How Infrastructure Shapes Comparative Advantage

Luis Baldomero-Quintana*

November 12, 2019

Please do not circulate

Most recent version: Google Drive, GitHub.

Abstract

This paper studies the effects of domestic trade costs on comparative advantage. I build a model of international trade and internal geography that considers both international shipping routes and input-output linkages. I use the model to simulate how a large infrastructure project, Ruta del Sol, affects the specialization of Colombia, a country whose exports are highly concentrated in the mining sector. This road improves the access to global markets for both mining and manufacturing regions. To quantify the model, I use customs administrative records, a transportation survey, and geospatial data generated from both physical and digital road maps. My results indicate that the road project weakens the comparative advantage of Colombia in mining. Lastly, I provide evidence that the change in comparative advantage is larger when industry linkages are considered, because access to tradable intermediate inputs benefits more the manufacturing sector relative to the mining sector. The results demonstrate that domestic trade costs shape national comparative advantage, beyond the elements typically analyzed in the literature such as labor, capital, technology and institutions.

*Department of Economics, University of Michigan, luisbq@umich.edu. I am grateful to my advisors Hoyt Bleakley, Andrei Levchenko, Jagadeesh Sivadasan and Sebastian Sotelo for their invaluable advice, patience and guidance. I am thankful for the comments and suggestions of Maria Aristizabal, Alberto Arredondo, Vybhavi Balasundharam, Barthelemy Bonadio, Brian Cevallos, Benjamin Glass, Xing Guo, Meera Mahadevan, Luis Monroy-Gomez-Franco, Pinghui Wu, and seminar participants at the University of Michigan's International Economics Seminar, Economic Development Seminar, Trade Breakfast Student Workshop, Macroeconomics Lunch and H2D2 seminar. The staff of the Stephen S. Clark Library at the University of Michigan provided technical advice with historical maps and GIS software, particularly Nicole Scholtz and Tim Utter. I am also thankful to Ministerio de Transporte, DANE and INVIAS for providing access to data. This work was supported with funding from the University of Michigan Library and the Michigan Institute of Teaching and Research in Economics (MITRE). Jordan Hewitt, Jack Pedlow and Callista Zingas provided excellent research assistance.

1 Introduction

Comparative advantage is a fundamental idea in international trade theory. Standard trade models typically examine the role of technology, institutions, and factor endowments such as land, labor, and capital to explain the sources of specialization. However, this approach is limited by the fact that we only observe the comparative advantage of regions well connected to the global markets. This is specially true for developing nations, as the quality of infrastructure varies within developing nations (Oxford Economics, 2017; IADB, 2013). Therefore, to have a more comprehensive view of comparative advantage, it is necessary to consider the spatial distribution of domestic trade costs.

This paper shows that domestic trade costs are determinants of comparative advantage in a developing country context. As new infrastructure projects change the structure of the national transportation network or how industry linkages propagate shocks across regions and sectors, it is necessary to use a quantitative model to understand the mechanisms by which changes in domestic trade costs affect comparative advantage. Therefore, I build an international trade and internal geography model with input-output linkages, road networks, and international shipping routes. I use the model to understand the effects of completing a large infrastructure project currently in construction (*Ruta del Sol*) on the comparative advantage of Colombia. I show that the completion of the project increases the share of manufacturing exports and reduces the share of mining exports. Therefore, the highway project shifts the comparative advantage of Colombia away from the mining sector and towards manufacturing products.[1]

Colombia is an ideal context to analyze the impact of infrastructure on comparative advantage because the country is similar to several developing nations in a few dimensions. First, Colombia's exports are concentrated in a few goods, particularly mining products. This concentration of exports has been common among developing economies. Second, there is variation in the access to global markets among Colombian departments.[2] In many developing countries there is a similar situation, with regions with excellent access to global markets, while at the same time there are regions that are almost isolated due to poor infrastructure. Third, there is heterogeneity in the comparative advantage of Colombian regions, Many large middle-income nations share this characteristic.

Informed by the previous facts, I develop a framework in which departments in Colombia trade with each other and with the rest of the world. The model includes input-output linkages between three tradable sectors (agriculture, mining, and manufacturing) and a non-tradable sector (services). This characteristic allows trade costs to affect both output prices and production costs. Lastly, I include a realistic transportation feature: the existence of different shipping routes when departments and the rest of the world trade with each other. The model produces a tractable expression for the international trade flows between a department and the rest of the world, that use specific ports of exit or entry (a department-port gravity equation).

[1] Throughout the paper, I measure the comparative advantage of Colombia in a sector by using the share of exports in such sector. This works as a proxy to measure comparative advantage because, when the Balassa Index of Revealed Comparative Advantage is used for small open economies and highly aggregated sectors, the denominator of the index is fixed. French (2017) documents that revealed comparative advantage is useful to analyze the patterns of comparative advantage for different economies.

[2] A department is the official administrative region of Colombia.

To take the model to the data, I combine four data sources: detailed customs administrative data with information about the port of exit or entry, a survey of transportation flows, and geospatial data that I create using digital and scanned physical road maps. The customs data allow me to obtain international trade flows between departments and the rest of the world, with information about the port used for exit or entry. The transportation survey allow me to obtain a proxy of domestic sectoral trade flows. Finally, using the geospatial data and the Dijkstra's algorithm, I obtain travel times between any location within Colombia for both modern and historical road networks.

There are two parameters that govern my model. The first parameter defines the relationship between trade costs and travel times, and the second parameter defines the heterogeneity of the use of shipping routes for goods traded between Colombian regions and the rest of the world. To recover the values of these parameters, I estimate a department-port gravity equation using an instrumental variable approach. My instrument is the distance between locations using historical road networks during periods in which the characteristics of the Colombian economy were very different compared to the current economic circumstances (this approach is similar to Baum-Snow, 2007; Michaels, 2007; Duranton, Morrow, and Turner 2014; and Duranton, 2015). After obtaining the value of the parameters of my model, I run counterfactual simulations.

My main counterfactual experiment considers the effects of the infrastructure program *Ruta del Sol* on the sectoral exports of Colombia. The project's objective is to modernize the highway that connects Cundinamarca with the Atlantic seaports. This department is the main exporter of manufacturing and agricultural products. Given the structure of the road system in Colombia, *Ruta del Sol* also improves substantially the access to international markets for several departments that specialize in the mining sector. Hence, the expected effect of this highway project in the national sectoral exports is unclear a priori. Additionally, given the structure of the input-output linkages in Colombia, the benefits of the reduction in domestic trade costs propagate in such a way that one sector benefits more than others.

The results of my counterfactual experiment show that the completion of the infrastructure project increases the share of manufacturing exports by four percentage points. To understand the importance of this result, note that for the past three decades, the share of mining exports of Colombia has grown substantially. This result implies that the road project can potentially reverse the upward trend of the specialization of Colombia in mining goods and shift the comparative advantage of Colombia towards the manufacturing sector. This result does not imply that the non-manufacturing exports fall, but rather the manufacturing exports grow more than the exports of other sectors. So, the infrastructure project can potentially revert any existing crowd-out effects of the mining boom on the Colombian manufacturing sector, due to potential Dutch disease effects (Alcott and Keniston, 2018).

To analyze the main forces driving my results, I run alternative counterfactual exercises, in which I isolate the different effects of *Ruta del Sol*. I consider separately the effects of the road project on domestic trade costs, international trade costs, and on both domestic and international trade costs without including input-output linkages. My alternative simulations show that industry linkages help to increase the manufacturing exports substantially. When I simulate the effects of *Ruta del Sol* without industry linkages, the increase in the share of manufacturing exports is one third of the growth observed in my main counterfactual experiment, which does consider these linkages. This is due to the fact that the manufacturing sector benefits more from access to tradable intermediate inputs, compared to the mining sector.

My work contributes to the international trade literature on the determinants of comparative advantage. My main contribution is to show that domestic trade costs are a source of national comparative advantage. This finding is specially relevant in developing countries where domestic trade costs are high, thereby generating differences in regional access to global markets within a country (Atkin and Donaldson, 2015). To my knowledge, recent international trade literature has provided little attention to the direct link that exists between the spatial distribution of domestic trade costs and national comparative advantage.

The main idea of this paper, how internal trade costs shape comparative advantage, is related to Deardoff (2014). His main idea is that the transportation costs of an economy to those countries that are geographically close, matter for its comparative advantage. He defines the term *local comparative advantage*, which measures comparative advantage considering transportation costs. With this term, it is possible to explain situations in which a country has a comparative advantage in a specific sector, even though his production costs are high. In such case, the comparative advantage exists due to low transportation costs between the economy and its neighboring nations. In my framework, I focus exclusively on how the comparative advantage of a country is shaped exclusively by its internal transportation costs, while Deardoff (2014) focuses on transportation costs to the neighboring economies.

The closest work to this paper, is Duranton, Morrow, and Turner (2014) and Duranton (2015). These papers use applied microeconomics methods to show that urban centers with better infrastructure can specialize in sectors that produce heavy goods. Different from these papers, I focus on how roads affect specialization at a national level. Additionally, I depart from such work by using an international trade model to run counterfactual scenarios that examine how a large infrastructure project can change comparative advantage. Besides, my theoretical framework considers the role of industry linkages.

Other work regarding the determinants of comparative advantage includes papers regarding how migration affects specialization (Pellegrina and Sotelo, 2019; Arkolakis, Lee and Peters, 2018; Bahar and Rapoport, 2018; Morales, 2019), how the quality of institutions is a source of comparative advantage (Levchenko, 2007) or how domestic trade costs influence crop choices in developing countries (Allen and Atkin, 2018; Morando, 2019). This paper also speaks to the theoretical research regarding the dynamics of comparative advantage (Matsuyama, 1992; Krugman, 1987; Levchenko and Zhang, 2016; Hanson, Lind, and Muendler 2015).

In economics, there is an increasing interest in the effects of infrastructure projects. This includes work on how infrastructure improvements affect either domestic outcomes, or trade flows between a country and the rest of the world (Alder, 2019; Allen and Arkolakis 2019; Coatsworth 1979; Cosar and Demir, 2016; Donaldson 2018; Donaldson and Hornbeck, 2016; Faber, 2014; Fogel 1962; Holl, 2016; Xu, 2016, Xu 2018). To my knowledge, only two papers consider jointly domestic outcomes and international trade: Fajgelbaum and Redding (2018) on the structural transformation of Argentina during the early 20th century, and Sotelo (2019) on how roads affect agricultural trade in Peru. I depart from the existing literature by highlighting the role of industry linkages when I examine the effects of infrastructure in economic outcomes. Specifically, I show that input-output linkages propagate the effects of lower domestic trade costs. Although the previous work on infrastructure considers the effects of domestic trade costs on exports by sector, the interactions between industry linkages and infrastructure have not been examined in detail. However, to understand the impact of infrastructure on sectoral exports, it is crucial to consider the industry linkages. This is because the existence of such linkages generates uneven effects of changes in domestic trade costs on exports across sectors.

Lastly, my results are relevant for the literature of the Dutch disease. The work of this topic has been extensive, as Van der Ploeg (2011) points out. My results show that improving the domestic integration of regional markets in specific ways can minimize the specialization of an economy in a single sector. More precisely, my paper is closely related to one of the main mechanisms of the Dutch disease, the crowd-out of the manufacturing sector after a resource boom, Allcott and Keniston (2018). In my case, changes in transportation infrastructure have the potential to generate improvements in the manufacturing sector. In a country in which industry linkages are such that the access to intermediate inputs has a major impact on the costs of the manufacturing sector, specific improvements in transportation can offset the crowd-out of the manufacturing sector caused by a commodity boom.

The rest of the paper is organized as follows. Section 2 describes the data and provides motivating facts. Section 3 presents the model. Section 4 describes how I take the model to data. Section 5 reports the results of my counterfactual exercises. Section 6 concludes.

2 Data and basic patterns

2.1 Data

This paper combines five datasets that allow me to measure domestic sectoral trade flows between Colombian departments, international trade flows between departments and the rest of the world by sector and port of exit/entry, input-output linkages, domestic trade costs, and international trade costs. My analysis focuses in four sectors (agriculture, mining, manufacturing and services) and considers data for 2013.

Customs data. I use a dataset created by the National Directorate of Taxes and Customs (DIAN, in Spanish) and the National Administrative Department of Statistics (the official statistical agency of Colombia, or DANE in Spanish) that contains all the shipments of exports and imports of Colombia. The data includes information such as harmonized system code, the department of origin/destination, and the city-port of exit/entry.[3]

Transportation and geography. I create a fully digitized road network that represents the primary highway system of Colombia,[4] based on physical and digital maps of the Ministry of Transporta-

[3] I define a city-port as the location through which the products exit/enter the country. In the customs data, there is a total of 19 city-ports that are actively used for international shipments. The use of a city-port is based on the fact that goods could exit via a specific city, through different methods. For example, firms could use the seaport or the international airport located in Cartagena. In such cases, I do not differentiate by the method of transportation. Hence, in this example I would define Cartagena as a city-port of exit.

[4] Given that the transportation of goods mainly occurs via trucks, I do not consider the secondary road system (composed by roads administered by the Departments) nor the tertiary road system (managed by municipalities) and I focus exclusively in the primary road system. I do this because I do not have the status of the secondary or tertiary

tion and the National Institute of Roads (INVIAS). My main analysis focuses on roads, given that the share of total shipments (measured in tons) shipped via road is 73%, as of 2013 (ANIF, 2014).[5] For each highway segment, I have information on whether the road is paved, if it crosses a city, and whether the road is under public management or administered by a public-private partnership via the legal figure of *concesion*. Roads under the legal status of *concesion* are paved and tend to have better geographical and topographical characteristics than the rest of the roads.[6]

I estimate the travel times using Dijkstra's algorithm. I assign a speed of 30 km/hour for unpaved roads. The speeds for paved roads are 50 km/hour for paved roads in urban areas, 80 km/hour for paved highways outside urban centers, and 100 km/hour for paved roads under the legal figure of *concesion*. The speeds for paved and unpaved roads are like the ones used by Allen and Atkin (2016) for the Indian highway system, with the difference that I define different speeds for paved roads under *concesion*. I describe in the Appendix A why I consider the roads under *concesion* to be of higher quality, which leads me to assign them higher speed values.

Survey of cargo flows. I use the 2013 Survey of Origin/Destination of Cargo Transportation of the Ministry of Transportation to obtain proxies of domestic trade flows for the agricultural and manufacturing sectors. Specifically, I use the data on total weight cargo flows between different Colombian locations, measured in metric tons. Additionally, I use data regarding oil production and refining from the Ministry of Energy and Mines and the public oil company Ecopetrol, to generate domestic trade flows for the mining sector.

Input-output linkages. Data to calibrate the parameters of input-output linkages come from two sources: the World Input-Output Table of 2013 (Timmer, Dietzenbacher et al., 2015) and Colombia's input-output table produced by DANE for the year 2010.

2.2 Motivating facts

This section describes four empirical facts about Colombian departments that motivate the theoretical framework. First, Colombian exports are concentrated in a few goods, mostly mining ones. Second, the Colombian departments specialize in different sectors. Third, departments differ in their access to international markets, which generates differences in the international trade costs between departments and the rest of the world. Lastly, when the departments trade with the rest of the world, they do not use a single city-port to trade.

Fact 1, Colombian exports are concentrated in a few goods. Figure 1 plots the share of exports of *traditional products* as a fraction of total exports. This category was created by the Colombian government agencies for specific goods, given the historical concentration of exports in these prod-

roads. Moreover, there are maps elaborated by the Ministry of Transportation, which contains graphical data about the annual flow of trucks by road. These maps show that most of the truck traffic use the primary road system. See IGAC (2005) for the most recent maps regarding truck flows across the country.

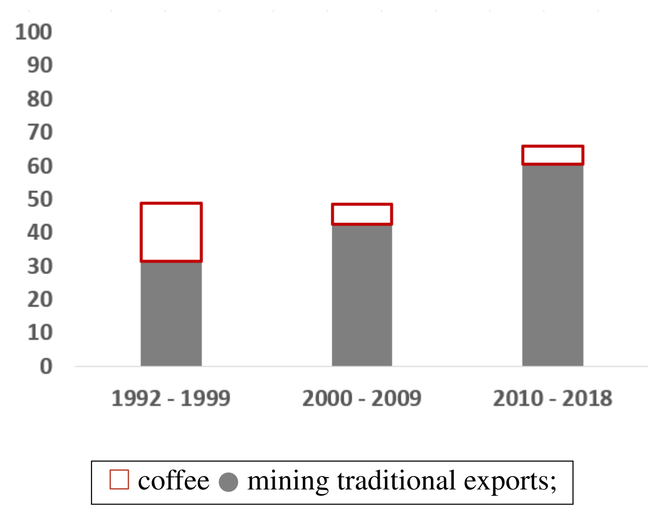
[5]The use of fluvial shipments is very limited, the railroad network is used exclusively for a specific route for the transportation of commodities, and the use of air cargo for domestic trade is relatively small (Duranton, 2015)

[6]Pachon and Ramirez (2006) explain that since the mid-90s, the Colombian government partially privatized some segments of the primary road system under the legal figure of public-private partnerships (concesiones, in Spanish). These roads were renovated/built by private companies, and the payments are split in two types: a direct government payment and the income generated by charging a fixed-fee to users of the highways).

ucts.[7] As figure 1 shows, during the past three decades, Colombia experienced an upward trend in the specialization of mining goods.

Colombia was considered the standard case of an agricultural commodity-dependent nation by international agencies due to its dependence on coffee exports (FAO, 2002). More recently, an oil boom has reduced the share of coffee in the national exports. Recent official documents elaborated by the Colombian government highlight the dependence of the country on commodity exports (DNP, 2019).

Figure 1. Share of "traditional exports" according to Colombia's statistical agency DANE (%)



Notes: The bars show the average annual share of "traditional exports" with respect to total exports, for the period indicated in the x-axis. The source of the data is the official website of DANE.

Fact 2, Colombian departments specialize in different sectors. Using customs data from 2013, I build a Regional Index of Revealed Comparative Advantage (RCA) for every department. My objective is to show how a department specializes in a sector, relative to the specialization of Colombia in this same industry. The formula of this index is

$$RCA_{s,d} = \Big(\frac{Exports_{s,d}}{Total\ Exports_d}\Big) / \Big(\frac{Exports_{s,Colombia}}{Total\ Exports_{Colombia}}\Big)$$

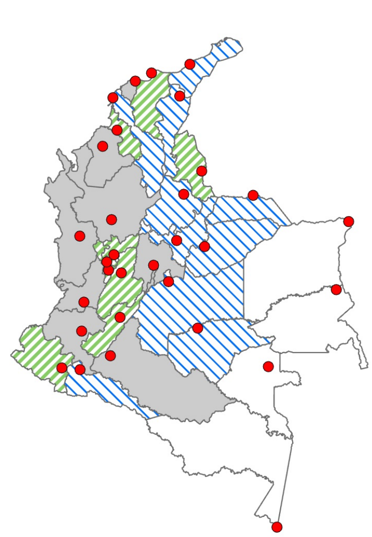
where s stands for a sector and d is a department. The index is the proportion of the exports of a department in sector s, divided by the proportion of Colombia's exports in industry s.

Intuitively, if the value of this ratio is high, a department is more specialized in sector s relative to the level of specialization of the entire Colombian economy in this industry. To obtain the Balassa Index of Revealed Comparative Advantage of every region (Balassa, 1966), the Regional index needs to be multiplied by the Balassa Index for Colombia. I use a regional index, instead of the Balassa index because I want to measure how every region is different than the Colombian economy, in its trade with the rest of the world.

[7]This term is commonly used by government agencies such as the National Department of Planning or the statistical agency DANE. It groups the following products: coal, oil, coffee, and nickel-alloy.

After I obtain the values of the index, I select the sector in which every department shows the highest level of specialization. With this information, I construct figure 2 to provide evidence that there is variation in the sectoral specialization of Colombian regions.

Figure 2. Map indicating the sector with the strongest comparative advantage of every department (highest value of the Balassa Index)

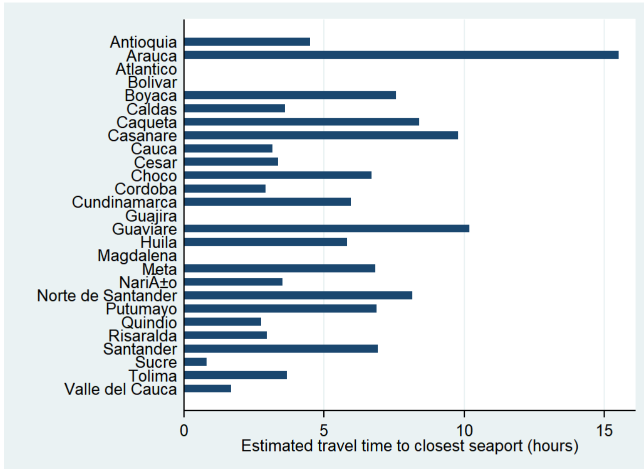


//// agriculture ; \\\\ mining; ■ manufacturing; • city-port

Notes: I do not consider the departments of Guainia, Leticia, San Andres y Providencia, Vaupes, and Vichada. Additionally, I merge Bogota with the department of Cundinamarca. See Appendix A for more details.

Fact 3, Colombian regions do not have uniform access to international markets. Colombian departments have heterogeneity in their access to global markets, given the existing geography of the country and the structure of the transportation network. To show this, Figure 3 displays the estimated travel times between the capitals of every department and the seaports of the country. (Given that 86% of exports and 70% of imports in 2013 exit or entered the country via seaports, figure 3 helps to illustrate the access to international markets of every Colombian department). The figure illustrates how some departments have immediate access to seaports, while for others it takes more than five hours to reach these ports.

Figure 3. Estimated travel times between the capital of the department and the closest seaport



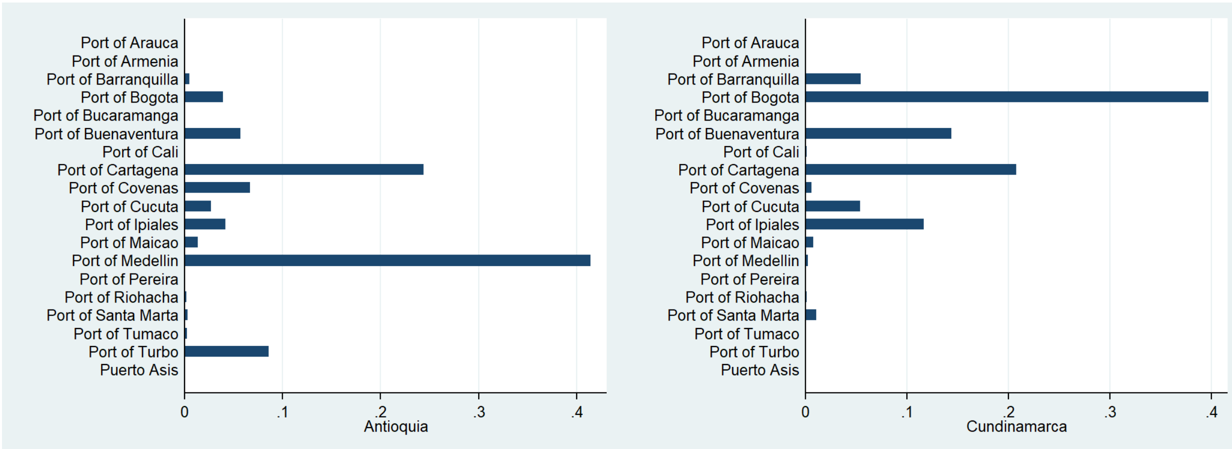
Notes: I estimate the travel times between the capital of every department and the closest seaport using Dijkstra's algorithm, according to the speed values described in section 2.1. I do not consider the departments of San Andres y Providencia, Guainia, Leticia, Vichada and Vaupes. See Appendix A for details about this.

Fact 4, Colombian departments use multiple ports to trade with the rest of the world. Several departments have enough logistical infrastructure to trade with the rest of the world, such as airports, international land bridges, and seaports. In spite of this, most of the firms in the departments use different city-ports to trade with the global markets.

Figure 4 shows that the goods exported by the largest two departments of Colombia (Cundinamarca and Antioquia) are sent to other countries via different city-ports, even though Cundinamarca and Antioquia have large city-ports to serve international trade shipments.[8] The main explanation for this is that every city-port has logistical advantages for the shipment of specific goods, even within the same sector. For example, if I look at manufacturing goods, the seaport of Covenas is ideal for naphta products (a chemical manufacturing good), the airport of Bogota has excellent logistical conditions for the shipment of textiles, while the seaport of Santa Marta has very good logistical capacity for handling steel and cement products.

[8]The city of Bogota located in the department of Cundinamarca posses the largest airport in the country, El Dorado International Airport, which has capacity to handle cargo shipments. The city of Medellin located in Antioquia has the Jose Maria Cordova International Airport, which also has infrastructure for the shipment of cargo.

Figure 4. Use of city-ports to export goods by the largest two Colombian departments ($\%$ of total department exports)



Notes: The vertical axis considers the 19 city-ports included in the customs data. For more details about the city-ports, see Appendix A.

3 Model

In this section, I describe my theoretical framework, define the equilibrium concept, provide an expression for a gravity equation, and explain how to translate changes in the road system into changes on the trade costs.

3.1 General framework

Geography. Consider an economy composed of Colombian departments and the rest of the world. These locations trade with each other. The departments are indexed by d and the rest of the world is indexed by RoW. The set of Colombian departments is $D = \{1, ..., \bar{d}\}$ and the set of all locations is $Z = \{1, ..., \bar{d}, RoW\}$. Each location is indexed by subscripts $n, j \in Z$. Trade between departments and the rest of the world require the use of city-ports ρ (see figure 5). There is a total of $\bar{\rho}$ city-ports. The set of city-ports is $\mathbb{P} = \{1, 2, ..., \bar{\rho}\}$.

I define an international shipping route as an ordered pair that consists of a department d and a city-port ρ. An *export route* consists of an ordered pair department, city-port $r_x = (d, \rho)$. There is a total of $\bar{d}\bar{\rho}$ export routes. The set of export routes is $R_x = D \times \mathbb{P}$. The subset of export routes for a department d is defined as $R_{x,d} = \{(d, \rho) : \rho \in \mathbb{P}\}$. An *import route* consists of an ordered pair city port-department $r_m = (\rho, d)$. There are $\bar{d}\bar{\rho}$ import routes. The set of import routes is $R_m = \{\mathbb{P} \times D\}$. The subset of import routes for a department d is defined as $R_{m,d} = \{(\rho, d) : d \in D\}$

Goods. There are two types of goods, intermediates and composite goods. There are four sectors in the economy: agriculture (a), mining (m), manufacturing (i) and services (z). Sectors are indexed by $k \in \{a, m, i, z\}$. Intermediate good firms in location n and sector k produce intermediate good.

Firms that produce composite goods buy from suppliers across different locations and produce an aggregated composite using a Dixit-Stiglitz aggregator. The market structure in all sectors is perfect competition.

Trade costs between departments and the rest of the world. International trade between a department, d, and the rest of the world, RoW, require specialized traders, as in Allen and Arkolakis (2019). There is a continuum of specialized traders $\iota \in [0,1]$. Traders choose among all the shipping routes when they export or import goods.[9] These traders face capacity constraints when moving goods internationally.

Figure 5 helps to understand the concept of international shipping routes. For example, when department 1 trades with the rest of the world, *Route 1A* can be used, which implies that the shipment of goods occurs via *Port A*. Or *Route 1B* can be used, therefore, the *Port B* will be chosen for the shipment of goods. A similar logic occurs when department 2 trades with the rest of the world.

Every specialized trader faces a productivity shock that is specific to the international shipping route and to every sector k. This implies that the cost of a specialized trader ι when it uses an international shipping route r_t is $\tau_{r_t,k}/z_{r_t,k}(\iota)$. I define the international shipping cost for trader ι as the lowest international shipping cost across different routes, when the trader ships a good between department d and RoW, that is

$$\tau(\iota) = \underset{r_t}{min}\, \frac{\tau_{r_t,k}}{z_{r_t,k}(\iota)} \text{ for } t \in \{x, m\} \tag{1}$$

where τ_{r_t} is the shipping cost along route r_t for goods of sector k, $z_{r_t,k}(\iota)$ is the productivity draw for a specific international shipping route r_t to transport goods of sector k, and subscript t defines whether the shipping route is used to export or import goods. This productivity draw follows a Frechet distribution with parameters $(A_{r_t,k}, \theta_k)$. The Frechet parameter $A_{r_t,k}$ is the scale parameter of the Frechet distribution. The shape parameter θ_k represents the heterogeneity of productivities of city-ports regarding the transportation of sector-k goods. The higher the value of θ_k, the lower the heterogeneity in the productivities of city-ports. Thus, high values of θ_k imply that traders tend to use the same city-port to move goods between departments and the rest of the world.

When agents buy exported or imported goods, they are randomly assigned with specialized traders. Thus, the iceberg trade cost between a department d and the rest of the world RoW is the expected trade cost across the continuum of traders, as in Allen and Arkolakis (2019).

$$\tau_{dRoW,k} \equiv E[\tau(\iota)] = E\Big[\underset{r_t,k}{min} \frac{\tau_{r_t}}{z_{r_t,k}(\iota)}\Big] \tag{2}$$

Using the properties of Frechet distribution, the expression for the iceberg trade cost between any

[9] Intuitively, firms choose logistical companies to ship goods between a department and the rest of the world (e.g. Fedex, UPS, McLane Company, JR Freight, etc.)

department d and the rest of the world becomes

$$\tau_{dRoW,k} = \Phi_x^{-\frac{1}{\theta}} \Gamma\Big(\frac{1+\theta_k}{\theta_k}\Big) \text{ where } \Phi_k = \sum_{r_t} A_{r_t} \tau_{r_t}^{-\theta_k} \tag{3}$$

where Γ is the gamma function.

International shipping costs. Following Duranton, Morrow, and Turner (2014), I define the international shipping cost of route $r_t = (d, \rho)$ as $\tau_{r_t} \equiv \tau_\rho \tau_{d\rho} \tau_d$. This implies that the international shipping cost of a route depends on logistical characteristics of department d, denoted by τ_d, the logistical capacity of the port ρ, represented by τ_ρ, and the connectivity between department d and port ρ, expressed as $\tau_{\rho d}$. The latter is a function of the travel times between d and ρ, T_{dp}, therefore $\tau_{d\rho} = f(T_{d\rho})$.[10]

Trade costs between departments in Colombia. There are standard iceberg trade costs for every sector. I denote the trade costs between department $d_1 \in D$ and department $d_2 \in D$ for sector-k goods as $\tau_{d_1 d_2,k}$. Iceberg trade costs between departments are a function of travel times along the least cost route that connects these departments ($T_{d_1 d_2}$), that is $\tau_{d_1 d_2} = f(T_{d_1 d_2})$. Notice that in figure 5 there is only one route to move goods between department 1 and 2.

Domestic traders are homogeneous, hence they always choose the same optimal road when sending goods from d_1 to d_2. Implicitly, this implies that all the trade flows are shipped through the least cost road between d_1 and d_2. If I consider the existence of traders for domestic trade, this assumption can be interpreted as having a very high value for the shape parameter θ that represents the heterogeneity in the use of roads across two locations within Colombia. The assumption is consistent with Allen and Arkolakis (2019), who find that domestic traders moving goods across two cities within a country tend to choose the same least cost road.

Preferences. Consumers' preferences are represented by a Cobb-Douglas utility function given by

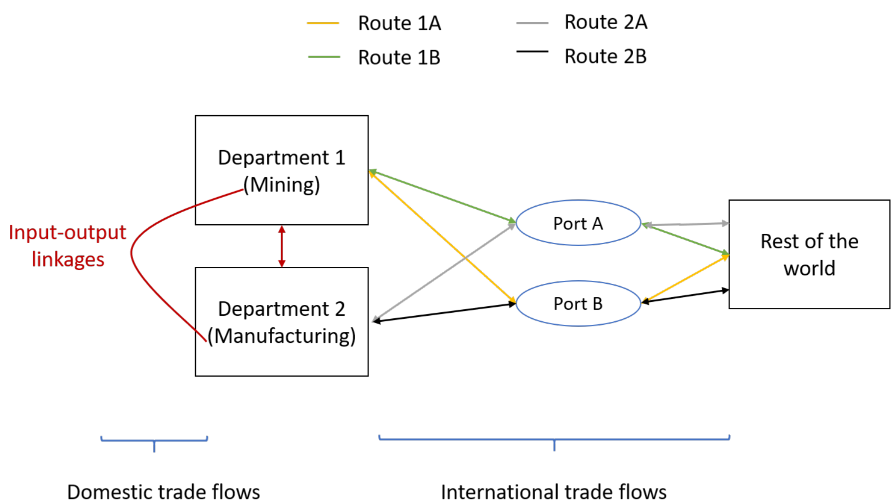
$$U_j = \prod_{k=1}^{K} (C_j^k)^{\alpha_j^k} \text{ , with } \sum_{k=1}^{K} \alpha_j^k = 1 \tag{4}$$

where α_j^k is the share of sector k in final demand and C_j^k is the level of consumption of the composite good. The income of households is denoted by I_n. Households' income are the sum of payments to labor and transfers, that is $I_n = w_n L_n + D_n$. The transfers are equal to deficits as in Dekle, Eaton and Kortum (2008).

Labor supply. Agents live in location $n \in Z$ and supply one unit of labor. There are L_n workers in location n. There is perfect labor mobility across sectors, but no labor mobility across locations (this implies no labor mobility across Colombian departments).

[10] The assumptions have implications about the symmetry of shipping costs of export and import routes $\tau_{r_x} = \tau_{r_m} = \tau_d \tau_{dp} \tau_\rho$, $if\ r_x = (d, \rho)\ and\ r_m = (\rho, d)$

Figure 5. Economic environment



3.2 Production

Production of intermediates. The production of intermediate goods requires labor and composite goods from all sectors. Technology has constant returns to scale and it is defined by

$$q_{n,k} = A_{n,k} l_{n,k}^{\beta_n^{l,k}} \left[\prod_{s\{a,m,i,z\}} m_{s,k}^{\beta_n^{s,k}} \right] \tag{5}$$

where $\beta_n^{l,k} + \sum_s \beta_n^{s,k} = 1 \, \forall \, n$. I denote by $m_{s,k}$ the amount of composite good of sector s used in the production of sector k, $\beta_n^{s,k}$ is the parameter that defines the share of composite goods from sector s used in the production of intermediates for sector k goods, $\beta_n^{l,k}$ is the share of value added of sector k, $A_{z,k}$ is the productivity of sector k, l_n^k is the amount of labor necessary for the production of good of sector k in city n,

Firms price at unit cost $\frac{c_{n,k}}{A_{n,k}}$, where $c_{n,k}$ is the unit cost of an input bundle. This can be expressed as

$$c_{n,k} = \phi_{n,k} (w_n)^{\beta_n^{l,k}} \prod_s (P_{n,s})^{\beta_n^{s,k}} \tag{6}$$

where $\phi_{n,k} \equiv (\beta_n^{l,k})^{-\beta_n^{l,k}} (\beta_n^{a,k})^{-\beta_n^{a,k}} (\beta_n^{i,k})^{-\beta_n^{i,k}} (\beta_n^{m,k})^{-\beta_n^{m,k}} (\beta_n^{z,k})^{-\beta_n^{z,k}}$ is a constant, and $P_{n,s}$ is the price of a composite intermediate good from sector s in location n. The cost function captures the input-output linkages between industries: if the price of the composite good in one industry changes, it will affect the unit cost of the rest of the sectors.

Production of composite goods. Firms that produce composite goods in location n for sector k purchase the intermediate goods from suppliers across different locations. The production technology of composite goods uses a Dixit-Stiglitz aggregator:

$$Q_{n,k} = \Big[\sum_j (q^c_{jn,k})^{\frac{\sigma_k - 1}{\sigma_k}} \Big]^{\frac{\sigma_k}{\sigma_k - 1}} \tag{7}$$

where $Q_{n,k}$ is the number of units that the firms supply, σ_k is the elasticity of substitution between intermediates of sector k and $q^c_{jn,k}$ is the demand of intermediate good of sector k by city n produced in city j.

Prices. Given the existence of perfect competition, the price of a good of sector k consumed by location n and produced in j considers the unit cost and the trade costs between locations, that is

$$p_{jn,k} = \frac{c_{j,k} \tau_{jn,k}}{A_{j,k}} \tag{8}$$

using this expression, I derive the price of the composite good of sector k in location n

$$P_{n,k} = \Big[\sum_j p_{jn,k}^{1-\sigma_k} \Big]^{\frac{1}{1-\sigma_k}} = \Big[\sum_j \Big(\frac{\tau_{jn} c_{j,k}}{A_{j,k}} \Big)^{1-\sigma_k} \Big]^{\frac{1}{1-\sigma_k}} \tag{9}$$

where the second equality comes from using (8). Using the previous prices of sector k, I can obtain the price index of location n:

$$P_n = \prod_k \Big(\frac{P_{n,k}}{\alpha_{n,k}} \Big)^{\alpha_{n,k}} \tag{10}$$

3.3 Trade flows and expenditure shares

Solving the optimization problem of the firms that produce the composite good, I obtain an expression for the demand of intermediate good in sector k, denoted by $q^c_{jn,k}$. Combining it with the price of intermediate good $p_{jn,k}$ and aggregating, I derive an expression for the total expenditure by location n on goods from sector k produced in location j

$$X_{jn,k} = \Big(\frac{\tau_{jn} c_{j,k}}{A_{j,k}} \Big)^{1-\sigma_k} Q_{n,k} P_{n,k}^{\sigma_k - 1} \tag{11}$$

Following Anderson and van Wincoop (2003), the trade flows equation can also be expressed as

$$X_{jn,k} = (\tau_{jn})^{1-\sigma_k} \left(\frac{Y_{j,k}}{\Pi_{j,k}^{1-\sigma_k}} \right) Q_{n,k} P_{n,k}^{\sigma-1} \tag{12}$$

where $\Pi_{j,k}^{1-\sigma_k} \equiv \sum_m \tau_{jm}^{1-\sigma_k} X_{m,k} P_{m,k}^{\sigma_k - 1}$. The term $X_{m,k}$ is the total expenditure of location m in goods of sector k. Finally, let $\lambda_{jn,k}$ be the fraction of expenditure of j in sector-k goods produced by location n:

$$\lambda_{jn,k} \equiv \frac{X_{jn,k}}{\sum_l X_{ln,k}} = \left(\frac{\tau_{jn} c_{j,k}}{A_{j,k}} \right)^{1-\sigma_k} (P_{n,k})^{\sigma_k - 1} \tag{13}$$

3.4 Total expenditure and trade balance

The total expenditure of location n in sector-k goods $X_{n,k}$ is composed by the expenditure by firms on intermediates (that depends on total exports of location n) and the households' expenditure (which is a constant fraction $\alpha_{n,k}$ of the total income):

$$X_{n,s} = \sum_k \beta_n^{s,k} \sum_j X_{j,k} \lambda_{nj,k} + \alpha_{n,s} I_n \tag{14}$$

where I_n denotes the total income of sector n, composed by labor income and transfers. The total income in location n is $I_n = w_n L_n + D_n$, where D_n is the total deficit of n.

The total trade deficits sum up to zero across all locations ($\sum_n D_n = 0$) and the total trade deficits are the sum of sectoral trade deficits, $D_n = \sum_k D_{n,k}$. A sectoral trade deficit $D_{n,k}$ is defined as $D_{n,k} = M_{n,k} - E_{n,k}$ where $M_{n,k} = \sum_j X_{n,k} \lambda_{jn,k}$ represents the total imports of country n of sector-k goods and $E_{n,k} = \sum_j X_{j,k} \lambda_{nj,k}$ is the total exports of n of sector-k goods. I consider total trade deficits as exogenous, but the sectoral trade deficits are endogenous, as in Caliendo and Parro (2015).

Considering the definition of total trade deficit for any location n, I can express the trade balance equation as

$$\sum_k \sum_j X_{j,k} \lambda_{nj,k} = \sum_k \sum_j X_{n,k} \lambda_{jn,k} - D_n \tag{15}$$

Labor market clearing. By aggregating the total expenditure of location n in sector k, equation (14), across all sectors and combining it with the trade balance equation (15), I get an expression for the labor market clearing (see Appendix B).

$$w_n L_n = \sum_k \beta_n^{l,k} \sum_j X_{nj,k} = \sum_k \beta_n^{l,k} \sum_j X_{j,k} \lambda_{nj,k} \tag{16}$$

3.5 Equilibrium

In this section, I define the world equilibrium. Then, I describe the equilibrium in changes, which requires fewer parameters than the original equilibrium. By doing this, I simplify the estimation procedure.

3.5.1 Equilibrium in levels

Definition 1. World equilibrium in levels. *The equilibrium is a set of wages $\{w_{n,k}\}_{n\in Z,k\in\{a,m,i,z\}}$, prices $\{P_{n,k}\}_{n\in R,k\in\{a,m,i,z\}}$, and labor allocations $\{L_{n,k}\}_{n\in Z,k\in\{a,m,i\}}$ for all locations $n \in Z$ under the assumption of perfect labor mobility across sectors and immobile labor across locations that solve equations (6), (9), (13), (14) and (15).*

3.5.2 Equilibrium in changes

Solving the previous equilibrium requires the knowledge of many parameters that are difficult to estimate, such as the sectoral productivities $\{A_{j,k}\}$. An option to reduce the number of parameters needed to calibrate the model, is to express the equilibrium in changes.

Following Dekle, Eaton and Kortum (2008), let x' be the value of any variable in the new steady state and define the change in the value of variables between the old and the new equilibrium as $\hat{x} = x'/x$. Thus, I obtain an expression for any variable in the new equilibrium as $x' = \hat{x}x$. The following definition, considers the original equilibrium in terms of changes. This is similar to Caliendo and Parro (2015).

Definition 2: Equilibrium in terms of changes. *Let $(\mathbf{w}, \mathbf{P})$ be an equilibrium under trade costs $\{\tau_{\mathbf{jn}}\}_{\mathbf{j,n\in R}}$. Consider a different equilibrium $(\mathbf{w}', \mathbf{P}')$ under trade costs $\{\tau'_{\mathbf{jn}}\}_{\mathbf{j,n\in R}}$. Let $(\hat{w}, \hat{P})$ be an equilibrium under trade costs $\{\tau'_{\mathbf{jn}}\}_{\mathbf{j,n\in R}}$ relative to $\{\tau_{\mathbf{jn}}\}_{\mathbf{j,n\in R}}$, where variable $\hat{x}$ represents relative changes, that is $\hat{x} = \frac{x'}{x}$. Then, the equilibrium conditions (6), (9), (13), (14) and (15) can be expressed in relative changes:*

(i) *Good market clearing condition*

$$\hat{c}_{n,k} = (\hat{w}_n)^{\beta_n^{lk}} \prod_{s\in\{a,m,i,z\}} (\hat{P}_{ns})^{\beta_n^{sk}} \tag{17}$$

(ii) *Expenditure shares*

$$\hat{\lambda}_{jn,k} = (\hat{\tau}_{jn,k})^{1-\sigma_k} (\hat{c}_{j,k})^{1-\sigma_k} (\hat{P}_{n,k})^{\sigma_k-1} \tag{18}$$

(iii) *Prices*

$$\hat{P}_{nk} = \Big[\sum_j (\hat{\tau}_{jn}\hat{c}_{j,k})^{1-\sigma_k} \lambda_{jn,k}\Big]^{\frac{1}{1-\sigma_k}} \tag{19}$$

(iv) *Total expenditure*

$$X'_{n,s} = \sum_k \beta_n^{s,k} \sum_j X'_{j,k}\lambda'_{nj,k} + \alpha_{n,s} I'_n$$

$$X'_{n,s} = \sum_k \beta_n^{s,k} \sum_j X'_{j,k}\hat{\lambda}_{nj,k}\lambda_{nj,k} + \alpha_{n,s}[\hat{w_n} w_n L_n + D'_n] \tag{20}$$

(v) *Trade balance*

$$\sum_k \sum_j X'_{j,k}\lambda'_{nj,k} = \sum_k \sum_j X'_{n,k}\lambda'_{jn,k} - D'_n$$

$$\sum_k \sum_j X'_{j,k}\hat{\lambda}_{nj,k}\lambda_{nj,k} = \sum_k \sum_j X'_{n,k}\hat{\lambda}_{jn,k}\lambda_{jn,k} - D'_n \tag{21}$$

3.6 Department-port gravity equation

I generate an expression for international trade flows between department d and the rest of the world, RoW, that use a specific city-port ρ (or specific international shipping route r_t). For the case of those trade flows between the rest of the world and the departments, equation (12) becomes a different expression. This is necessary, given I need to include the role of the specialized traders on the international trade flows. To do this, I obtain the share of exports/imports that use route r_t and combine it with equation (3), which defines the relationship between international shipping costs and trade costs, to generate a department-port gravity equation.

Shares of international shipping routes. Using the properties of the Frechet distribution, it is possible to obtain an expression for the shares of trade flows that are shipped via a specific international shipping route r_t for $t \in \{m, x\}$. Define $G_{r_t}(c)$ as the probability that the *international shipping cost* of a good sent via route r_t is lower than c.

$$G_{r_t,k}(c) \equiv Pr\Big[\frac{\tau_{r_t}}{z_{r_t,k}(\iota)} \leq c\Big]$$

$$G_{r_t,k}(c) = 1 - exp[-A_{r_t}(\tau_{r_t})^{-\theta_k} c^{\theta_k}] \tag{22}$$

Let $G_{t,k}(c)$ be the probability that a good shipped via route r_t has an *observed cost* lower than c. This probability is expressed as

$$G_{t,k}(c) \equiv Pr\{\tau_s(\iota) \leq c\} = Pr\Big[\min_{r_t} \frac{\tau_{r_t,k}}{z_{r_t,k}(\iota)} \leq c\Big]$$

$$G_{t,k}(c) = 1 - exp[-c^{\theta}\Phi_{t,k}], \quad \text{where} \quad \Phi_k = \sum_{r_t} A_{r_t}\tau_{r_t}^{-\theta_k} \tag{23}$$

Finally, define π_{r_t} as the probability that a good is shipped via route r_t as

$$\pi_{r_t,k} = Pr\{\tau_{r_t,k}(\iota) \leq \underset{v_t \in R_{t,d} \setminus r_t}{min} \tau_{v_t,k}(\iota)\}$$

$$\pi_{r_t,k} = \frac{A_{r_t}\tau_{r_t}^{-\theta_k}}{\Phi_{t,k}} \tag{24}$$

Similar to Eaton and Kortum (2002), I can show that the distribution of international shipping costs is the same, no matter which route is used (see Appendix B). This implies that $\pi_{r_t,k}$ also represents the share of the value of exports/imports between a department d and RoW, sent via route r_t.

Trade flows between department and rest of the world via a city-port. I obtain an expression for the trade flows between departments and the rest of the world shipped via a specific route r_t. Consider as example, the export flows that use route $r_t = (d, p)$:

$$X_{dRoW,k,r_t} = X_{dRoW,k}\pi_{r_t,k}$$

$$X_{dRoW,k,r_t} = (\tau_{dRoW,k})^{1-\sigma_k}\left(\frac{Y_{RoW,k}}{\Pi_{RoW,k}^{1-\sigma_k}}\right)Q_{d,k}P_{d,k}^{\sigma-1}\pi_{r_t,k}$$

Inserting (3) and (24) into the expression for trade flows between any department d to RoW, that are sent via route $r_t = (d, p)$, I get:

$$X_{dRoW,k,dp} = \left[\Phi_k^{-\frac{1}{\theta}}\Gamma\left(\frac{1+\theta}{\theta}\right)\right]^{1-\sigma_k}\left(\frac{Y_{RoW,k}}{\Pi_{RoW,k}^{1-\sigma_k}}\right)Q_{d,k}P_{d,k}^{\sigma-1}\left[\frac{A_d A_\rho(\tau_d\tau_p\tau_{dp})^{-\theta_k}}{\Phi_{t,k}}\right] \tag{25}$$

To obtain the previous result, I assume that $A_{r_t} = A_{dp} = A_d A_\rho$. This implies that the scale parameter of the Frechet distribution, which governs the behavior of the productivities of the shipping routes, depends on a productivity transportation factor related to the department, and another productivity transportation factor related to the ports. The assumption is economically intuitive. To see this, notice that if any of these factors increases, then the international trade costs between departments and the rest of the world fall (see equation 3), and the probability that the route $r_t = (d, \rho)$ is used also increases (see equation 24).

For the international shipping costs, I use the expression $\tau_{r_t} = \tau_p\tau_{dp}\tau_p$. A similar expression can be obtained for imports using a particular international shipping route. Notice that the assumption regarding the productivity term for the international shipping routes implies symmetric trade costs.

There are two characteristics of the international shipping costs τ_{dp} that matter for the theoretical framework. First, they affect the share of trade flows X_{dRoW} and X_{RoWd} that are traded via port ρ through international shipping routes $r_x = (d, \rho)$ and $r_m = (\rho, d)$, respectively, through the term π_{r_t}. Second, the international shipping costs affect the trade costs between department d and the rest of the world, $\tau_{d\ RoW}$. Such effects are economically intuitive. Consider that τ_{dp} depends on the infrastructure that connect d and ρ. If an infrastructure project reduces the road distance between d and ρ, then port ρ will be used more often ($\uparrow \pi_{r_t}$), and the department d will better connected to the global markets ($\downarrow \tau_{dRoW}$).

3.7 Estimation of changes in trade costs due to new infrastructure projects

I can use the *equilibrium in changes* previously defined in section 3.5 only if I take as given a specific change in the vector of trade costs, $\hat{\tau}$. The objective of this paper is to evaluate how a new road infrastructure project change the national comparative advantage. Hence, I need to define how improvements in the Colombian road network lead to changes in trade costs. To facilitate the comprehension of this process, figure 6 illustrates how new infrastructure projects translate into changes in trade costs.

Estimation of the change in trade costs between departments and the rest of the world. Consider a large infrastructure project that changes the travel times across all international shipping routes from $\{T_{r_t}\}$ to $\{T'_{r_t}\}$. If the function between trade costs and travel times is known, $\tau = f(T)$, then it is possible to obtain both the old and the new international shipping costs along all routes, τ_{r_t} and τ'_{r_t}, respectively. I use the function $\tau_{r_t} = exp(\beta_{time} T_{r_t})$, which is a standard assumption in international trade and economic geography models. I discuss with detail how to obtain the value of the parameter β_{time} in section 4.

Using the exact algebra method of Dekle, Eaton and Kortum (2009) with the transportation model equations (3) and (24), I can obtain the change in shares of trade flows between d and RoW that use international shipping route r_t

$$\hat{\pi}_{r_t,k} = \frac{(\hat{\tau}_{r_t})^{-\theta_k}}{\sum\limits_{v_t \in R_t} \pi_{v_t,k} (\hat{\tau}_{v_t})^{-\theta_k}} \tag{26}$$

and the change in trade costs between department d and RoW is expressed as

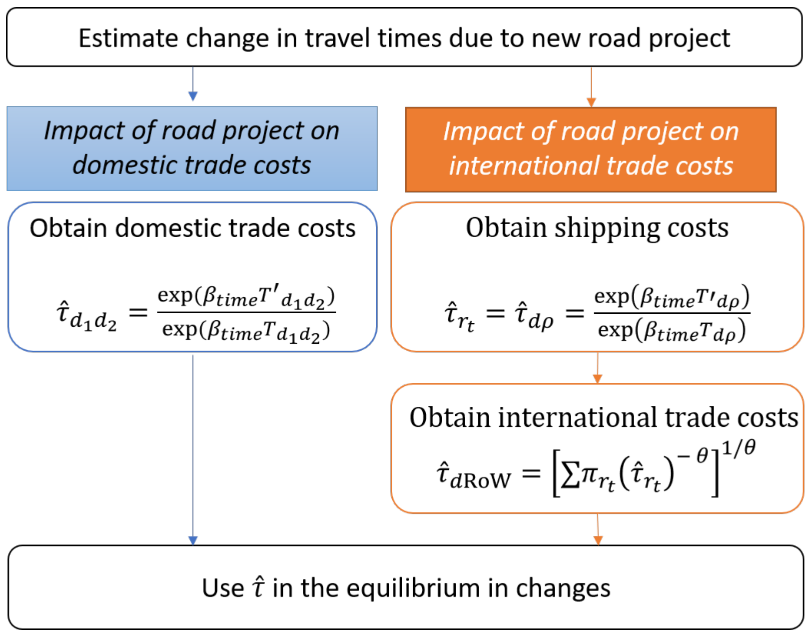
$$\hat{\tau}_{dRoW,k} = \left[\sum_{r_t \in R_t} \pi_{r_t,k} (\hat{\tau}_{r_t})^{-\theta_k} \right]^{-\frac{1}{\theta_k}} \tag{27}$$

where π_{dRoW,r_t} is the share of exports of department d to the rest of the world that use the route r_t. I can estimate this share using customs administrative data.[11].

[11] The shares of the export flows transported through a specific route might not necessarily be the same as the shares of imports shipped through this route (i.e. $\pi_{dRoW,r_t} \neq \pi_{RoWd,r_t}$). Hence, to make the counterfactual consistent with symmetric trade costs, I estimate the change in trade costs between any department and the rest of the world using the shares of total trade flows.

Estimation of the changes in trade costs between departments. I obtain the travel times before the infrastructure project is built, $\{T_{d_1 d_2}\}_{d_1,d_2 \in D}$, and after the highway is completed, $\{T'_{d_1 d_2}\}_{d_1,d_2 \in D}$. Then, I can get both the old and the new trade costs between departments ($\tau_{d_1 d_2}$ and $\tau'_{d_1 d_2}$, respectively) using directly the function $\tau_{d_1 d_2} = f(T_{d_1 d_2}) = exp(\beta_{time} T_{d_1 d_2})$. I do this because I assume there is no heterogeneity in the use of shipping routes between any two departments. Once I obtain the old and the new trade costs for the domestic trade model, I can calculate directly the change in trade costs for trade flows across departments, $\hat{\tau}_{d_1 d_2} = \frac{\tau'_{d_1 d_2}}{\tau_{d_1 d_2}}$.

Figure 6. Steps to obtain changes in trade costs



4 Taking the model to data

4.1 Parameters of the Armington model

Data sources to calibrate the model. I use the following datasets (i) customs data with records about individual export and import shipments, with information about the port of entry/exit, (ii) the World IO database (WIOD), (iii) the input-output table from the Colombian statistical agency for 2010, (iv) the 2013 Transportation Survey of Origin/Destination elaborated by the Colombian Ministry of Transportation, [12] (v) crude oil production data and refinery capacity, and (vi) the Economic Accounts produced by DANE to obtain variables such as value-added and gross output at a sectoral level . Appendix A provides more details.

Production and consumption parameters. I use the same value for the elasticity of substitution for all sectors, $\sigma_k = 6 \; \forall k$. I estimate the share of value added for the rest of the world and the

[12] This data was used to generate a proxy of the domestic trade flows of agriculture and manufacturing. Unfortunately, Colombia does not have a detailed Commodity Flow Survey like the United States that allows researchers to estimate good measures of domestic trade flows

departments using $\beta_n^{l,k} = (VA_k)/Y_k$, where VA_k is value added of sector k and Y_k is the gross production. Given the lack of input-output tables for Colombian departments, I assume the same value for this parameter for all departments. I estimated the share of sector s in the production of sector k using $\beta_n^{s,k} = (1 - \beta_n^{l,k})\frac{C_{intermediate,k,s}}{C_{intermediate,k,total}}$, where $C_{intermediate,k,s}$ is the intermediate consumption of sector k in goods from sector s, and $C_{intermediate,k,total}$ is the total intermediate consumption of sector k. I assume identical values of these parameters for all departments. Lastly, I estimate the share of final consumption in sector k with data from the input-output tables, using the formula $\alpha_{n,k} = C_{k,final,total}/C_{final,total}$, where $C_{k,final,total}$ is the final consumption in sector k and $C_{final,total}$ is the level of total final consumption.

Trade deficits and expenditure shares: agriculture and manufacturing. My estimation of trade deficits is limited by the information of transportation survey data that serves as proxy for domestic trade flows, for the agriculture and manufacturing sectors. The trade deficit of any department d can be considered as $D_{d,Total} = D_{d,Domestic\ trade} + D_{d,International\ trade}$. Unfortunately, I cannot obtain direct estimates of domestic trade flows using the cargo transportation survey. Hence, I assume that for Colombian departments, the deficit generated from the domestic trade is very small relative to the deficit the international trade. Hence, my deficit estimations exclusively consider the customs administrative data.

Expenditure shares: agriculture and manufacturing. For the case of the expenditure shares, $\lambda_{nj,k}$, table 1 illustrates the construction of the shares. I obtain the share of expenditures of Colombia on its goods, denoted by $\gamma_{Col,Col}$, using Colombia's input-output table, the share of expenditures of rest of the world on its goods, represented by $\gamma_{RoW\ RoW}$, using WIOD tables and the customs administrative dataset. Besides, using the customs data, I obtain the share of Colombian exports for every department, expressed as γ_{dRoW}, and the department share of national imports, characterized by γ_{RoWd}.

To obtain data on domestic trade flows, I rely on the transportation survey elaborated by the Ministry of Transportation for 2013. I assume this survey exclusively reflects patterns of domestic trade. I denote $\mu_{d_1d_2}$ as the shares of expenditures of a department d_2 in goods from d_1, exclusively considering domestic trade flows. Notice these are not the shares from the Armington model $\lambda_{d_1d_2}$ for $d_1, d_2 \in D$, because such shares consider both domestic and international trade. Unfortunately, I am not able to obtain values for the share of expenditures of departments on their goods for the case of domestic trade flows, μ_{dd}, therefore, I run my simulations under different values for this parameter ($\mu_{dd} = 0.3, 0.6$).

Table 1. Construction of the matrix of expenditure shares

Exporter ↓ Importer →	RoW	$\mathbf{d_1}$	...	...	$\mathbf{d_D}$
RoW	$\lambda_{RoWRoW} = \gamma_{RoW\ RoW}$	$\lambda_{RoWd_1} = (1-\gamma_{ColCol})\gamma_{RoWd_1}$			
$\mathbf{d_1}$	$\lambda_{d_1RoW} = (1-\gamma_{RoWRoW})\gamma_{d_1RoW}$	$\lambda_{d_1d_1} = \mu_{d_1d_1}\gamma_{ColCol}$			
...					
...					
$\mathbf{d_D}$	$\lambda_{d_DRoW} = (1-\gamma_{RoWRoW})$				$\lambda_{d_Dd_D} = \mu_{d_Dd_D}\gamma_{ColCol}$

Notes: μ_{ij} represents shares exclusively considering domestic trade flows between locations i and j, while γ_{mn} represents international trade flows between m and n

Trade deficits and expenditure shares: mining. I build the international and domestic expenditure shares of mining under the assumption that domestic trade flows are exclusively for crude oil between departments with oil fields and those with refineries,[13] while international trade flows include oil, coal, and minerals. I assume that those departments that are oil producers ship crude oil to the five refineries located in the departments of Bolivar, Santander, Casanare, Putumayo, and Meta.

Given that Colombia is a crude oil exporter, I presume that refineries only use crude oil produced domestically. To build these domestic trade flows, I infer that departments with refineries consume all the crude oil they produce, and if there is remaining capacity, they will import crude oil from other departments. The size of such domestic imports from each department is proportional to their oil production. This assumption allows me to obtain domestic trade flows for the mining sector. Additionally, I used the customs data to obtain international trade flows of the mining sector between departments and the rest of the world.[14]

4.2 Solving the model

I solve the model using the algorithm of Caliendo and Parro (2015). I make two adjustments: I do not need to consider how tariffs affect the expenditure function, and my measure of welfare does not need to consider tariff revenue.

4.3 Parameters of the transportation model

The department-port gravity equation (25) does not allow me to estimate the parameters that determine the dispersion of productivity of the shipping routes by sector, θ_k. To see this, consider the standard assumptions in international trade and economic geography models regarding the relationship between trade costs and travel time.

$$\tau_{d\rho} = exp(\beta_{time} T_{d\rho}) \tag{28}$$

where $T_{d\rho}$ is the travel time between department d and ρ and β_{time} is the parameter that defines the relationship between the shipping costs of a route $r_t = (d, \rho)$ and the travel time between department d and city-port ρ. By inserting this expression in the department-port gravity equation (25), and taking logs I obtain

$$\begin{aligned} ln(X_{dRoW,k,d\rho}) = \alpha + \alpha_{d,exporter} + \alpha_{d,importer} + \alpha_{RoW,exporter} + \alpha_{RoW,importer} \\ + \alpha_\rho - \mu_{t,k}(T_{d\rho}) + \epsilon_{d\rho} \end{aligned} \tag{29}$$

[13] 86 % of the gross domestic output of the mining sector is coal and crude oil. According to the Energy International Agency, Colombia exports most of its coal production. Hence, I assume that the domestic trade flows consisted mostly of crude oil from departments with oil fields to departments with oil refineries

[14] I could build a more precise measure of mining domestic trade flows using pipelines information. Unfortunately, I do not have accurate geospatial data about pipeline location and capacity.

where $\mu_{t,k} = \theta_k \beta_{time}$, and T_{dp} is the travel time between department d and city-port ρ. Using this structural regression, I get an estimate of $\mu_{t,k}$. Given that I cannot estimate separately the parameters β_t and θ_k, I use the value of β_{time} from previous literature. Specifically, I use the value reported by Allen and Arkolakis (2019). I elaborate about the value of this parameter in section 4.5.

4.4 Estimation of gravity equation

Although it is possible to use OLS to estimate $\mu_{t,k}$ using (29), there are concerns about the presence of endogeneity given the existence of unobservables correlated with both the travel time between a department d and city-port ρ and the international trade flows between such pair, $X_{dRoW,k,d\rho}$. Consider that $\epsilon_{d\rho}$ represents a bilateral cost/demand shifter of the international trade flows using the route $r_t = (d, \rho)$. The main source of endogeneity is the fact that the Colombian national government could target the pair department city-port, (d, ρ), through infrastructure policies that affect both the demand/cost shifter of international trade flows, $\epsilon_{d\rho}$, and the travel times $T_{d\rho}$.

To solve this endogeneity issue, I use an instrumental variable approach. This approach requires a valid instrument $Z_{d\rho}$. The instrumental variable needs to be relevant, $E[Z_{d\rho}T_{d\rho}] \neq 0$, and exogenous, $E[Z_{d\rho}\epsilon_{d\rho}] = 0$. I consider two instrumental variables: the distance between ports and capitals of departments using the road network of Colombia in 1938, and the distance between city-ports and the capitals of departments using the 17th-century colonial roads of the Viceroyalty of New Granada. These instrumental variables are similar to the ones used by Duranton (2015) to analyze the domestic trade between Colombian cities. I discuss the validity of the instrument below. Duranton, Morrow, and Turner (2014), Baum-Snow (2007), and Michaels (2007) also use a similar approach.

The road network of 1938 served specific regional purposes because railroads played a major role in the transportation of goods. Therefore, the transportation policies implemented by the Colombian national government focused on the expansion of the railroad network (Pachon and Ramirez, 2006; Alvear-Sanin, 2008). Also, as Duranton (2015) pointed out, the road infrastructure did not serve international trade purposes. For example, the two most populated Colombian cities (Medellin and Bogota) did not have a road connection to the Atlantic seaports (see Appendix B).

Duranton (2015) describes with detail the characteristics of the colonial road network (*caminos reales*). Some of the *caminos reales* were used by the indigenous tribes that lived in the country before the Spanish colonizers arrived. They mainly consisted of trails and paths used by the Spanish colonizers to travel to the interior of Colombia. To travel along these trails, it was necessary to use mules. Therefore, Duranton (2015) argues that internal trade was very small within colonial towns. Moreover, the first census implemented in Colombia at the beginning of the 19th century (two centuries after the colonial routes were established) indicates there were less than 2.4 million people in the country (DANE, 2019).[15] According to the 2018 census generated by DANE, Colombia had a population of 48.2 million persons. To sum up, the economic conditions that lead to the establishment of the colonial routes were very different, relative to the current economic circumstances

[15]The first census of Colombia was implemented in 1822, and included the nations of Venezuela, Panama, Colombia and Ecuador, which were part of the former Republic of Colombia.

that define which city-port a department uses to trade with the rest of the world.

The distance using an old road network is correlated with the travel times using the current road network, given that it is easier and less costly to build new roads using existing old paths or roads, relative to constructing new roads using new land. The exogeneity of my instrumental variables comes from the fact that given the economic conditions that explain the structure of the old road networks, it is highly likely that the current demand/cost shifters of the trade flows for a pair department city-port, (d, ρ), are uncorrelated with distance using old road networks, given that these network were built when the structure of the Colombian economy was different. In the 17th century, domestic trade in the country was relatively small. During 1938, Colombia was mainly an agricultural economy.

Given that for some department-port pairs, there is not a connection in the old road networks, I created two categorical indices based on the estimated road distances between locations using Dijkstra's algorithm, one for each road network. These indices include a category for those department-port pairs unconnected in the historical road networks. Table 2 reports the results of my estimation, combining both instrumental variables. There is no evidence of weak instrumental variables, given the value of the F-statistic of the first stage (Stock and Yogo, 2005). Moreover, the 2SLS estimates are more precise, compared to the OLS estimates.

Table 2. Empirical results of the gravity equation

Method	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS
Sector	agriculture		mining		manufacturing	
$-\mu_{t,k}(time_{dp})$	-0.3282	-0.5274	-0.2293	-0.5199	-0.2880	-0.6182
	(0.3410)	(0.0576)	(0.4100)	(0.0642)	(0.3190)	(0.0592)
F-statistic (1st stage)	-	13.94	-	13.94	-	13.94
N	1,026	1,026	1,026	1,026	1,026	1,026
R-squared	0.5430	0.5230	0.4600	0.4180	0.6910	0.6531

Notes: The categorical variables that I use as instrumental variables have a value of 1 if the department and the port are in the same city; a value of 2 if the distance between the locations is 1-300 kilometers for the 1938 road network, and 1-330 km for the colonial road network; a value of 3 if the distance between locations is 300-700 kilometers using the 1938 road system and 330-830 kilometers using the colonial path system; a value of 4 if the distance is larger than 700 km using 1938 roads, or the distance is longer than 830 kilometers using the 17th century roads; and a value of 5 for those locations unconnected using the old road network.

Given that there is a negative sign multiplying the parameter $\mu_{t,k}$ according to my structural model, then the value of this parameter is positive. Table 2 shows that the magnitude of the OLS estimate is smaller in absolute value, compared to the magnitude of the 2SLS estimate. This implies that any unobservable governments policies that are affecting both exports between department-port pairs and their travel times, are being targeted at regions with large travel times with respect to city-ports (or equivalently, poor infrastructure). This is consistent with the evidence provided by Pachon and Ramirez (2006) and Alvear-Sanin (2007) regarding infrastructure policies in Colombia.

4.5 Estimation of the parameter of the dispersion of productivity of shipping routes

To obtain estimates of the parameters $\theta_k \forall k \in \{a, m, i, z\}$, I use estimates of β_{time} from Allen and Arkolakis (2019). The authors consider the function $\tau_{nj} = exp(\beta_{time} T_{nj})$ in their estimation procedure, where T_{nj} is the travel time between locations n and j. They report $\beta_{time} = 0.08$ for a trade elasticity $\sigma = 9$. If I use the elasticity of substitution $\sigma = 6$, then $\beta_{time} = 0.13$. I report my estimates of the parameter θ_k in table 3.

A potential concern is that the estimate of β_{time} comes the context of the American network road system. The empirical evidence of Atkin and Donaldson (2015) shows that the relationship between intra-national trade costs and distance/travel times is very different in developing countries (Ethiopia and Nigeria) relative to the United States.

Although this may represent a concern, there is a caveat. First, data from the World Bank suggests that for the year 2012, Colombia's quality of infrastructure for trade and logistics was much higher compared to the African countries analyzed by Atkin and Donaldson (2015)[16]. This suggests that, even though the values for the parameter β_{time} may not be the same for the United States and Colombia, their differences must be much smaller than the reported by Atkin and Donaldson (2015) between the two African countries and the United States.

As a robustness check, I run my counterfactuals with other values of β_{time}. Specifically, I consider that the parameter can be 10% and 20% higher than the one from Allen and Arkolakis (2019) as it is shown in table 3. For the purpose clarity, Appendix E contains graphs on how the values of the parameters θ_a, θ_m and θ_i vary if I also consider the confidence intervals of my estimates of $\mu_{t,a}$, $\mu_{t,m}$ and $\mu_{t,i}$.

To interpret the magnitudes of the estimate of the parameter θ_k, it is necessary to recall that it is the shape parameter of the Frechet distribution. Economically, it represents the dispersion of the productivities of the international shipping routes (or equivalently, the dispersion of productivities of the city-ports). A high value for θ_k implies low heterogeneity in the productivity of the city-ports to export a good from sector k. A low value, represents high heterogeneity in the productivities of the city-ports. Given the values that I report in Table 3, this implies that for all three sectors, the city-ports show high levels of heterogeneity in their productivities. Intuitively, this implies that firms within a department tend to choose different city-ports to export and import goods from the rest of the world.

Table 3. Values for θ_k for different values of β_{time}

Parameter	$\theta_{agriculture}$	θ_{mining}	$\theta_{manufacturing}$
Values when $\beta_{time} = 0.13$ (Allen and Arkolakis, 2019)	4.06	4.00	4.76
Values when $\beta_{time} = 0.143$ (10% higher than baseline)	3.69	3.64	4.32
Values when $\beta_{time} = 0.156$ (20% higher than baseline)	3.38	3.33	3.96

Notes: To obtain the values of θ_k, I use the estimates of $\hat{\mu}_{t,k}$ shown in Table 2 for every sector $k \in \{a, m, i, z\}$ and the value of β_{time} from Allen and Arkolakis (2019). Then, I adjust the value of β_{time} upwards.

[16] In 2012, Colombia ranked 64th in the Logistics Performance Index of the World Bank (there are 168 positions). Ethiopia and Nigeria's positions were 141 and 118, respectively. The United States ranked 4th.

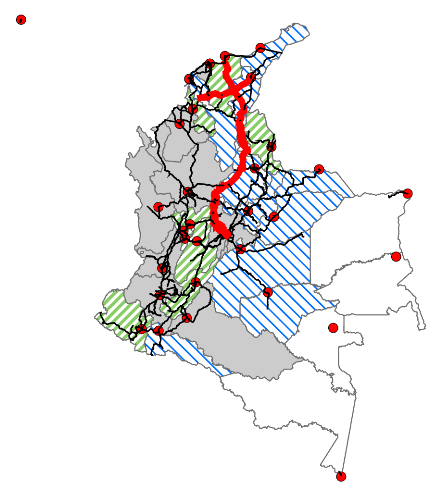
5 The impact of *Ruta del Sol* on comparative advantage

5.1 Expected impacts of the road project

I evaluate the effects of the construction of the infrastructure road project *Ruta del Sol*. The project consists of the construction, renovation, and expansion of lanes for 1,071 kilometers of the primary road system. The objective of the highway is to improve the connectivity between the center of the country and the Atlantic Ocean seaports. There was an unsuccessful attempt to start construction in 1997. A decade later, the Colombian government made a second attempt to start the project in 2009.

The project consists of three segments. The bidding process occurred in 2009, and contracts were negotiated and signed the following year (INCO, 2010a; INCO 2010b and INCO 2010c). The beginning of the construction for different segments started in the period 2010-2011. The project has faced multiple delays in its completion, although many sub-segments were inaugurated during the period 2014-2019 as the local media reported (El Espectador, 2019; La Republica, 2014, Semana 2019).

Figure 7. Location of the project "Ruta del Sol"



//// department specialized in agriculture \\\\ department specialized in mining ■ department specialized in manufacturing.

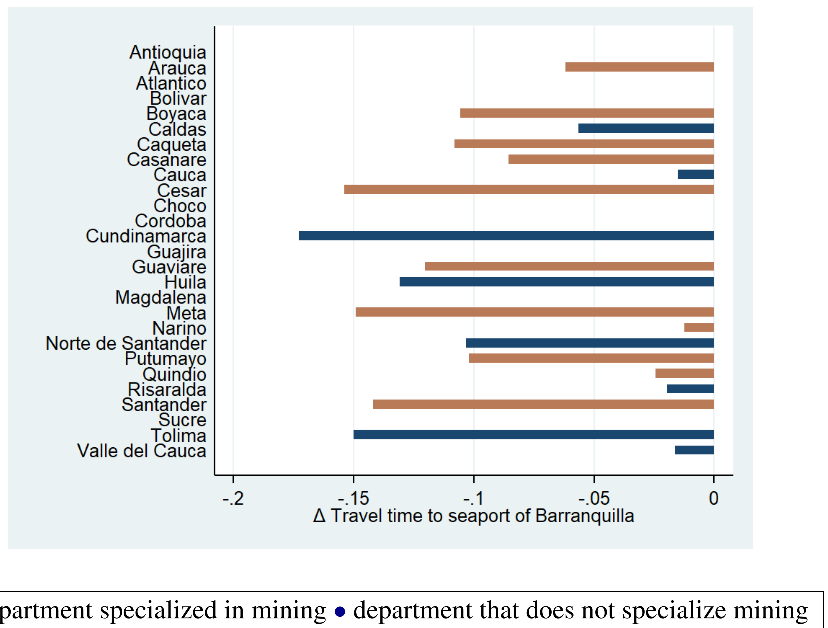
▲ city-port • capital of department — primary road network — Ruta del Sol (thick line)

Notes: The colors/figures that fill the area of every department show the sector with highest value of the Balassa index.

To measure the effects of the infrastructure project on travel times, I create a road network that includes improvements in the segments that already exist and those segments not built yet. I consider that after the completion of the project, the speed of the roads improves from 80 km/hour (approximately 50 miles/hour) to 100 km/hour (approximately 60 miles/hour). I chose a small change in speed derived from the completion of the project for the existing road segments. One of the main objectives of the project is to guarantee the existence of two lanes along the highway. This improvement particularly benefits trucks, by increasing physical maneuverability, particularly in the areas where the highways cross hilly regions. Such improvement has a direct impact on the speed of vehicles.

A priori, the effects of *Ruta del Sol* on the comparative advantage of Colombia are unclear. Figure 7 shows that the road crosses regions that specialize in different sectors. The project improves the connectivity between the department of Cundinamarca, which specializes in manufacturing, and the Atlantic seaports. But also reduces the travel times between departments that sepcialize in mining and the same seaports. Moreover, graphs in the Appendix D show that the international trade costs τ_{dRoW} fall for several departments and all tradable sectors, according to the predictions of my framework.

Figure 8. Reduction in travel distance between department and seaport of Barranquilla (%)



Notes: The change in travel times is measured as a fraction of the original travel times, without Ruta del Sol. I calculate the specialization of every department with the Balassa Index. I estimate the travel times between the capital of each department (shown in the vertical axis) and the seaport of Barranquilla for a baseline scenario and a new scenario. For the baseline, I assume that the existing segments that already exist (but will be improved) have a speed of 80 km/h. For the scenario in which the project is completed, these existing segments will have a speed of 100 km/h. In addition, I also include the planned new segments. Thus, the new scenario, in which the road project is completed, includes both the new and the improved road segments of *Ruta del Sol*.

5.2 Relevance of the parameter that defines relationship between trade costs and travel times

A potential concern regarding the evaluation of the effects of new infrastructure on sectoral exports is the choice of the value of β_{time}, which comes from the American road system. The value of the parameter affects the results given that it determines how changes in travel times in the Colombian primary road system lead to changes in domestic and international trade costs.

Although I do not have the true value of β_{time} for Colombia, there are reasons to believe the value of this parameter is higher in Colombia compared to the United States. This idea is supported by the empirical evidence of Atkin and Donaldson (2015), which suggests that travel times have a larger effect on trade costs in developing nations, relative to the United States' context. As a robustness check, I report the results of simulations using different values of β_{time}. These results are in Appendix E.

Appendix J shows the estimates of the effect of the completion of the highway *Ruta del Sol* in the trade costs $\tau_{dRoW,k}$ for different values of β_{time}. These graphs illustrate that using the value of β_{time} from Allen and Arkolakis (2019) leads to conservative estimates of the change in trade costs caused by the completion of *Ruta del Sol*.

5.3 Results of the main simulations

I report the effects of my simulations on the share of agricultural, mining, and manufacturing exports in table 4. As I mentioned before, for small open economies and aggregated sectors, the share of exports for a specific sector is a good proxy to measure shifts in the comparative advantage of a country in a specific sector (this share is the numerator of the Balassa index of RCA). As French (2017) documents, RCA indices are useful to measure patterns of comparative advantage.

As a robustness check, I implement my simulations under different values for the share of expenditures on own goods for the case of domestic trade of Colombian departments, μ_{dd}.[17]

Table 4. Results of the simulation under different parameters

	Counterfactual	$\bar{\mu}_{dd}$	$\frac{X_{agriculture,Col}}{X_{total,Col}}$	$\frac{X_{mining,Col}}{X_{total,Col}}$	$\frac{X_{manuf.,Col}}{X_{total,Col}}$	**Change in the share of manuf. exports (p.p.)**
	No new road	0.3	7.70 %	54.19%	38.10%	
I	Completion of Ruta del Sol	0.3	7.53 %	50.15 %	42.32 %	**+4.22**
	No new road	0.6	7.39%	55.76 %	36.85 %	
II	Completion of Ruta del Sol	0.6	7.43 %	51.37 %	41.21 %	**+4.35**

Note: μ_{dd} is the share of expenditures of a department in its own goods for the agricultural and manufacturing sector (only considering the domestic trade flows)

Under different values of the parameters that govern the trade and transportation framework, the results are similar: the infrastructure project *Ruta del Sol* leads to a higher share of manufacturing exports, even though it increases the connectivity of several mining departments.

[17] See section "Taking the model to the data" for more details

5.4 Mechanisms: the role of regions and their specialization

I analyze which departments contribute to the increase in the share of manufacturing exports in my simulations. The improvements in the connectivity of the department of Cundinamarca and the capital district of Bogota could be the main source of this growth. This is because the three main manufacturing regions in Colombia are located in the metropolitan areas of Bogota (capital district and capital of the department of Antioquia), Medellin (capital of the department of Antioquia), and Cali (capital of the department of Valle del Cauca). Nevertheless, among these three manufacturing cities, Bogota is the one that would observe the largest reductions in the travel times to the Atlantic seaports, as figure 8 shows.

To examine the regional contributions to the changes in the share of manufacturing exports, I analyze the change in manufacturing exports between two different equilibria using the following expressions

$$\Delta\ Manuf.\ Share\ Exports = \frac{X'_{i,Colombia}}{X'_{Colombia}} - \frac{X_{i,Colombia}}{X_{Colombia}}$$

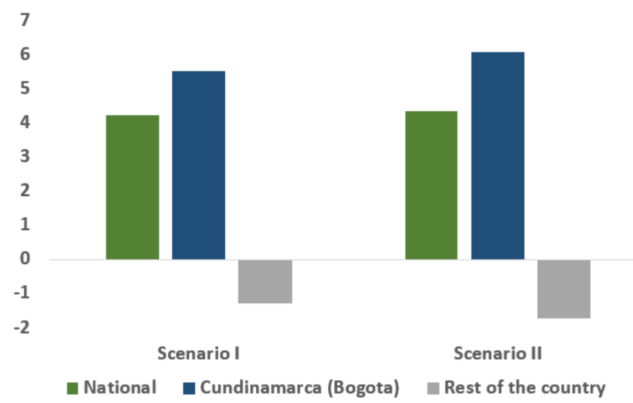
$$\Delta\ Manuf.\ Share\ Exports = \Big[\Big(\frac{X'_{i,b}}{X'_{Colombia}}\Big) + \Big(\frac{\sum_{d\neq b} X'_{d,i}}{X'_{Colombia}}\Big)\Big] - \Big[\Big(\frac{X_{i,b}}{X_{Colombia}}\Big) + \Big(\frac{\sum_{d\neq b} X_{i,d}}{X_{Colombia}}\Big)\Big]$$

$$\Delta\ Manuf.\ Share\ Exports = \Big[\frac{X'_{i,b}}{X'_{Colombia}} - \frac{X_{i,b}}{X_{Colombia}}\Big] + \Big[\frac{\sum_{d\neq b} X'_{i,d}}{X'_{Colombia}} - \frac{\sum_{d\neq b} X_{i,d}}{X_{Colombia}}\Big] \tag{30}$$

where $X_{Colombia}$ and $X'_{Colombia}$ are total exports of Colombia under the old and new equilibrium, respectively; $X_{i,d}$ and $X'_{i,d}$ are the manufacturing exports of the department d for the case of the old and new equilibrium, respectively.

Using (30), the change in share of manufacturing exports is the contribution of any department b (first term parenthesis) plus the contribution of the rest of the departments (second term in parenthesis). Hence, I decompose the growth in the share of manufacturing exports for different scenarios. I display the results of the decomposition in figure 9. These results show that the increase in manufacturing exports of Cundinamarca is the main driver of the change of national comparative advantage towards manufacturing.

Figure 9. Decomposition of growth in share of manufacturing exports (%)



Note: μ_{dd} is the share of expenditures of a department in its own goods for the agricultural and manufacturing sector (only considering the domestic trade flows)

5.5 Mechanisms: the role of industry linkages and the structure of the road system

To understand the forces driving the shift of the comparative advantage of Colombia towards manufacturing, I analyze the increase in the share of manufacturing exports under different counterfactual scenarios that consider separately the road network effects of *Ruta del Sol*, with and without input-output linkages.

In the first alternative counterfactual (scenario A), I close the input-output linkages but I allow for the impact of the road infrastructure project on both domestic and international trade costs (see equation 5). This implies that firms producing intermediate goods exclusively use labor as input. The second alternative counterfactual (scenario B) allows for the existence of input-output linkages, but only takes into account the effects of *Ruta del Sol* on the domestic trade costs. Lastly, I run a third alternative counterfactual simulation (scenario C), in which I consider industry linkages, but I assume the road project only affects international trade costs, and does not change domestic trade costs. I report the results of these alternative counterfactual experiments in columns 2 and 3 of table 5.

Table 5. Results of alternative simulations

Scenario		Increase in the share of manufacturing exports	
		($\bar{\mu}_{dd}$ = 0.3)	($\bar{\mu}_{dd}$ = 0.6)
Main	All the effects of *Ruta del Sol*	+4.2	+ 4.4
A	Impacts of *Ruta del Sol* without considering input-output linkages	+1.2	+1.8
B	*Ruta del Sol* only affects domestic trade costs	+0.7	+0.7
C	*Ruta del Sol* only affects international trade costs	+3.6	+2.0

Note: When I use the Balassa index for Colombia, the share of exports of a sector is a good proxy of its comparative advantage, since the denominator of the index is given for a small open economy and if the sectors are not very disaggregated. μ_{dd} is the share of expenditures of a department in its own goods for the agricultural and manufacturing sector (only considering the domestic trade flows).

The alternative counterfactual simulations provide two interesting insights about the forces driving my results. The first insight is that improvements in infrastructure lead to better access to intermediate inputs. This specially benefits manufacturing exports. To see this, the results of the scenario B are informative. In this alternative simulation, I consider that *Ruta del Sol* only improves access to domestic inputs. As a result, the national share of manufacturing exports increases by 0.7 percentage points. Hence, the improvement in the access of domestic inputs alone helps to increase the share manufacturing exports.

The second insight is that the existence of industry linkages propagate the positive effects generated by the road project. The presence of such linkages benefit the manufacturing sector the most. To see this, notice that in scenario A, in which input-output linkages are not considered, the reductions in trade costs lead to an increase in the specialization of Colombia in manufacturing goods, but this growth is one third of the increase from the main counterfactual (scenario A vs. main scenario).

The alternative counterfactuals show the relevance of industry linkages when we measure the impact of road projects, using international trade models. These linkages are not usually considered in existing studies regarding the general equilibrium effects of infrastructure improvements. Failure to consider these linkages will result in the estimation of smaller effects of lower trade costs on the trade flows of specific sectors.

6 Conclusion

The main conclusion of this paper is that domestic trade costs are determinants of comparative advantage. This idea is especially relevant for those countries with low quality of infrastructure. Quality of roads influence the spatial distribution of trade costs, thus influencing the availability of factor endowments and inputs for the production of goods and services across regions within an economy. Hence, to have a more comprehensive view of the comparative advantage of a country, it is necessary to consider the structure of its road system.

This idea also has policy implications. Infrastructure projects can be a tool for those policymakers whose objective is to shift the comparative advantage of a country in a particular direction. Given that reductions in trade costs lead to welfare gains through multiple channels, as recent economic literature predicts, the construction of roads seems to be a feasible policy alternative to change the national comparative advantage.

Specifically, in the context of Colombia, one of the most important infrastructure policy projects *Ruta del Sol*, has the potential to change the comparative advantage of the country, by weakening the comparative advantage of the country in mining goods, while strengthening the comparative advantage of the nation in the manufacturing sector. My results indicate that the share of manufacturing exports would grow 4 percentage points in the long run due to the completion of the project. The importance of this magnitude is supported by the fact that in the past three decades, the share of exports for two mining products (coal and oil) has observed a substantial increase, from 30% in 1992, to 58% in 2018. Moreover, the reduction of the concentration of Colombian exports in mining goods is aligned with the objectives of public officials.

The change in the comparative advantage of Colombia caused by *Ruta del Sol* is driven by two forces. First, the road project increases access to global markets of the department of Cundinamarca, which specializes in manufacturing goods. Second, the improvement in access to inputs benefits the manufacturing firms the most, given the structure of input-output linkages of the country.

Lastly, my results highlight the relevance of input-output linkages when considering how infrastructure shapes comparative advantage. I show that when industry linkages are not considered, the increase in the share of manufacturing exports is one third of the growth observed for simulations that consider industry linkages. This result is specially relevant for previous work regarding the economic effects of infrastructure projects, given that little attention has been paid to the relationship between infrastructure and input-output linkages.

7 Bibliography

1. Acemoglu, D. and F. Zilibotti (1997). ”Was Prometheus Unbounded by Chance? Risk, Diversification and Growth”. *Journal of Political Economy*, Vol. 105, No. 4, 709-751.

2. Alder, S. (2019). ”Chinese Roads in India: The Effect of Transport Infrastructure on Economic Development”, mimeo, University of North Carolina, Chapel Hill.

3. ANIF (2014). “Costos de Transporte, Multimodalismo y la Competitividad de Colombia”. Centro de Estudios Economicos de la Asociación Nacional de Instituciones Financieras. Reporte elaborado para la Cámara Colombiana de la Infraestructura. Bogota, D.C., Colombia.

4. Allcott, H. and D. Keniston (2018). ”Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America”. *Review of Economic Studies*. Vol. 85, No. 2, pp. 596-731

5. Allen, T. and D. Atkin (2016). ”Volatility and the Gains from Trade”. NBER Working Paper No. 22276.

6. Allen, T. and C. Arkolakis (2019). ”The Welfare Effects of Transportation Infrastructure Improvements”. NBER Working Paper No. 25487

7. Alvear-Sanin, J. (2008). ”Historia del Transporte y la Infraestructura en Colombia, 1492-2007”, Ministerio de Transporte, Bogota D.C., 1a Edicion: Diciembre de 2008.

8. Arkolakis, C., S.K. Lee and M. Peters (2018). ”European Immigrants and the United States’ Rise to the Technological Frontier in the 19th Century”. 2019 Meeting Papers 1420, Society for Economic Dynamics.

9. Atkin, D. and D. Donaldson (2015). ”Who’s Getting Globalized? The Size and Implications of Intra-National Trade Costs”, NBER Working Paper 21439.

10. Balassa, B. (1965). ”Trade liberalization and ’revealed’ comparative advantage”. *The Manchester School*. Vol. 33, Issue 2, pp. 99-123.

11. Bartelme, D., A. Costinot, D. Donaldson and A. Rodriguez-Clare (2019). ”The Textbook Case for Industrial Policy: Theory Meets Data”, mimeo University of Michigan.

12. CAF (1993). ”Proyectos Viales de Integracion Andina”. Corporacion Andina de Fomento, Santafe de Bogota, Enero 1993.

13. Caliendo, L. and F. Parro (2015). ”Estimates of the Trade and Welfare Effects of Nafta”. *The Review of Economic Studies*, Vol. 82, No. 1, pp. 1-44.

14. Caselli, F., M. Koren, M. Lisicky, and S. Tenreyro (2019). ”Diversification through Trade. The Quarterly Journal of Economics”, *The Quarterly Journal of Economics* forthcoming.

15. Chandra, V., J. Boccardo and I. Osorio (2007). ”Export Diversification and Competitiveness in Developing Countries”, mimeo, World Bank.

16. Coatsworth, J. (1979). ”Indispensable Railroads in a Backward Economy: The Case of Mexico”. *The Journal of Economic History*, Vol. 39, No. 4, pp. 939-960.

17. Cosar, K. and P. Fajgelbaum (2016). ”Internal Geography, International Trade and Regional Specialization”. *American Economic Journal: Microeconomics*, 8(1): 24-56.

18. DANE (2019). ”DANE en el Bicentenario: Historia de la división político-administrativa y los censos en Colombia”. Available online https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/censo-nacional-de-poblacion-y-vivenda-2018/cuantos-somos

19. Deardoff, A. (2014). ”Local comparative advantage: Trade costs and the pattern of trade”. International Journal of Economic Theory, Vol. 10, Issue 1, pp. 9-35.

20. Dekle, R., J. Eaton and S. Kortum (2008). Global Rebalancing with Gravity: Measuring the Burden of Adjustment. *IMF Staff Papers*, Vol. 55, No. 3, pp. 511-540.

21. DiGiovanni, J. and A. Levchenko (2012). NBER International Seminar on Macroeconomics, University of Chicago Press, vol. 8(1), pages 97-151.

22. DNP (2014). “Logística: Factor Clave para el Comercio Exterior”. Presentation by the Director of Infrastructure and Sustainable Energy Dimitri Zaninovich Victoria. Available online https://www.dnp.gov.co/DNPN/direcciones/dirección-de-infraestructura-y-energia-sostenible.

23. Donaldson, D. (2018). ”Railroads of the Raj: Estimating the Impact of Transportation Infrastructure”. *American Economic Review*, Vol. 108 (4-5), pp. 899-934.

24. Donaldson, D. and R. Hornbeck (2016). ”Railroads and American Economic Growth: A ’Market Access’ Approach”. *The Quarterly Journal of Economics*, Vol. 131, Issue 2, pp. 799-858.

25. Duranton, G., P. Morrow, and M.A. Turner (2014). ”Roads and Trade: Evidence from the US”. *The Review of Economic Strudies*. 81 (2), 681–724.

26. Eaton, J. and S. Kortum (2002). ”Technology, Geography and Trade”, *Econometrica*, Vol. 70, Issue 5, September 2002, pp. 1741-1779.

27. Duranton, G. (2015). ”Roads and Trade in Colombia”. *Economics of Transportation*, Vol. 4, pp. 16-36.

28. El Espectador (2014). ”Primer tramo de la Ruta del Sol, en veremos”. October 14th, 2014. Available online. https://www.elespectador.com/noticias/bogota/primer-tramo-de-ruta-del-sol-veremos-articulo-522631

29. El Tiempo (2007). ”Si de Procuraduria al caso Commsa”, Redaccion El Tiempo, August 15th, 2007. Available online in https://www.eltiempo.com/archivo/documento/MAM-2618824

30. Faber, B. and C. Gaubert (2019). ”Tourism and Economic Development: Evidence from Mexico’s Coastline”. *American Economic Review*, Vol. 109, No. 6, June 2019, pp. 2245-93.

31. FAO (2002). ”Dependence on Single Agricultural Commodity Exports in Developing Countries: Magnitude and Trends”, mimeo, discussion paper presented at the FAO Round Table on Special and Differential Treatment, WTO Negotiations on Agriculture, Geneva, February 1st 2002.

32. Fajgelbaum, P. and S. Redding (2018). ”Trade, Structural Transformation and Development: Evidence from Argentina 1869-1914”. NBER Working Paper 20217.

33. FEDESARROLLO (2015). Plan Maestro de Transporte Intermodal (PMTI) 2015-2035: Infraestructura Para el Comercio Exterior, el Desarrollo Regional y la Integración del Territorio. Informe Presentado a Financiera de Desarrollo Nacional.

34. French, S. (2017). ”Revealed comparative advantage: What is it good for?”. *Journal of International Economics*, Vol. 106, pp. 83-103.

35. Fogel, R. (1962). ”A Quantitative Approach to the Study of Railroads in American Economic Growth: A Report of Some Preliminary Findings”. *The Journal of Economic History*, Vol. 22, No. 2, pp. 163-197.

36. Hanson, G.H., N. Lind and M.A. Muendler (2015). ”The Dynamics of Comparative Advantage”, NBER Working Paper 21753.

37. Holl, A. (2016). ”Highways and Productivity in Manufacturing Firms”. *Journal of Urban Economics*, Vol. 93, pp. 131-151.

38. IADB (2014). ”Rethinking Productive Development: Sound Policies and Institutions for Economic Transformation”. Edited by Gustavo Crespi, Eduardo Fernández-Arias, and Ernesto Stein.

39. IGAC (2002). ”The Atlas of Colombia 2002”. Instituto Geográfico Agustín Codazzi. Santa Fé, Bógota, 5a edición revisada y aumentada.

40. INCO(2010a). ”Acta de Inicio del Contrato de Concesion No 002 del 14 de Enero de 2010, Proyecto Ruta del Sol - Sector 1: Tobiagrande/Villeta - El Korán”. Legal document of the Instituto Nacional de Concesiones, República de Colombia.

41. INCO(2010b). ”Acta de Inicio del Contrato de Concesion No 001 del 14 de Enero de 2010, Proyecto Ruta del Sol - Sector 2: Puerto Salgar - San Roque”. Legal document of the Instituto Nacional de Concesiones, República de Colombia.

42. INCO(2010c). ”Acta de Inicio del Contrato de Concesion No 001 del 4 de Agosto de 2010”. Legal document of the Instituto Nacional de Concesiones, República de Colombia.

43. Kremer, M. and S. Jayachandran (2006). ”Odious debt”, *The American Economic Review*, Vol. 96, No. 1, March 2006, pp. 82-92.

44. Koren, M. and S. Tenreyro, (2007). ”Volatility and Development”. *The Quarterly Journal of Economics*, Vol. 122, No. 1, 243-287.

45. Kruger, A. B. Tuncer (1982). ”An Empirical Test of the Infant Industry Argument”. *The*

American Economic Review, Vol. 72, No. 5, December 1982, pp. 1142-1152.

46. Krugman, P. (1987). ”The narrowing moving band, the Dutch disease, and the competitive consequences of Mrs. Thatcher: Notes on trade in the presence of dynamic scale economies”, *The Journal of Development Economies*, Vol. 27, Issues 1-2, October 1987, pp. 41-55.

47. La Republica (2014). ”Mintransporte inauguró obra en el Sector 2 de la Ruta del Sol”. May 3rd, 2014. Available online https://www.larepublica.co/infraestructura/mintransporte-inauguro-obra-en-el-sector-2-de-la-ruta-del-sol-2117241

48. Levchenko, A. (2005). ”Financial Liberalization and Consumption Volatility in Developing Countries”. *IMF Staff Papers*, 52:2, 237-259.

49. Levchenko, A. (2007). Institutional Quality and International Trade. *The Review of Economic Studies*, Vol. 74, No. 3, pp. 791-819.

50. Levchenko, A. and J. Zhang (2016). ”The Evolution of Comparative Advantage: Measurement and Welfare Implications”. *Journal of Monetary Economics*, Vol. 78, pp. 96-111.

51. Matsuyama, K. (1992). ”Agricultural Productivity, Comparative Advantage and Economic Growth”, *Journal of Economic Theory*, 58, pp. 317-334.

52. Michaels, G. (2008). The Effect of Trade on the Demand for Skill - Evidence from the Interstate Highway System. *The Review of Economics and Statistics*, Vol. 90, No. 4, pp. 683-701.

53. Ministerio de Transporte (2015). “Transporte en Cifras. Estadisticas 2015”. Oficina Asesora de Planeacion, Grupo de Planificacion Sectorial.

54. MIT Media Lab (2019). ”The Observatory of Economic Complexity”. Available online https://atlas.media.mit.edu

55. Morales, N. (2019). ”High-Skill, Multinational Companies and the Location of Economic Activity”, mimeo, University of Michigan.

56. Morando, B. (2019). The Reversal of Comparative Advantage? Transportation Costs and Cropping Patterns in Uganda, mimeo, Trinity College Dublin.

57. Oxford Economics (2017). Global Infrastructure Outlook: Infrastructure Investment Needs 50 Countries, 7 Sectors to 2040 Available online https://outlook.gihub.org/

58. Pachon, A. and M.T. Ramirez (2006). ”La infraestructura de transporte en Colombia durante el siglo XX”. Bogota: Banco de la Republica. Ediciones Fondo de Cultura Economica.

59. Pack, H. and Saggi K. (2006). ”Is there a case for industrial policy? A critical survey (English)”. *The World Bank research observer*, Vol. 21, No. 2, 267-297.

60. Pacione, M. (2005). ”City profile: Dubai”. *Cities*, Vol. 22, Issue 3, June 2005, pp. 255-265.

61. Pellegrina, H. and S. Sotelo (2019). ”Migration, Specialization and Trade: Evidence from the Brazilian March to the West”, mimeo, University of Michigan.

62. Semana (2006). ”Arreglo de Commsa: ¿no era mejor un buen pleito?”, September 12th, 2006. Available online https://www.semana.com/on-line/articulo/arreglo-commsa-no-mejor-buen-pleito/80948-3

63. Semana (2006). ”El calvario de Commsa”, November 18th, 2006. Available online

https://www.semana.com/nacion/articulo/el-calvario-commsa/82096-3

64. Semana (2019). ”La tragedia de la ruta del Sol”, Reportaje Especial. Available online https://especiales.semana.com/ruta-del-sol/

65. Sotelo, S. (2019a). ”Domestic Trade Frictions and Agriculture”. *Journal of Political Economy*, forthcoming

66. Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R. and de Vries, G. J. (2015), ”An Illustrated User Guide to the World Input–Output Database: the Case of Global Automotive Production”, *Review of International Economics*, Vol. 23, pp. 575–605

67. U.S. Energy Information Administration (2019). ”Country Analysis Executive Summary: Colombia”. Last updated: January 7th, 2019.

68. Van der Ploeg, F. (2011). ”Natural Resources: Curse or Blessing?”. *Journal of Economic Literature*, Vol. 49, pp 366-420.

69. Vieira-Posada, E. (2008). ”La Formacion de Espacios Regionales en la Integracion de America Latina”, 1a Edicion, Pontificia Universidad Javeriana: Convenio Andres Bello.

70. Xu, Hangtian (2016). ”Domestic railroad infrastructure and exports: Evidence from the Silk Route”. *China Economic Review*, Vol. 41, 129-147.

71. Xu, Mingzhi (2018). ”Riding on the New Silk Road: Quantifying the Welfare Gains from High-Speed Railways”, mimeo, University of California, Davis.

Appendix

A. Data

The following list contains detailed notes about data. This includes the geospatial dataset as well as the data regarding the calibration of all parameters. Unless otherwise indicated, I use data for 2013 in all cases.

Departments merged or dropped for the analysis

I merge or drop six departments when I take the model to the data

- San Andres y Providencia. The department is an island.
- Leticia. This department trades with the rest of the world exclusively because there is a regional dynamic between two border towns.
- Bogota (merged). The data from Bogota D.C. was merged with Cundinamarca. This gives us a total of 30 departments for the trade model.
- Vaupes, Vichada and Guainia. The states are not connected to the primary road system. Additionally, their international trade flows are small, and these flows are linked to the regional economic activity of the small border towns in Venezuela or Brazil.

Speed values for public-private roads

I assume higher speed values for public-private roads given that the characteristics of the public-private infrastructure projects suggest higher quality for these roads, relative to the standard ones. Such characteristics are publicly available via documents published by the National Agency of Infrastructure, the government office in charge of public-private infrastructure projects. Such documents include the legal contracts with information about design specifications and fines in case of violations by the construction company, as well as inspection documents.

There exists evidence that the Colombian government enforces these contracts, particularly for very expensive projects. Specifically, Alvear-Sanin (2008) documents a legal case in which the Colombian government sued an conglomerate of construction companies for breach of contract (the legal case of Commsa). The Colombian government attempted to impose the largest fine and persisted through different judiciary instances for nine years until a settlement was reached. Hence, it is safe to assume that the quality of public-private roads is higher compared to the standard roads that are directly administered by the Colombian government.

Trade flows

- **Oil exports**. The customs data does not record the department of origin for 55% of mining exports. This data corresponds to shipments with HS2012 codes 2709, 2710 and 2711 (petroleum and oil products). I use production data at a department level from the Colombian public oil company Ecopetrol to define the source of such flows. I assign the export flows without information about the department of origin proportionally to every department that produces oil, according to the production shares.

- **Trade between departments**.

  Agriculture and manufacturing. I use data of the estimated weight for the annual cargo flows from the Transportation Survey of Origin/Destination 2013 from the Ministry of Transportation in Colombia to create a matrix of domestic trade flows. I assume the domestic trade flows are the same for both sectors.

  Mining. I use data regarding oil production from the Ministry of Energy and Mines. I assume that only crude oil is domestically traded given that production of coal and oil represent 88% of the output of the mining sector according to the input-output matrix of Colombia created by DANE, for the year 2010. Additionally, coal is mostly exported by Colombia, according to data from the U.S. Energy Information Administration (2019). Therefore, I assume that most of the trade that occurs between departments will be crude oil from the oil fields to the states with refineries.

- **Purchases of location i to itself**.

  - Purchases of the RoW to itself, $\mu_{RoW,RoW}$. I estimated this value using data from WIOD 2013 to obtain $C_{world,final,k}$ and $C_{world,intermediate,k}$ and the customs data of Colombia to obtain this parameter.

  - Purchases of Colombia to itself, μ_{ColCol}. I estimated this using the input-output matrix produced by DANE for the year 2010.

  - Purchases of a department to itself or $\mu_{dd,k}$. I assume this number for the agricultural and manufacturing sectors. For the case of the mining sector, I obtained a proxy of this parameter for every department. To do so, I assume that all the domestic trade of mining is exclusively crude oil from the oil fields to the refineries, given that 88% of the mining production is coal and crude oil according to DANE, and that Colombia does use very little coal for energy consumption (less than 9%) according to the U.S. Energy International Agency (2019).

Trade deficits

- *Trade deficits beween departments and RoW*. I use customs administrative data from DANE for the year 2013.

- *Trade deficits between departments*.

Agriculture and manufacturing. Use data from the Transportation Survey of Origin/Destination 2013 produced by Ministry of Transportation in Colombia. I assume the trade deficits between departments are very small for agriculture and manufacturing, compared to the deficits of departments with the Rest of the world.

Mining. Similar to the way I obtained the trade flows shares, I calculate this variable assuming that domestic trade between departments is mostly crude oil from departments with oil fields to departments with refineries.

Input-output parameters

- Share of value added. Given that global input-output table of WIOD does not have data for Colombia, to estimate the parameter I consider the data for the entire world. This seems feasible given that Colombia is a very small economy, therefore it is likely that the value of this parameter for the world is the same with/without including the Colombian economy.
- Share of sector k in final demand $\beta_{i,k}$.
    - *Rest of the world*. Use final consumption column of the WIOT 2016. Due to constraints in WIOD data, I estimate the parameter for the entire world.
    - *Colombia*. I use input-output table produced by DANE for the year 2010.

Data sources

The following list provides the sources for every variable used in this paper.

1. **WIOD data**. It contains data for all European countries and other major economies. Colombian data is contained in the rest of the world, thus it is not reported individually. See Timmer et al. (2015). I use the input-output table corresponding to the year 2013 (version 2016).
2. **Colombian statistical agency DANE**
    (a) Input-output matrix
    (b) Value added data
    (c) Sectoral GDP data
3. **Colombian statistical agency (DANE)**. Provides the customs administrative data used to estimate trade flows between departments and rest of the world.
4. **Ministry of Transportation of Colombia (Ministerio de Transporte)**.
    (a) Physical maps regarding the primary road system. This allows me to obtain the road distance between Colombian departments. To create the map of 2013, I use as baseline

the digital road map created by the National Institute of Roads (INVIAS) for the year 2014.

(b) Data regarding the estimated weight of the cargo transported between the capitals of Colombian departments.

5. **International Monetary Fund**. Daily data for the exchange rate Colombian peso per dollar.

6. **Ministry of Mines and Energy**. Data on oil production for the year 2013 and the capacity of all refineries in Colombia.

Geospatial data

I obtain information regarding the location of city-ports and capitals of departments via two sources: the main topographic world map generated by ArcGIS software, and coordinates obtained through Google Maps. For some cases, the location of the city-port was assigned to specific coordinates to make sure that the trade costs from a location to itself was normalized to 1. I describe these cases below.

1. All the goods eported via the international bridge of San Miguel are assigned to Puerto Asis in the customs data. For the purpose of the estimation of distances, I use the actual location of the port of San Miguel.

2. I merged the Port of Coveñas and the Port of Cartagena given that they are located in the same city (Cartagena).

3. When the port is located within the city limits, then I situated the capital in the same location as in the port. The cases where this occurs are: Cartagena, Santa Marta, Pereira, Barranquilla and Bogota, .

   The cases where the port of trade is located outside the city limits are: Medellin, Arauca, Cali, Armenia and Bucaramanga.

4. I considered all the goods that are exported via the Port of Coveñas as exported via the Port of Cartagena given that they are located in the same city (Cartagena).

5. I did not use customs data from the ports of Inirida, Leticia, San Andres, Puerto Carreño. This is because the international trade flows of these towns are mainly influenced by the local border regions. For example, the trade flows observed in the port of Leticia, Colombia are mainly driven by regional dynamics between Leticia and Tabaratinga, Brazil.

B. Historical maps

Map of Colombia's road network in 1938 from the Atlas de Colombia (IGAC, 2002)

Map of the colonial routes of the Viceroyalty of New Granada available in the Atlas de Colombia(IGAC, 2002)

C. Derivations

Obtaining the expression for trade flows

By solving the firm's problem I obtain the demand of the composite good

$$q^c_{jn,k} = \frac{p^{-\sigma_k}_{jn,k}}{P^{1-\sigma_k}_{n,k}} Q_{n,k}$$

where $P_{n,k}$ is the price of the composite intermediate good and $p_{n,k}$ is the price of the intermediate good in location n.

Given the existence of perfectly competitive markets, the price charged by a firm located in j that sells good of sector k to composite goods firms in location n is

$$p_{jn,k} = \frac{\tau_{jn} c_{j,k}}{A_{j,k}}$$

Plugging this into the equation for the price of the composite intermediate, $P_{j,k}$, I obtain

$$P_{n,k} = \Big[\sum_j p^{1-\sigma_k}_{jn,k}\Big]^{\frac{1}{1-\sigma_k}} = \Big[\sum_j \Big(\frac{\tau_{jn} c_{j,k}}{A_{j,k}}\Big)^{1-\sigma_k}\Big]^{\frac{1}{1-\sigma_k}}$$

To obtain the expression for trade flows, combine the demand of composite good with the price, to get

$$x_{jn,k} = p_{jn,k} \cdot q^c_{jn,k} = \frac{p^{1-\sigma_k}_{jn,k}}{P^{-\sigma_k}_{nk}} Q_{n,k} \iff$$

$$X_{jn,k} = \Big(\frac{\tau_{jn,k} c_{j,k}}{A_{j,k}}\Big)^{1-\sigma_k} Q_{n,k} P^{\sigma_k - 1}_{n,k}$$

Obtaining the labor market clearing

By aggregating the total expenditure of location n in sector-k goods (14) across all sectors, I obtain the total expenditure of location n

$$X_n = \sum_s X_{n,s} = \sum_s \Big[\sum_k \Big(\beta^{s,k}_n \sum_j X_{j,k} \lambda_{nj,k}\Big) + \alpha_{n,s} I_n\Big]$$

$$M_n = X_n = \sum_s \sum_k \beta^{s,k}_n \sum_j X_{j,k} \lambda_{nj,k} + w_n L_n + D_n$$

$$E_n = \sum_k \sum_j X_{j,k}\lambda_{nj,k} = M_n - D_n = \sum_k (1-\beta_n^{l,k}) \sum_j X_{j,k}\lambda_{nj,k} + w_n L_n$$

where he first equality comes from the trade balance equation. After some algebra, I can obtain an expression for labor market clearing.

$$w_n L_n = \sum_k \beta_n^{l,k} \sum_j X_{nj,k} = \sum_k \beta_n^{l,k} \sum_j X_{j,k}\lambda_{nj,k}$$

Definition of equilibrium in levels (detailed).

The equilibrium is a set of wages $\{w_{n,k}\}_{n\in Z, k\in\{a,m,i\}}$, prices $\{P_{n,k}\}_{n\in Z, k\in\{a,m,i\}}$, and labor allocations $\{L_{n,k}\}_{n\in Z, k\in\{a,m,i\}}$ for all locations $n \in Z$ under the assumption of labor mobility across sectors and immobile labor across locations, given the following parameters:

(a) *trade costs $\{\tau_{ij}\}_{n,j\in R}$,*

(b) *share of value added of sector s in the production of sector k $\{\beta_n^{s,k}\}_{n\in R, s,k\in\{a,m,i,z\}}$,*

(c) *elasticity of substitution $\{\sigma_k\}_{k\in\{a,m,i,z\}}$,*

(d) *labor endowments $\{L_n\}_{n\in R}$,*

(e) *and total trade deficits $\{D_n\}_{n\in R}$*

that solve the following system of equations:

(i) *Wages.*

$$w_i = w_{i,k} \forall k$$

(ii) *Cost of an input bundle*

$$c_{n,k} = \phi_{n,k}(w_n)^{\beta_n^{l,k}} \prod_{s\in\{a,m,i,z\}} (P_{n,s})^{\beta_n^{s,k}}$$

(iii) *Prices.*

$$P_{n,k} = \Big[\sum_j \Big(\frac{\tau_{jn}c_{j,k}}{A_{j,k}}\Big)^{1-\sigma_k}\Big]^{\frac{1}{1-\sigma_k}}$$

(iv) *Trade flows shares.*

$$\lambda_{jn,k} = (\tau_{jn})^{1-\sigma_k}(c_{j,k})^{1-\sigma_k}(P_{n,k})^{\sigma_k-1}{A_{j,k}}^{\sigma_k-1}$$

(v) *Total expenditure*.

$$X_{n,s} = \sum_k \beta_n^{s,k} \sum_j X_{j,k}\lambda_{nj,k} + \alpha_{n,s} I_n$$

where

$$I_n = w_n L_n + D_n$$

(vi) *Trade balance*[18].

$$\sum_k \sum_{j\in R} X_{j,k}\lambda_{nj,k} = \sum_k \sum_{j\in R} X_{n,k}\lambda_{jn,k} - D_n$$

Transportation framework

Probability that the shipping cost offer is lower than c

Consider a shipping route $r_t \in R_t$ for $t \in \{x, m\}$. Denote the potential shipping cost of a trader ι as $\tau^o_{r_t,k}$. This offer depends on the shipping cost along route r_t and a productivity draw $z_{r_t,k}(\iota)$, which follows a Frechet distribution with parameters $(A_{r_t,k}\theta_k)$.

$$\tau^o_{r_t,k}(\iota) = \frac{\tau_{r_t,k}}{z_{r_t,k}(\iota)}$$

It can be noticed that the higher the value of the draw, the lower the shipping cost offer along route r_t. The probability that the shipping cost offer is lower than c is given by

$$G_{r_t,k}(c) = Pr\Big[\tau^o_{r_t,k}(\iota) \leq c\Big] \iff$$

$$G_{r_t,k}(c) = Pr\Big[\frac{\tau_{r_t,k}}{z_{r_t}(\iota)} \leq c\Big] \iff$$

$$G_{r_t,k}(c) = Pr\Big[z_{r_t}(\iota) \geq \frac{\tau_{r_t,k}}{c}\Big] \iff$$

$$G_{r_t,k}(c) = 1 - Pr\Big[z_{r_t}(\iota) \leq \frac{\tau_{r_t,k}}{c}\Big] \iff$$

$$G_{r_t,k}(c) = 1 - F\Big(\frac{\tau_{r_t,k}}{c}\Big) \iff$$

[18]This condition implies labor market clearing $w_n L_n = \sum_k \beta_n^{l,k} \sum_{j\in R} X_{j,k}\lambda_{nj,k}$

$$G_{r_t,k}(c) = 1 - exp[-A_{r_t}(\tau_{r_t})^{-\theta_k} c^{\theta_k}]$$

Let $\tau_s(\iota)$ be the actual shipping cost of trader ι from department d to the rest of the world. This cost is the minimum shipping price among all potential shipping cost offers across city-ports, that is

$$\tau_s(\iota) = \underset{r_t}{min}\ \tau^o_{r_t,k}(\iota) = \underset{r_t}{min}\ \frac{\tau_{r_t,k}}{z_{r_t,k}(\iota)}$$

Probability that the observed shipping cost is lower than c

Let $G_{t,k}(c)$ be the probability that the *observed* shipping cost $\tau_s(\iota)$ is lower than c. Therefore, I have

$$G_{t,k}(c) \equiv Pr\Big[\tau_s(\iota) \leq c\Big] = Pr\Big[\underset{r_t}{min} \frac{\tau_{r_t,k}}{z_{r_t,k}(\iota)} \leq c\Big] \iff$$

$$G_{t,k}(c) = 1 - Pr\Big[\underset{r_t}{min} \frac{\tau_{r_t,k}}{z_{r_t,k}(\iota)} \geq c\Big]] \iff$$

$$G_{t,k}(c) = 1 - Pr\Big[\bigcap_{r_t \in R_t} \frac{\tau_{r_t,k}}{z_{r_t,k}(\iota)} \geq c\Big] \iff$$

$$G_{t,k}(c) = 1 - Pr\Big[\bigcap_{r_t \in R_t} \frac{\tau_{r_t,k}}{z_{r_t,k}(\iota)} \geq c\Big] \iff$$

$$G_{t,k}(c) = 1 - \prod_{r_t \in R_t} Pr\Big[\frac{\tau_{r_t,k}}{z_{r_t,k}(\iota)} \geq c\Big] \iff$$

$$G_{t,k}(c) = 1 - \prod_{r_t \in R_t} [1 - G_{r_t}(c)]$$

Plugging the expression $G_{r_t}(c) = 1 - exp[-A_{r_t}(\tau_{r_t})^{-\theta_k} c^{\theta_k}]$ into the previous equation, I obtain

$$G_{t,k}(c) = 1 - \prod_{r_t \in R_t} exp\Big[-A_{r_t}(\tau_{r_t})^{-\theta_k} c^{\theta_k}\Big] \iff$$

$$G_{t,k}(c) = 1 - exp\Big[-c^{\theta_k} \sum_{r_t} A_{r_t}(\tau_{r_t})^{-\theta_k}\Big] \iff$$

$$G_{t,k}(c) = 1 - exp\Big[-c^{\theta_k} \Phi_t\Big]$$

where $\Phi_t \equiv \sum_{r_t} A_{r_t}(\tau_{r_t})^{-\theta_k}$.

Probability that any good is shipped via route r_t

Denote $\pi_{r_t,k}$ the probability that any good is shipped via route $r_t \in R_t$. Similar to Eaton and Kortum (2002), given that specialized traders have i.i.d. draws that are sector k specific in my framework, then $\pi_{r_t,k}$ is also the fraction of goods of sector k that are shipped via route r_t.

$$\pi_{r_t,k} \equiv Pr\Big[\tau^o_{r_t,k}(\iota) \leq \min_{v_t \in R_t \setminus r_t} \tau^o_{v_t,k}(\iota)\Big] \iff$$

$$\pi_{r_t,k} = \int_0^\infty Pr\Big[\min_{v_t \in R_t \setminus r_t} \tau^o_{v_t,k}(\iota) \geq c\Big] dG_{r_t,k}(c) \iff$$

$$\pi_{r_t,k} = \int_0^\infty Pr\Big[\bigcap_{v_t \in R_t \setminus r_t} \{\tau^o_{v_t,k}(\iota) \geq c\}\Big] dG_{r_t,k}(c) \iff$$

$$\pi_{r_t,k} = \int_0^\infty \prod_{v_t \in R_t \setminus r_t} [1 - G_{v_t}(c)] dG_{r_t,k}(c) \iff$$

$$\pi_{r_t,k} = \int_0^\infty \prod_{v_t \in R_t \setminus r_t} [1 - G_{v_t}(c)] dG_{r_t,k}(c) \iff$$

Using the expressions $G_{v_t,k}(c) = 1 - exp[-A_{v_t}(\tau_{v_t})^{-\theta_k} c^{\theta_k}]$, and $dG_{r_t,k}(c) = \frac{d}{dc}[1 - exp(-A_{r_t}(\tau_{r_t})^{-\theta_k} c^{\theta_k})]dc$, I obtain

$$\pi_{r_t,k} = \int_0^\infty \prod_{v_t \in R_t \setminus r_t} \Big[exp\Big(-A_{v_t}(\tau_{v_t})^{-\theta_k} c^{\theta_k}\Big)\Big] \Big[\frac{d}{dc}[1 - exp(-A_{r_t}(\tau_{v_t})^{-\theta_k} c^{\theta_k})]\Big] dc \iff$$

$$\pi_{r_t,k} = A_{r_t}(\tau_{r_t})^{-\theta_k} \int_0^\infty \theta_k c^{\theta_k - 1} [exp(-c^{\theta_k} \Phi_t)] dc$$

$$\pi_{r_t,k} = \frac{A_{r_t}(\tau_{r_t})^{-\theta_k}}{\Phi_t} \Big[- exp(-c^{\theta_k} \Phi_t|_0^\infty)\Big]$$

$$\pi_{r_t,k} = \frac{A_{r_t}(\tau_{r_t})^{-\theta_k}}{\Phi_t}$$

Why π_{r_t} is the fraction of trade flows between department d and the rest of the world that are shipped via route r_t

So far, I have shown that π_{r_t} is the fraction of exports/imports by department d to/from the rest of the world, RoW. But this is not the same as the percentage of the value of trade flows shipped via route r_t. Hence, I need to show that the distribution of shipping cost offers is independent of the shipping route. If this is true, then I can consider r_t as the fraction of exports/imports shipped via route r_t.

I express the probability that the shipping cost offer is lower than $\bar{c}$ conditional on route r_t offering the lowest price as

$$Pr[\tau^o_{r_t}(\iota) \leq \bar{c}|\tau^o_{r_t}(\iota) \leq \underset{v_t \in R_t \backslash r_t}{min} \tau^o_{v_t}(\iota)] = \frac{1}{\pi_{r_t}} \int_0^{\bar{c}} Pr[\underset{v_t \in R_t \backslash r_t}{min} \tau^o_{v_t}(\iota) \geq c] dG_{r_t}(c)$$

$$= \frac{1}{\pi_{r_t}} \int_0^{\bar{c}} \prod_{v_t \in R_t \backslash r_t} [1 - G_{v_t}(c)] dG_{r_t}(c)$$

Combining $G_{v_t,k}(c)$ and $dG_{r_t,k}(c)$ with my last expression, I get

$$Pr\Big[\tau^o_{r_t}(\iota) \leq \bar{c}|\tau^o_{r_t}(\iota) \leq \underset{v_t \in R_t \backslash r_t}{min} \tau^o_{v_t}(\iota)\Big] = \int_0^{\bar{c}} \prod_{v_t \in R_t \backslash r_t} [exp(-A_{v_t}\tau_{v_t})] \frac{d}{dc}[1 - exp(-A_{r_t}(\tau_{r_t})^{-\theta_k} c^{\theta_k})] dc$$

$$Pr\Big[\tau^o_{r_t}(\iota) \leq \bar{c}|\tau^o_{r_t}(\iota) \leq \underset{v_t \in R_t \backslash r_t}{min} \tau^o_{v_t}(\iota)\Big] = \frac{1}{\pi_{r_t}} \frac{A_{r_t}(\tau_{r_t})^{-\theta_k}}{\Phi_t} \Big[- exp(-c^{\theta_k} \Phi_t|^{\bar{c}}_0)\Big]$$

$$Pr\Big[\tau^o_{r_t}(\iota) \leq \bar{c}|\tau^o_{r_t}(\iota) \leq \underset{v_t \in R_t \backslash r_t}{min} \tau^o_{v_t}(\iota)\Big] = \frac{1}{\pi_{r_t}} \frac{A_{r_t}(\tau_{r_t})^{-\theta_k}}{\Phi_t} \Big[1 - exp(-\bar{c}^{\theta_k} \Phi_t)\Big]$$

$$Pr\Big[\tau^o_{r_t}(\iota) \leq \bar{c}|\tau^o_{r_t}(\iota) \leq \underset{v_t \in R_t \backslash r_t}{min} \tau^o_{v_t}(\iota)\Big] = G_{t,k}(\bar{c})$$

The distribution of shipping cost offers is the same for department d, independently of the route r_t used to transport the good. Therefore, the average value of the shipment sold/purchased by department d is independent of the route taken. This implies that we can express the fraction of the value of exports/imports that use shipping route r_t as π_{r_r}. This intuition is similar to the intuition of the result of Eaton and Kortum (2002), the best routes are more efficient, therefore such routes transport a larger share of goods to/from department d from/to the rest of the world, up to the level where the shipping cost offers are equal to the distribution of the observed shipping costs.

Trade costs between a department and the rest of the world

Using the results of the model with traders of Allen and Arkolakis (2019), define the trade cost between a department d and the rest of the world, as

$$\tau_{dRoW} \equiv E\Big[\tau_s(\iota)\Big]$$

$$\tau_{dRoW} = \int_0^\infty p_s(\iota) \iff$$

$$\tau_{dRoW} = \int_0^\infty p \; dG_t(p) \iff$$

$$\tau_{dRoW} = \int_0^\infty pdG_t(p) \iff$$

$$\tau_{dRoW} = \int_0^\infty p \, \frac{d}{dp}[1 - exp(-p^\theta \Phi_t)]dp \iff$$

$$\tau_{dRoW} = \int_0^\infty p\frac{d}{dp}[1 - exp(-p^\theta \Phi_t)]dp \iff$$

$$\tau_{dRoW} = \int_0^\infty \theta\Phi_t p p^{\theta-1} exp(-p^\theta \Phi_t)dp$$

Now, use change of variables, where $x = p^\theta \Phi_t$ and $dx = \theta p^{\theta-1}\Phi_t$. Therefore, I can express the integral as

$$\tau_{dRoW} = \int_0^\infty \Big(\frac{x}{\Phi_t}\Big)^{\frac{1}{\theta}} exp(-x)dx \iff$$

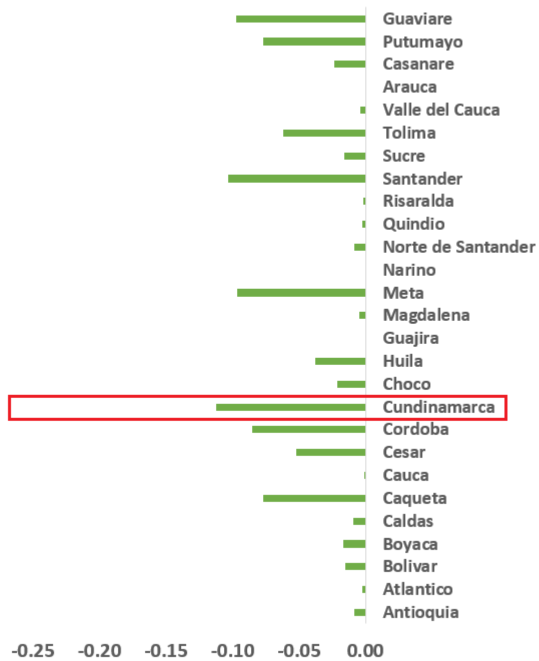
$$\tau_{dRoW} = \Phi_t^{-\frac{1}{\theta}} \int_0^\infty x^{\frac{1}{\theta}} e^{-x} dx$$

Recall that $\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}dx$. If I consider $(t-1) = \frac{1}{\theta} \iff t = \frac{1+\theta}{\theta}$, then I can express the trade cost between a department d and the rest of the world as

$$\tau_{dRoW} = \Phi_t^{\frac{1}{\theta}} \Gamma\Big(\frac{1+\theta}{\theta}\Big)$$

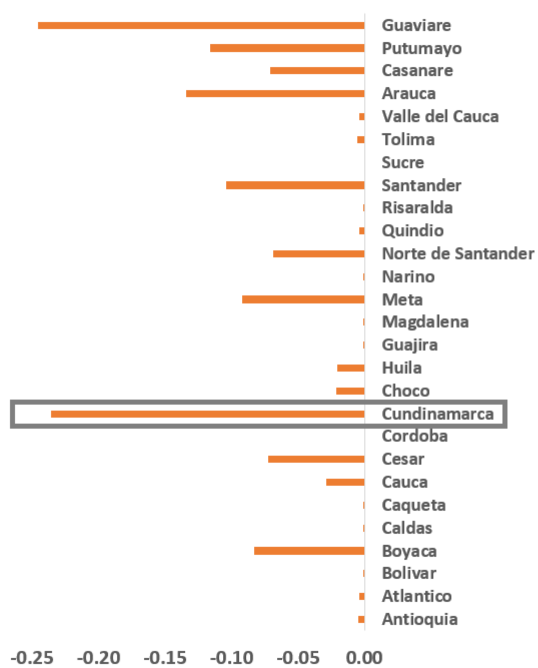
D. Reductions in trade costs $\tau_{dRoW,k}$ generated by the completion of Ruta del Sol

Figure D1. Reductions of $\tau_{dRoW,agriculture}$ caused by Ruta del Sol



Notes: I simulate the change in international trade costs using the value of β_{time} from Allen and Arkolakis (2019).

Figure D2. Reductions of $\tau_{dRoW,agriculture}$ caused by Ruta del Sol



Notes: I simulate the change in international trade costs using the value of β_{time} from Allen and Arkolakis (2019).

Figure D3. Reductions of $\tau_{dRoW,agriculture}$ caused by Ruta del Sol

Notes: I simulate the change in international trade costs using the value of β_{time} from Allen and Arkolakis (2019).

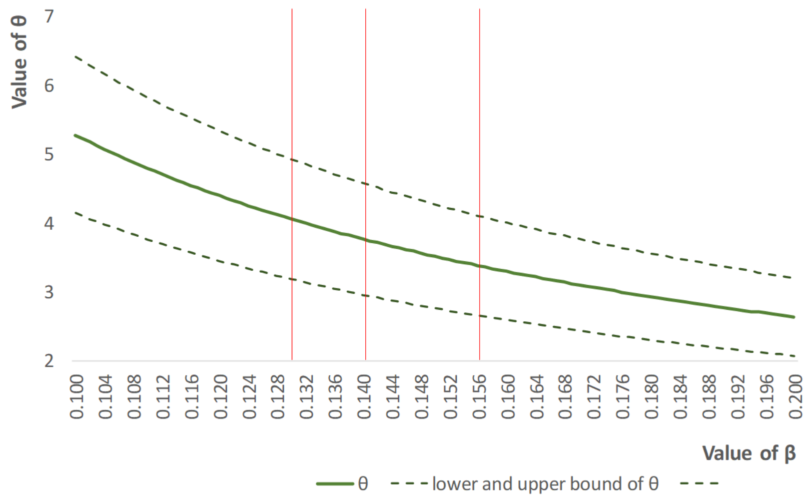
E. Robustness checks for the simulations of Ruta del Sol

Table E1. Effects of the ”Ruta del Sol” infrastructure project in sectoral exports of Colombia under different values of β_{time}

Scenario	$\bar{\mu}_{dd}$	Value of β_t that defines impact of project on τ_{dRoW}	$X_{agr.}/X_{total}$	X_{mining}/X_{total}	$X_{manuf.}/X_{total}$	Δ **share of manufacturing exports**
		No project (baseline scenario)	7.70 %	54.19%	38.10%	
A	0.3	0.13	7.53 %	50.15 %	42.32 %	**+4.22**
B		0.143	7.14 %	49.13 %	43.73 %	**+5.63**
C		0.156	6.9 %	47.3 %	45.8 %	**+7.72**
		No project (baseline scenario)	7.39%	55.76 %	36.85 %	
D	0.6	0.13	7.43 %	51.37 %	41.21 %	**+4.35**
E		0.143	7.09 %	50.41 %	42.51 %	**+5.65**
F		0.156	7.05 %	48.44 %	44.51 %	**+7.65**

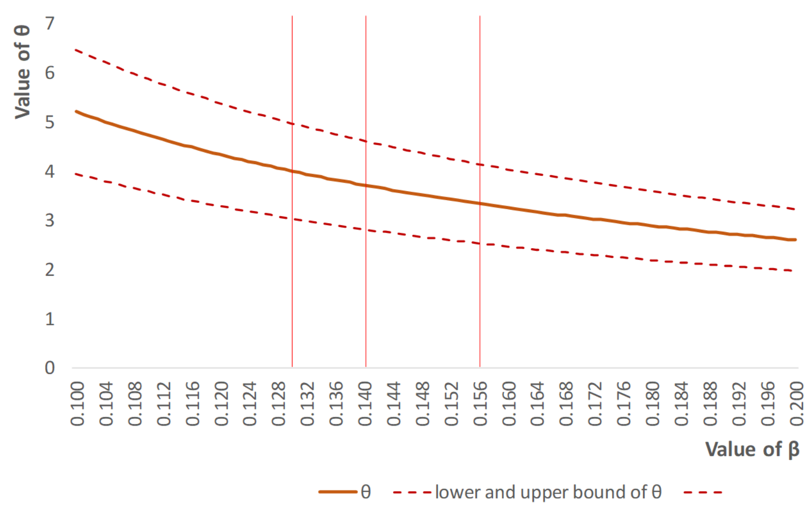
F. Values of θ_k considering the confidence intervals of $\hat{\mu_k}$

Figure F1. Value of parameter θ_a when considering the confidence interval of $\hat{\mu_a}$



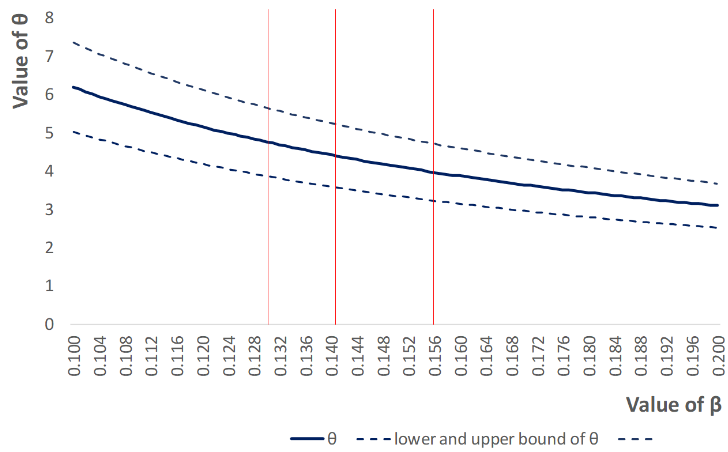
Notes: I use the estimate of $\hat{\mu_a}$ obtained using 2SLS. See Table 2.

Figure F2. Value of parameter θ_m when considering the confidence interval of $\hat{\mu_m}$



Notes: I use the estimate of $\hat{\mu_m}$ obtained using 2SLS. See Table 2.

Figure F3. Value of parameter θ_i when considering the confidence interval of $\hat{\mu}_i$



Notes: I use the estimate of $\hat{\mu}_i$ obtained using 2SLS. See Table 2.

G. Simulated change in trade costs after the completion of the highway "Ruta del Sol"

Figure G1. Simulated change in trade costs before/after Rutal del Sol is finished for agricultural sector

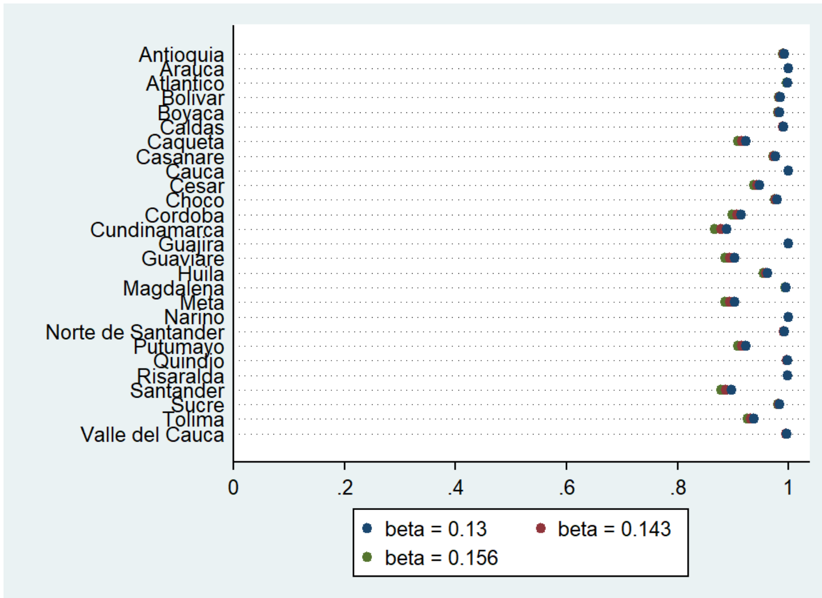


Figure G2. Simulated change in trade costs before/after Rutal del Sol is finished for mining sector

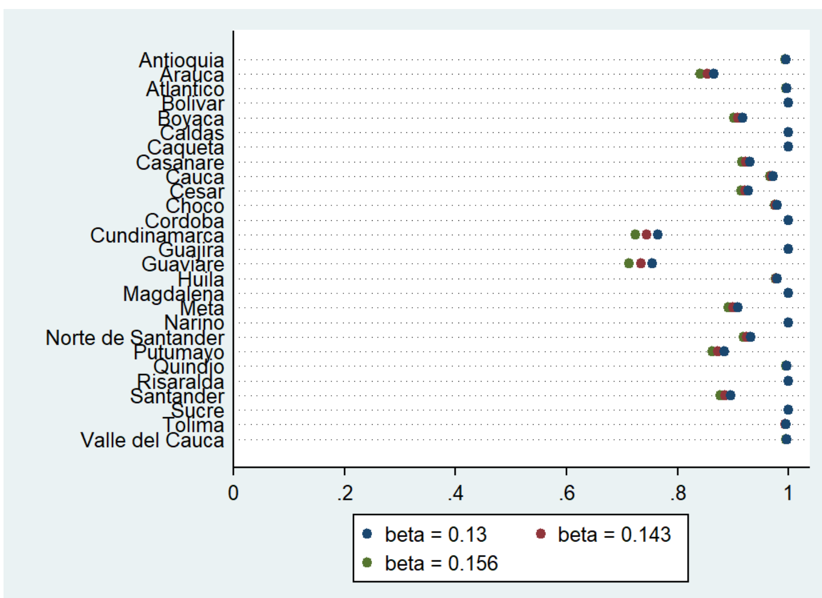


Figure G3. Simulated change in trade costs before/after Rutal del Sol is finished for manufacturing sector