*, 1994, 24 (6), 723–751.
- Jaworski, Taylor and Carl T. Kitchens, “National Policy for Regional Development: Historical Evidence from Appalachian Highways,” *Review of Economics & Statistics*, 2019, 101 (5), 777–790.
- Michaels, Guy, “The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System,” *Review of Economics and Statistics*, 2008, 90 (4), 683–701.
- Nagy, David Krisztián, “Hinterlands, city formation and growth: evidence from the U.S. westward expansion,” Economics Working Papers 1717, Department of Economics and Business, Universitat Pompeu Fabra February 2020.
- National Research Council, *Highway Capacity Manual* 2000.

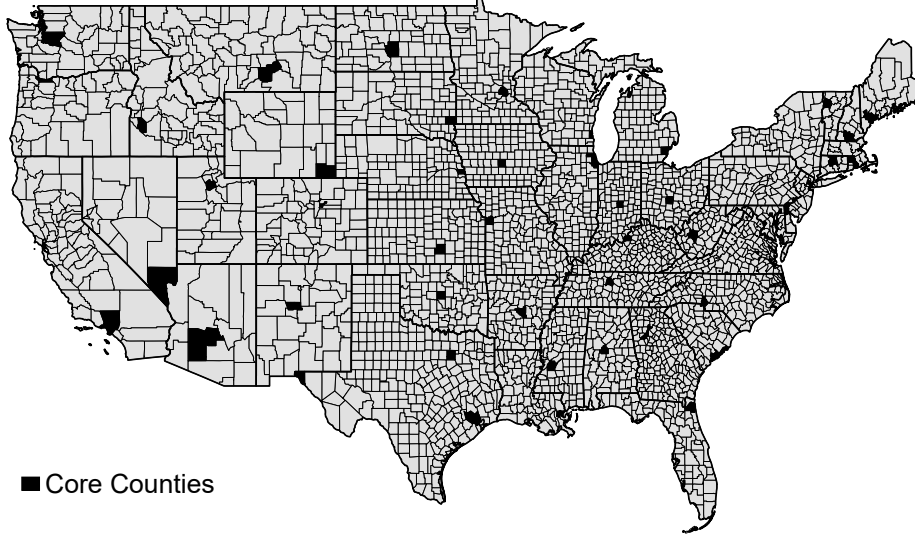


- Ossa, Ralph, “Why trade matters after all,” *Journal of International Economics*, 2015, *97* (2), 266–277.
- Ramondo, Natalia, Andrés Rodríguez-Clare, and Milagro Saborío-Rodríguez, “Trade, Domestic Frictions, and Scale Effects,” *American Economic Review*, 2016, *106* (10), 3159–84.
- , —, and —, “Internal Geography, Labor Mobility, and the Distributional Impacts of Trade,” *American Economic Journal: Macroeconomics*, 2019, *11* (3), 252–288.
- Redding, Stephen J., “Goods Trade, Factor Mobility and Welfare,” *Journal of International Economics*, 2016, *101*, 148–167.
- Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B Schroeder, Matthew Sobek et al., “Integrated public use microdata series: Version 5.0 [Machine-readable database],” *Minneapolis: University of Minnesota*, 2010, *42*.
- Timmer, Marcel P., Erik Dietzenbacher, Bart Los, Robert Stehrer, and Gaaitzen J. De Vries, “An Illustrated User Guide to the World Input-Output Database: The Case of Global Automotive Production,” *Review of International Economics*, 2015, *23* (3), 575–605.
- US Census Bureau, *Highway Statistics Series* 2010.
- , *TIGER/Line shapefiles* 2012.
- US Department of Transportation, *Freight Facts and Figures* 2013.

## Appendix – For Online Publication

### A Core Counties

Figure A1: Core Counties



### B Extension with Agglomeration and Congestion

In this section we extend the model presented to Section 3 to consider the role of agglomeration (and congestion) effects. On the one hand, higher  $L_i^c$  has positive effects of local productivity, which reflects agglomeration effects. On the other hand, higher  $L_i^c$  makes  $c \in i$  relatively less attractive destination for migrants, which reflects congestion effects. This setup follows [Allen and Arkolakis \(2014\)](#).

When agglomeration externalities exist, counterfactual changes in county-sector-specific unit cost are negatively affected by the change in total labor force in that county:

$$\widehat{\kappa}_i^c(s) = (\widehat{L}_i^c)^{-ag} \widehat{w}_i^{\gamma_i(s)} \left( \prod_{\dot{s} \in S} \widehat{P}_i^c(\dot{s})^{\eta_i(\dot{s}s)} \right)^{1-\gamma_i(s)},$$

where we set  $ag = 0.1$  as in [Allen and Arkolakis \(2014\)](#). Congestion externalities affect counterfactual spatial allocation of labor according to the following equation:

$$\omega_{ij}^{cm'} = \frac{(\widehat{L}_j^c)^{-ce} \widehat{V}_j^m \omega_{ij}^{cm}}{\sum_{k,n} (\widehat{L}_n^k)^{-ce} \widehat{V}_n^k \omega_{in}^{ck}},$$

where  $ce = 0.3$ .

With these two changes to the model, we consider the same counterfactual of removing the IHS and decompose this value into its domestic and international components. The first row of **B1** reports the baseline results for comparison; the second row reports the results allowing for agglomeration and congestion forces. The aggregate results for the total value of the IHS are not substantially different. However, allowing for agglomeration and congestion increases the value of the domestic component and decreases the value of the international component. These results are intuitive since agglomeration and congestion forces affect labor and productivity in US counties but not in other countries. Removing the IHS distorts the spatial distribution of labor and productivity levels in the United States, which increases the losses from removing the IHS for the domestic portion of trade.

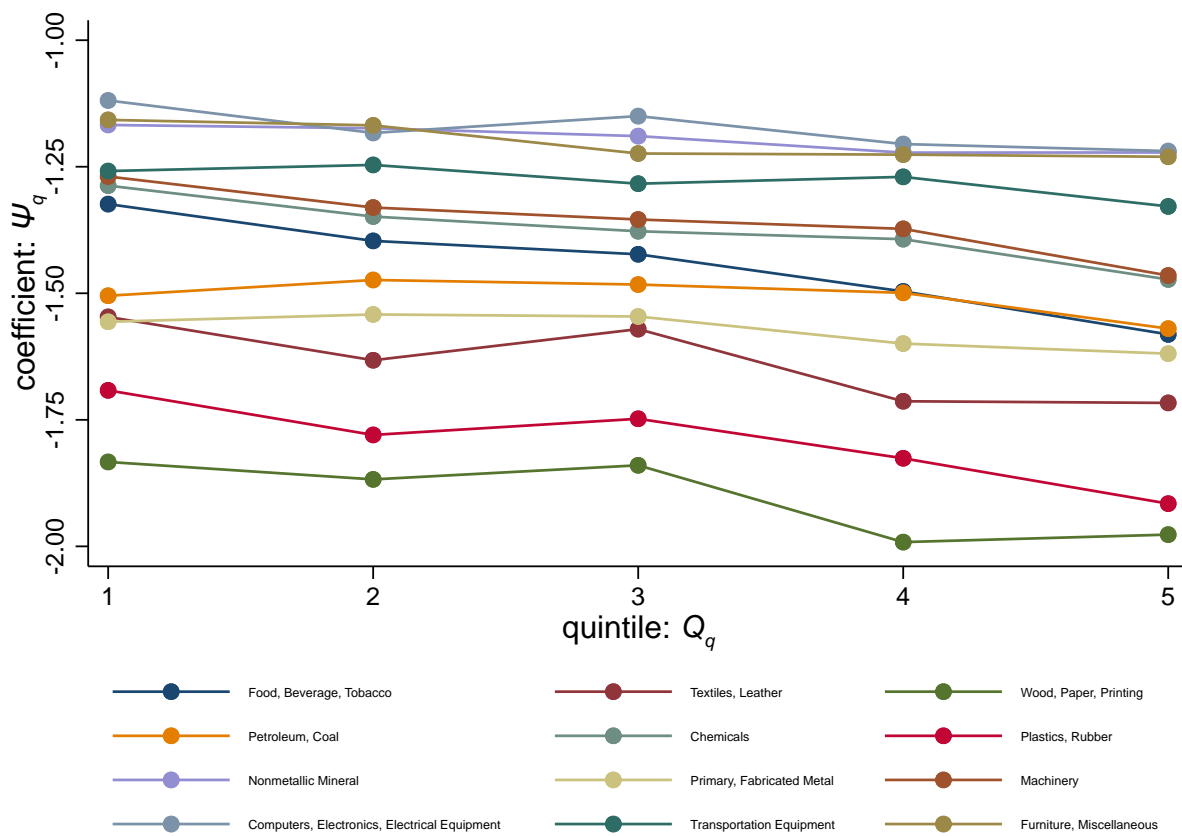
**Table B1: Total Losses from Removing the IHS with Agglomeration/Congestion**

	Total (1)	Domestic (2)	International (3)
Baseline, No IHS	601.4	445.9	155.6
With Agglomeration/Congestion	601.7	448.1	153.6

*Notes:* The table shows additional results from counterfactual exercises removing the Interstate Highway System. Column 1 shows the reduction in real GDP from removing the IHS for both the domestic and international components of trade costs. Column 2 shows the reduction in real GDP from removing the IHS domestic components of trade costs. Column 3 shows the difference between columns 1 and 2, which is the reduction in real GDP from removing the IHS foreign components of trade costs. The results in row 1 are for the baseline version of the model and row shows show results for a versions of the model with agglomeration and congestion.

## C Additional Tables and Figures

Figure C1: Results for International Distance Coefficients by Sector



*Notes:* This figure shows the results of estimating equation (30) for each sector  $s$ . The dependent variable are the trade flows between each US port  $r$  and each foreign trading partner  $j$ . Each line plots the coefficients associated with the quintiles of distance (in miles) for a given sector. All specifications include port and country fixed effects.

## D Details for Data Sources and Variable Construction

*Locations and Sectors:* We calibrate the model to domestic locations in the United States including 2,894 counties in 48 states and Washington, DC, using data from 2012 as the benchmark year. We exclude Alaska and Hawaii. The foreign locations are 35 countries (Australia, Austria, Belgium, Brazil, Canada, China, Cyprus, Czech Republic, Germany, Denmark, Estonia, Finland, France, Great Britain, Greece, Hungary, Indonesia, India, Ireland, Italy, Japan, Korea, Lithuania, Mexico, the Netherlands, Poland, Portugal, Romania, Russia, Slovakia, Slovenia, Spain, Sweden, Turkey, and Taiwan) and the rest of the world. Finally, we calibrate the model to 22 sectors, including 12 manufacturing sectors, 8 service sectors, construction, and combined wholesale and retail trade. Appendix Table D1 shows how we aggregate sectors from North American Industry Classification System (NAICS) to the sectors used in the empirical work.

*Domestic and International Trade Flows:* Data on domestic trade flows for the United States are drawn from the Commodity Flow Survey for 2012. We use this data to construct trade flows between US states as well as the domestic flow of exports from US states to foreign countries via US ports. Data on international trade flows are drawn from USA Trade Online for 2012. We use this data to construct trade flows between US states and foreign countries as well as between US ports and foreign countries. For domestic trade flows, the public use file for the 2012 Commodity Flow Survey is available for download at [this link](#). For international trade flows, data available for download or purchase from [USA Trade Online](#). International trade flows are drawn from the World Input-Output Database for 2012 (see [Timmer, Dietzenbacher, Los, Stehrer and De Vries, 2015](#)).

*Employment, Output, and Migration:* Employment and annual payroll data are drawn from the County Business Patterns for 2012. Migration data are drawn from the Internal Revenue Service for 2011-2012. The employment and payroll data can be downloaded at [this link](#). The migration data can be downloaded at [this link](#).

*State Production and Consumption Shares:* Value added in gross output shares, intermediate input shares, and Cobb-Douglas consumption shares are constructed from data drawn from the County Business Patterns for 20102 and World Input-Output Database for 2012 (see [Timmer, Dietzenbacher, Los, Stehrer and De Vries, 2015](#)). The World Input-Output Database can be downloaded [here](#).

*Transportation Network Database and Trade Costs:* The domestic and international transportation network is based on the US highway network—for routes between locations within the United States (i.e., counties and ports)—and international shipping—for routes between US ports and foreign countries.

Each location (i.e., counties, ports, countries) is represented as a centroid. Locations are connected via the transportation network which includes the US highway network from the US De-

partment of Transportation ([download here](#)), navigable waterways providing access to inland ports from the National Transportation Atlas Database ([download here](#)), international shipping lanes digitized from the CIA World Factbook ([download here](#)), and international transit between the United States and Canada or Mexico. The US highway network is comprised of all major roads including IHS segments, other federal-aid highways, and state highways. We assign travel speeds of 70, 55, and 45, respectively, to these portions of the US highway network. In addition, to ensure that all county and port centroids are connected to the highway network we build a network of “access roads” that provide direct connections. We assign a travel speed of 10 to the access road network.

To construct benchmark domestic and international trade costs we use ArcGIS to find the least cost route between centroids via the transportation network. In particular, for the domestic trade cost components, we use the network analyst tool to find the route between any pair of US counties or between US counties and US ports that minimizes *travel time*. These are used to the construct the interstate, intrastate, and state-to-port trade costs components. For the interstate trade cost component, for each state pair we identify the least cost route between CBSA’s in each origin destination pair. When computing the cost minimizing intertate route, we identify the county where the route exits the origin state and the county where the route enters the destination state. We use the exit and entry counties as the aggregation points to construct the intrastate trade costs. Intrastate trade costs are then constructed by measuring the travel time and distance from an origin county to the exit county or, from the entry county to the final destination county. Because the entry/exit counties will differ for each interstate trading pair, the intrastate trade costs are specific to the origin and destination pair. Similarly, for the state-to-port trade cost component, we find the minimum travel between US states and US ports. For international trade costs, we use the network analyst tool to find the route between US ports and foreign trading partners that minimizes *travel distance*.

To construct counterfactual domestic trade costs we again use ArcGIS to find the least cost route corresponding to the interstate, intrastate, and state-to-port trade cost components, after removing the a segment or several segments of the US highway network. In some cases we include or exclude segments from particular counterfactuals. For example, for the counterfactual removing I-95 from the highway network, we exclude I-95 and all associated loops and spurs of I-95 from the network. For each counterfactual we then find the route that minimizes the travel time and correspond to each of the domestic trade cost components.

To account for congestion in the benchmark and all counterfactual scenarios we use ArcGIS to identify the which of the roughly 331,000 pieces of the US highway network are used for particular interstate routes. Each piece of the highway network has a tabulated annual daily traffic entry based on data collected by the Federal Highway Administration and used to construct level of service. We use this data for the benchmark scenario to quantify the relationship between the level of service and

observed trade flows. For the counterfactuals, we then use the estimated relationship between level of service and observed trade flows in the benchmark scenario to assign trade-generated traffic and the corresponding level of service to the relevant pieces of the highway network for counterfactual routes.

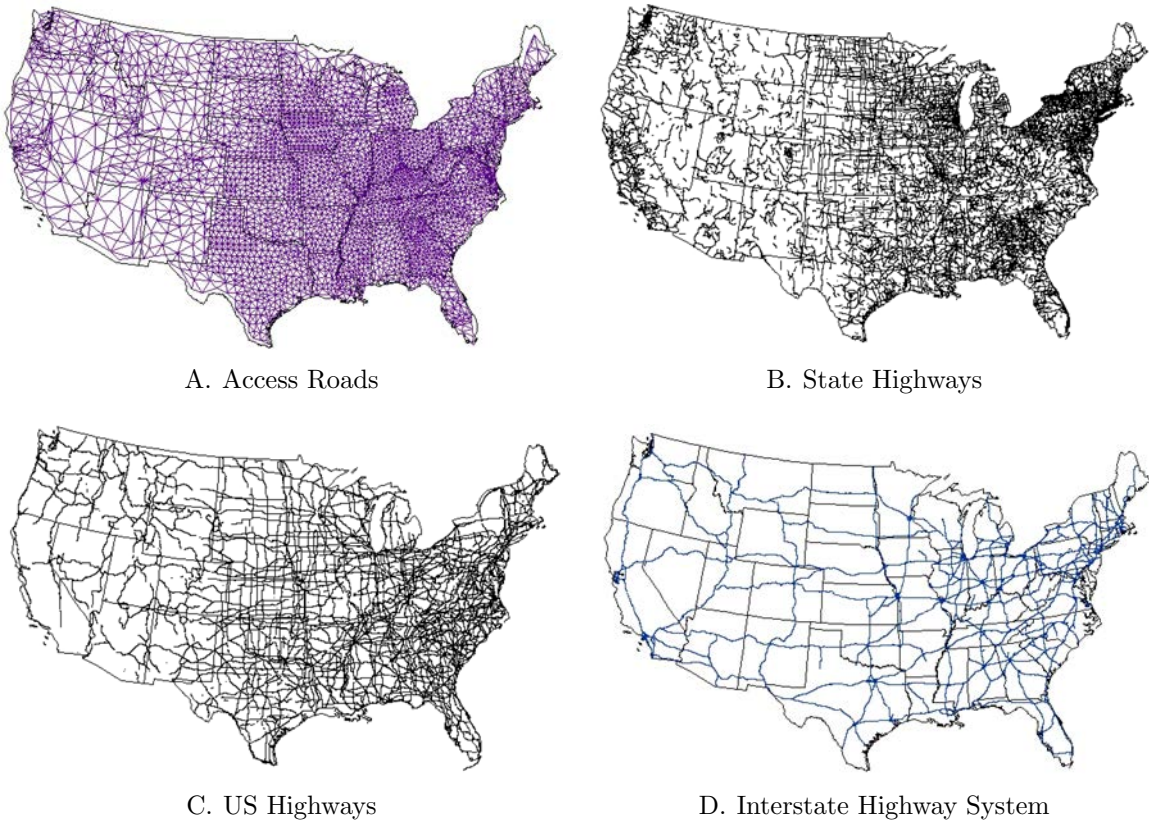
**Table D1: Aggregation of NAICS Sectors**

Sector	Name	NAICS	WIOD
1	Food, Beverage, and Tobacco Products	311-312	5
2	Textile and Leather Products	313-316	6
3	Wood Products, Paper, Printing, and Related Products	321-323	8-9
4	Petroleum and Coal Products	324	10
5	Chemical Products	325	11-12
6	Plastics and Rubber Products	326	13
7	Nonmetallic Mineral Products	327	14
8	Primary Metal and Fabricated Metal Products	331-332	15-16
9	Machinery	333	19
10	Computer, Electronic Products, Electrical Equipment	334-335	18
11	Transportation Equipment	336	20-21
12	Furniture and Related Products, and Misc.	337-339	22
13	Transport Services	481-488	31-34
14	Information Services	511-518	37-40
15	Finance and Insurance Services	521-525	41-43
16	Real Estate Services	531-533	44
17	Education Services	61	52
18	Health Care Services	621-624	53
19	Accommodation and Food Services	721-722	36
20	Other Services	493, 541, 55, 561, 562, 711-713, 811-814	54-51
21	Wholesale and Retail Trade	42-45	28-30
22	Construction	236	27

*Notes:* This table shows the aggregation of the industries used in this paper based on the North American Industrial Classification and World Input-Output Database.



**Figure D1: Components of the US Highway Network**



*Notes:* This figure shows the four components of the US highway network used to calculate travel time and trade costs. Panel A shows the access road network with assigned speed of 10 miles per hour, Panel B shows the state highway network with an assigned speed of 45 miles per hour, Panel C shows the US highway network with an assigned speed of 55 miles per hour, and Panel D shows the Interstate Highway System with an assigned speed of 70 miles per hour.