ISSN 2279-9362

Collegio Carlo Alberto

Roadways, Input Sourcing, and Patterns of Specialization

Esteban Jaimovich

No. 586
April 2019

Carlo Alberto Notebooks

www.carloalberto.org/research/working-papers

Roadways, Input Sourcing, and Patterns of Specialization*

Esteban Jaimovich[†]

April 2019

Abstract

We propose a model where the internal transport network facilitates the sourcing of intermediate goods from geographically diffuse locations. A denser internal transport network promotes thus the growth of industries that rely on a large variety of inputs. The model shows that heterogeneities in internal transport infrastructures can become a key factor in shaping comparative advantage and specialization. Evidence based on industry-level trade data grants support to the main prediction of the model: countries with denser road networks export relatively more in industries that exhibit broader input bases. We show that this correlation is robust to several possible confounding effects proposed by the literature, such as the impact of institutions on specialization in complex goods. Furthermore, we show that a similar correlation arises as well when the density of the local transport network is measured by the density of their internal waterways, and also when road density is instrumented with measures of terrain roughness.

Keywords: International Trade, Comparative Advantage, Internal Transportation Costs.

JEL Classifications: F11, O18, R12

*I would like to thank Facundo Albornoz, Ainhoa Aparicio-Fenoll, Holger Breinlich, Giovanni Mastrobuoni, Vincenzo Merella and Ignacio Monzon, as well as seminar participants at the Collegio Carlo Alberto, University of Nottingham, University of Stirling, and University of Surrey for comments and suggestions.

[†]University of Surrey. Mailing address: School of Economics, Guildford, GU2 7XH, United Kingdom. Email: e.jaimovich@surrey.ac.uk

1 Introduction

The spatially diffuse distribution of economic activities means that transportation costs represent a major factor influencing countries' output, trade flows and specialization. Apart from few exceptions, the vast majority of the past trade literature has centered their attention on the cost of shipping goods internationally.[1] However, the evidence at hand suggests that internal transport costs are far from being a secondary component that can be disregarded when confronted with transboundary costs.[2] Furthermore, the impact of internal transport costs on specialization gets magnified by the fact that local infrastructures differ quite substantially between countries, especially when comparing economies at different stages of development.

Being able to efficiently transport commodities across space is crucial to keep total costs low. Yet, owing to specificities of their physical characteristics and of their production processes, some commodities turn out to be inherently more transport-intensive than others. This means that the efficiency of the local transportation infrastructure may unevenly affect the development of different industries. This paper studies a specific channel by which the internal transport network may shape countries' comparative advantages and specialization. One key role of the internal transportation network is that it facilitates the sourcing of intermediate inputs from geographically diffuse locations. As a result, industries that require a broader variety of intermediate inputs tend to make more intense use of the network.[3]

To illustrate this idea, we introduce a simple model with two intermediate sectors and a continuum of final good industries, which are geographically spread out within countries. A denser road network allows cheaper transportation of the intermediate inputs to the location site of final good producers. A crucial feature of the model is that final goods differ in terms of the breadth of their intermediate input requirements. In particular, some final industries have production functions that are very intensive in only one intermediate sector, while others require a more balanced mix of the two intermediate sectors. Since transportation of inputs is costly, those industries that require a relatively balanced combination of the intermediate inputs turn out to benefit relatively more (in terms of cost reduction) from a denser road network.

[1] For a few papers that have incorporated internal transport costs into international trade models, see Allen and Arkolakis (2014), Felbermayr and Tarasov (2015), Coşar and Fajgelbaum (2016), Ramondo, Rodriguez-Clare and Saborio-Rodriguez (2016), Redding (2016), Matsuyama (2017).

[2] See, e.g., Limao and Venables (2001), Anderson and Van Wincoop (2004), Hillberry and Hummels (2008), Mesquita Moreira *et al* (2013), Atkin and Donaldson (2015), Donaldson (2018), Agnosteva *et al* (2019).

[3] This idea was first suggested by Clague (1991a, 1991b) who argued that countries with poor infrastructure will specialize in 'self-contained' sectors (i.e., sectors that do not intensively rely on inputs from other sectors).

This simple mechanism yields a very clear prediction in terms of specialization within a framework with open economies. Countries that enjoy a denser local transport network tend to display a comparative advantage in the goods whose production process requires a broader mix of the intermediate inputs. This is because these are the industries that make heavier use of the local transport network to source their inputs. Conversely, countries with underdeveloped transport networks tend to specialize in industries with narrow input bases, as this allows them to economize on input sourcing costs.

After presenting the model we provide evidence consistent with its main prediction. To do so, we proceed as follows. Firstly, we index industries by their degree of input breadth using the information contained in the US input-output matrix. Secondly, we measure the density of local transport networks of countries by the length of their roadways per square kilometer. Finally, we correlate countries specialization by industries (measured by their total exports at the industry level) with an interaction term between industries' input breadth and countries' roadways density. We find that countries with denser road networks export relatively more in industries that exhibit broader input bases.

The correlation between road density and specialization in industries with broader input bases may obviously be driven by other mechanisms to the one suggested by our model. We show however that this correlation is robust to the inclusion of a large set of possible confounding covariates. In particular, one important channel related to ours works through institutions, as industries that rely on a wide set of inputs tend to be more dependent on contract enforcement [Levchenko (2007) and Nunn (2007)]. We show that the correlation predicted by our model is still present once we also control for the effect of rule of law. In that respect, our findings complement the previous studies that have interpreted the degree of input variety as a sign of product complexity, showing that industries with broad input bases seem also to be strongly reliant on the internal transport network.

One additional concern is whether the found correlation can be interpreted at all as evidence of *causation* from road density to specialization in transport-intensive industries. Roadways are the result of investment choices. Hence, road infrastructure may positively respond to transport needs resulting from patterns of specialization, reversing thus the direction of causation. Interestingly, we show that an analogous correlation to that one found with road density arises when using *waterways* density as an alternative measure of the depth of the local transport network. Moreover, this correlation is especially strong and significant in the case lower-income countries, which are exactly the types of economies that tend to suffer from sparser road networks.

Arguably, while waterways cannot be molded and expanded as flexibly as road networks, and hence they are less sensitive to issues of reverse causation, their evidence does not directly address this concern. In order to address more directly the possibility that reverse causality is behind our empirical results, drawing on Ramcharan (2009), we also instrument the density of a country's road network with topographical measures of terrain roughness.[4] The instrumental variable approach confirms the previous findings, granting further support to the hypothesis that the density of the internal road network is an important determinant of comparative advantage in industries with broader input bases.

There is a growing literature studying the impact of the local transport infrastructure on international and intra-regional trade and specialization. For example, Volpe Martincus and Blyde (2013) study the access to foreign markets and international trade across regions in Chile, Coşar and Demir (2016) does so for Turkey, and Volpe Martincus, Carballo and Cusolito (2017) for Peru. Donaldson (2018) looked at reductions of price and output distortions across Indian regions after expansions of the local railroad network, and Donaldson and Hornbeck (2016) assess how the expansion of the railroad network in the US enhanced market access of US counties. Fajgelbaum and Redding (2014) and Coşar and Fajgelbaum (2016) investigate the regional location of export-oriented activities given the local infrastructure in the cases of Argentina and China, respectively. Closer to our main focus, Duranton, Morrow and Turner (2014) and Coşar and Demir (2016) have tried to capture whether there is an effect of road infrastructure on specialization in transport-intensive activities. Duranton *et al* (2014) show that US cities with more highways tend to produce goods of higher weight per physical unit, while Coşar and Demir (2016) find a similar effect for Turkey. Our paper focuses on a different channel whereby the local transport infrastructure impacts comparative advantages: the notion that the spatial distribution of activities makes industries that need to source a large variety of intermediate inputs relatively more reliant on the internal transport network.

The internal transportation channel studied in this paper was first suggested by Clague (1991a, 1991b). There it is argued that poorer economies specialize in 'self-contained' sectors, as they lack a sufficiently developed infrastructure needed to sustain the production of industries that require a large variety of inputs. These articles, however, do not articulate this hypothesis within an international trade model, nor do they empirically assess whether trade flows at the industry level are associated with actual measures of the internal transport network in a way consistent with it.[5] We formulate the hypothesis that the local transport infrastructure

[4] Ramcharan (2009) shows that countries with rougher topography tend to exhibit less dense road networks. He argues that this is partly due to the impact of terrain roughness and grade variation on the cost of building and maintenance of transport networks.

[5] These articles provide evidence that the relative efficiency of underdeveloped economies is worse in industries

matters relatively more for industries with wider input bases within a trade model, where transport costs, location choices and comparative advantage are explicitly modeled. This leads to an endogenous determination of trade flows and specialization patterns, which respond to heterogeneities in transport infrastructures. In addition, we present evidence supporting the relevance of this mechanism exploiting cross-country variation in the density of road networks.[6]

Finally, our paper also relates to several strands of literature that have expanded upon the traditional Ricardian/Heckscher-Ohlin trade models based on heterogeneities in factor productivities/endowments. One set of papers have looked at enforcement institutions as a source of comparative advantage in industries producing complex goods requiring large variety of input-specific relationships [Antràs (2005), Acemoglu, Antràs and Helpman (2007), Levchenko (2007), Nunn (2007), Costinot (2009), Ferguson and Formai (2013)]. Another strand of literature has delved into the role of financial markets fostering exports in industries that are heavy users of external finance [Beck (2002), Svaleryd and Vlachos (2005), Becker, Chen and Greenberg (2012), Manova (2013)]. Finally, another institutional source of comparative advantage is presented by Cuñat and Melitz (2012), who show that countries with more flexible labor market regulations tend to export more in industries subject to higher volatility.[7] Our paper highlights the impact of local infrastructures when industries differ in their dependence on internal transportation of inputs.

The rest of the paper is organized as follows. Section 2 briefly discusses the empirical evidence on geographic clustering and coagglomeration of industries, which motivates some of the key assumptions in the model. Section 3 introduces the main features of the model in the case of a closed economy. Section 4 extends the model to a two-country setup, and derives the main predictions in terms of trade flows. Section 5 contrasts the main predictions of the model with the data. Section 6 discusses some endogeneity issues and alternative interpretations of the empirical results. Section 7 concludes. All relevant proofs are relegated to Appendix A.

that rely on a large variety of intermediate inputs. While this could be the result of poorer economies having less developed transport networks, it could also be the result of other factors usually associated with underdeveloped economies, like weaker institutions, lower levels of human capital, etc.

[6]Yeaple and Golub (2007) show that the stock of roads affects total factor productivity and sectoral composition across 10 industries for a panel of 18 countries. While their analysis highlights that roads may be a source of comparative advantage in some industries, it does not link the effect of heterogeneities in transport infrastructures to specialization in industries with different degrees of input diversity.

[7]See Chor (2010) for a paper that aims at quantifying the importance of all these institutional sources of comparative advantage, alongside the more traditional ones stemming from heterogeneities in factor productivities and endowments as in Romalis (2004).

2 Geographic Clustering and Coagglomeration of Industries: Evidence in the Literature

The main argument raised by this paper is the fact that countries with denser internal road networks tend to exhibit a comparative advantage in industries with broader input bases. The reason behind this result is that industries that require a wider variety of intermediate inputs end up being heavier users of the internal transportation network. The strength of this mechanism rests crucially however on some specific features of the geographic location structure of activities in the economy: different industries must be *unevenly* distributed across geographically diffuse locations. In fact, if the different industries in the economy were uniformly distributed across space, then whether a particular sector requires a wide or a narrow set of intermediate inputs would in principle have no differential impact on the incurred transport costs of inputs. In other words, if all geographic locations in an economy were able to offer local producers the same variety of intermediate inputs and at the same cost, then the density of the internal transport network would no longer unequally impact the cost of production of industries with various degrees of input breadth.

The differential effect of road density across sectors with broader versus narrower input bases is therefore intrinsically linked to two features of the geographic distribution of economic activities: *i*) concentration of specific industries in certain geographic locations; *ii*) geographic coagglomeration of industries with strong input-output links. Both of these features have been vastly documented in a number of empirical studies.

Geographic Concentration by Industry

Examples of geographic concentration of specific industries abound. Probably, the most often-cited ones in the economic geography literature are the high-tech firms in Silicon Valley and the automobile industry in Detroit. In addition to these paradigmatic cases, a growing number of articles have shown that industry geographic clustering is a quite prevalent feature among a vast number of different industries, and also in different countries.

Ellison and Glaeser (1997) study the degree of concentration of U.S. manufacturing industries, and find that 446 out of 495 four-digit SIC sectors display excess geographic concentration (relative to the degree of geographic concentration that would be observed if firms in all industries would pick their locations following an identical random process). Moreover, they find that over one quarter of the U.S. manufacturing industries exhibit what they deem as 'high geographic concentration'. Similar results are found for France by Maurel and Sedillot (1999), for the U.K. by Devereux, Griffith and Simpson (2004) and by Duranton and Overman (2005), for Japan by Mori, Nishikimi and Smith (2005), and for Belgium by Bertinelli and Decrop (2005). Following this literature, the model to be presented in the next two sections will aim at generating an uneven concentration of different industries in different regions of the economy.

Geographic Coagglomeration of Strongly Linked Industries

The other crucial geographic feature instrumental to an unequal impact of the internal transport network density on the cost of input transportation in different industries is the possibility for strongly linked sectors to locate near one another. Intuitively, the coagglomeration of downstream firms near their main input source helps reducing the transportation cost of inputs proportionally more for those firms that rely heavily on very few intermediate inputs. The economic geography literature has also documented quite categorically features consistent with this behavior. For example, Ellison and Glaeser (1997) and Ellison, Glaeser and Kerr (2010) show that manufacturing industries in the U.S. with strong input-output links tend to locate in close physical proximity. Duranton and Overman (2008) produce similar evidence for the case of the U.K. Quite remarkably, Ellison, Glaeser and Kerr (2010) show that the presence of input-output links between upstream and downstream industries represents quantitatively the strongest factor leading to coagglomeration of industries, amongst the three traditional Marshallian reasons for coagglomeration[8]. Furthermore, that article highlights that the key channel stimulating such patterns of coagglomeration is that of economizing on transport cost of main inputs via physical proximity. The model developed in the next two sections will aim at embedding this feature, so as to generate analogous patterns of geographic coagglomeration in equilibrium.

3 A Closed Economy with Internal Transportation Cost

This section presents the environment and main features of the model in the specific case of a closed economy. This proves helpful for two reasons. First, it allows an easier description of the main building blocks of the model. Second, it facilitates the exposition of the main intuition for how the density of the transport network may heterogeneously affect the cost of production in sectors with different degrees of input breadth.

[8] These three reasons are: *i*) to save on transport cost of goods; *ii*) to allow pooling of workers and save thus on transport of people; *iii*) to allow spillovers of ideas across firms/industries.

3.1 Setup and Environment

3.1.1 Geography and Population

The economy is inhabited by a continuum individuals with mass L. Each individual is endowed with one unit of labor time that is supplied inelastically to firms. The economy comprises two geographically separate regions denoted by $R = A, B$. Region R is inhabited by a mass of individuals L_R, where $L_A + L_B = L$. There are no restrictions to individuals mobility within the economy. For brevity, throughout the paper, we abstract from explicitly modelling the determination of L_A and L_B. However, those two variables could be thought of as resulting from the underlying equilibrium in the regional labor markets together with individuals' optimal location choices.[9]

Both region A and B are, in principle, able to host the production of any existing intermediate and final good. Production activities are carried out by firms, which must choose a specific location. All markets are assumed to be perfectly competitive.

3.1.2 Intermediate Sectors

There exist two intermediate sectors (or industries), indexed by $S = 0, 1$. Each sector S comprises a unit continuum of intermediate goods (or varieties), which we index by $i_0 \in [0, 1]$ and $i_1 \in [0, 1]$, for $S = 0$ and $S = 1$, respectively. Whenever creating no confusion, we will skip the subscript $S = 0, 1$ from $i_{S=0,1}$, and index intermediate goods in either sector by a generic $i \in [0, 1]$. Henceforth, we will refer to each i in sector S interchangeably as intermediate good i or variety i of sector S.

The technology in the intermediate sectors draws heavily from Eaton and Kortum (2002).[10] Producing one unit of variety i in sector S in region R requires $1/Z_{R,S}(i)$ units of labor, where each $Z_{R,S}(i)$ is the result of an independent draw from a Fréchet distribution with location parameter $T_{R,S} > 0$ and shape parameter $\theta > 1$. That is,

$$F_{R,S}(Z) = \exp\left(-T_{R,S}\, Z^{-\theta}\right). \tag{1}$$

For the rest of the paper, with regards to the parameters $T_{R,S} \in \{A, B\} \times \{0, 1\}$ we assume:

Assumption 1 *i)* $T_{A,0} = T_{B,1} = 1 + T$, *where* $T > 0$*; ii) and* $T_{A,1} = T_{B,0} = 1$.

[9] In Appendix D.1 we explicitly show how L_A and L_B may arise as equilibrium objects in a context where individuals optimally choose the specific region where to live, consume, and work.

[10] Other models with an internal geography within an Eaton-Kortum framework can be found in Ramondo *et al* (2016), Cosar and Fajgelbaum (2016), Redding (2017), Donaldson (2018), and Caliendo *et al* (2018).

Assumption 1 entails that region A will display (on average) higher productivity in intermediate varieties belonging to sector 0, and region B will exhibit (on average) higher productivity in intermediate varieties of sector 1.[11] The main intention behind Assumption 1 is to open up room for some degree of regional specialization across the two intermediate sectors.[12]

3.1.3 Final Goods Sector

In addition to the intermediate good sectors (and their respective varieties), there also exists a unit continuum of final goods, indexed by $j \in [0,1]$. Final goods are purchased by individuals with preferences given by

$$U = \int_0^1 \ln(y_j)\, dj, \tag{2}$$

where y_j denotes the consumed amount of j.

Final goods are produced by combining intermediate good varieties within two-level structured production functions, where sectoral CES aggregators of intermediate varieties are nested into Cobb-Douglas production functions of the sectoral aggregators. In particular, total output of final good $j \in [0,1]$ is given by:

$$Y_j = \Psi_j \left[\left(\int_0^1 x_{0,i}^{(\sigma-1)/\sigma} di \right)^{\frac{\sigma}{\sigma-1}} \right]^{1-\alpha_j} \left[\left(\int_0^1 x_{1,i}^{(\sigma-1)/\sigma} di \right)^{\frac{\sigma}{\sigma-1}} \right]^{\alpha_j}, \tag{3}$$
$$\text{where: } \alpha_j \in [0,1], \;\; \Psi_j \equiv \alpha_j^{-\alpha_j} (1-\alpha_j)^{-(1-\alpha_j)}, \text{ and } \sigma > 1.$$

In (3), $x_{0,i}$ and $x_{1,i}$ denote, respectively, the amount of intermediate variety i from sector 0 and sector 1 used in the production of final good j. The parameter σ denotes the elasticity of substitution between varieties of intermediate goods within each intermediate sector.

A crucial feature of the production functions in (3) is the fact that the Cobb-Douglas weights $1-\alpha_j$ and α_j are assumed to be sector-specific. This implies that final good sectors may differ in terms of the intensity requirements of inputs from each of the two intermediate sectors. Final

[11] The implicit symmetry across regions' productivity in sectors 0 and 1 implied by Assumption 1 is posed to keep the model's structure as simple as possible. Our main qualitative results will carry over under more general assumptions, as long as each region retains *some* degree of aggregate advantage in one of the sectors. In particular, we could replace Assumption 1 by a more general setting, satisfying: *i)* $T_{A,0} > T_{A,1}$; *ii)* $T_{B,0} < T_{B,1}$; *iii)* $T_{A,0} > T_{B,0}$; *iv)* $T_{A,1} < T_{B,1}$; and the main results of the model will remain essentially unchanged.

[12] Notice, however, that since each $Z_{R,S}(i)$ is independently drawn from a Fréchet distribution, both regions will end up producing in equilibrium a *positive* mass of intermediate varieties in *both* sector 0 and 1 (in other words, it is only the *degree* of specialization across intermediate varieties from sector 0 and 1 that will differ in equilibrium between region A and B).

sectors with a small α_j (resp. a large α_j) will tend to use intermediate varieties from sector 0 (resp. sector 1) more intensively. On the other hand, sectors whose α_j lies close to 0.5 will tend to use a relatively balanced mix of varieties from both intermediate sectors.[13]

For the remainder of the paper, we will assume that, when considering the whole unit continuum of final goods, the values of α_j will turn out to be uniformly distributed within the unit interval. Abusing a bit the notation, we can thus index a generic final good j by the specific value of $\alpha_j \in [0, 1]$.

3.1.4 Internal Transportation Cost

We assume that transporting intermediate goods across the two regions entails an iceberg trade cost $d_{A,B} = d_{B,A} = d > 1$, where $d_{R,-R}$ indicates the amount of intermediate good that must be shipped from R to $-R$ for one unit of the good to arrive in $-R$. In the cases of intra-regional trade, transportation of intermediate goods is assumed to be costless; that is, $d_{A,A} = d_{B,B} = 1$.

Regions A and B are assumed to be linked to each other by an internal road network of density (or length) $r > 0$. A denser internal road network lowers the interregional trade cost of intermediate goods.[14] Namely,

Assumption 2 *$d = d(r)$, where $d(r) > 1$ and $d'(r) < 0$, for all $r \in \mathbb{R}_+$.*

With regards to final goods, to simplify the analysis, we will assume that they can be *costlessly* transported between region A and B. In our context, costless transportation of final goods will imply that, in equilibrium, the wage must be the same in A and B. Appendix D describes the implications of introducing also an iceberg trade costs for interregional transportation of final goods.[15]

[13] The sector-specific parameters Ψ_j are just posed so that we work with production functions where, in the cases where the intermediate sector aggregators turn out to yield the same price, the marginal cost of production would be identical for all $j \in [0, 1]$ – see equation (4) ahead in the text.

[14] See Felbermayr and Tarasov (2015), Fajgelbaum and Schaal (2017), Santamaria (2018), and Allen and Arkolakis (2019), for papers that explicitly model how increases in the density of the internal transportation infrastructure leads to lower transportation costs of goods.

[15] The reason for the wage to the same in A and B is that, since utility (2) depends only on consumption of final goods, in the absence of internal trade cost of final goods, all final goods will priced identically in both regions, and thus free internal mobility requires that earnings should be equalized in A and B for individuals to be indifferent. Appendix D.1 shows that the presence of interregional trade cost for final goods does not *necessarily* lead to unequal wages in A and B. In particular, given the symmetric structure across regions implied by Assumption 1 and the log utility function (2), when all final goods are subject to the same iceberg cost, the

3.2 Location Choice and Cost of Production of Final Goods

Perfect competition implies that, in equilibrium, each final good j will be sold at a price equal to its marginal cost. Using (3), we can observe that the marginal cost of final good j produced in region R, denoted by $c_{j,R}$, is given by

$$c_{j,R} = p_{0,R}^{1-\alpha_j}\, p_{1,R}^{\alpha_j}, \tag{4}$$

where $p_{0,R}$ and $p_{1,R}$ are, respectively, the CES price indices of intermediate inputs from sector 0 and 1, faced by a firm located in region R.

Letting w denote the wage per unit of labor, and using (1), Assumption 1, $d_{A,B} = d_{B,A} = d(r)$ and $d_{A,A} = d_{B,B} = 1$, the results in Eaton and Kortum (2002) applied to our context imply that for firms located in A:

$$p_{0,A} = \gamma\, w\, \left[(1+T) + d(r)^{-\theta}\right]^{-\frac{1}{\theta}} \quad \text{and} \quad p_{1,A} = \gamma\, w\, \left[1 + (1+T)\, d(r)^{-\theta}\right]^{-\frac{1}{\theta}}, \tag{5}$$

while for firms located in B :

$$p_{0,B} = \gamma\, w\, \left[(1+T)\, d(r)^{-\theta} + 1\right]^{-\frac{1}{\theta}} \quad \text{and} \quad p_{1,B} = \gamma\, w\, \left[d(r)^{-\theta} + (1+T)\right]^{-\frac{1}{\theta}}, \tag{6}$$

where $\gamma \equiv \left[\Gamma\left(1 + \frac{1-\sigma}{\theta}\right)\right]^{1/(1-\sigma)}$ and $\Gamma(\cdot)$ is the Gamma function. Henceforth, we restrict the parameters to satisfy $\sigma < 1 + \theta$, so as to ensure having well-defined price indices.

From (4), (5) and (6), it follows that the marginal costs of j, depending on where it is being produced, are given by

$$c_{j,A} = \gamma\, w\, \left[(1+T) + d(r)^{-\theta}\right]^{-\frac{1-\alpha_j}{\theta}} \left[1 + (1+T)\, d(r)^{-\theta}\right]^{-\frac{\alpha_j}{\theta}} \tag{7}$$

and

$$c_{j,B} = \gamma\, w\, \left[1 + (1+T)\, d(r)^{-\theta}\right]^{-\frac{1-\alpha_j}{\theta}} \left[(1+T) + d(r)^{-\theta}\right]^{-\frac{\alpha_j}{\theta}}. \tag{8}$$

Since final goods are assumed to be transported costlessly between regions, each j will be produced in the region with the lower marginal cost.[16] Comparing $c_{j,A}$ vis-a-vis $c_{j,B}$ for different values of α_j yields the following result:

internal equilibrium will still entail an identical wage in regions A and B. Under alternative assumptions, the equilibrium wage may well differ across regions. Yet, the main qualitative results of the model will in general still hold true even in the presence of interregional income inequality.

[16] This result no longer holds true when final goods are also subject to internal transport cost. In those cases, some final goods will end up being produced in both regions, and will be only sold locally, instead of being traded between regions – see Lemma 5 in Appendix D.1.

Lemma 1 *In equilibrium, the geographic distribution of final good producers is as follows: i) all producers of final goods for which $\alpha_j < 0.5$ locate in region A; ii) all producers of final goods for which $\alpha_j > 0.5$ locate in region B; iii) the producers of the final good for which $\alpha_j = 0.5$ are indifferent between the two regions, and choose randomly their location.*

The result in Lemma 1 states an intuitive geographic agglomeration result. Final good producers choose to locate in the region that exhibits a comparative advantage in the intermediate inputs they use more intensively. Intuitively, by setting up their firms in those regions, final good producers are able to economize relatively more on the transportation of the inputs they more strongly rely on.

Using the result in Lemma 1, together with (7) and (8), we can now write down the marginal cost of final good j:

$$c_j(r) = \begin{cases} \gamma\, w\, \left[(1+T) + d(r)^{-\theta}\right]^{-\frac{1-\alpha_j}{\theta}} \left[1 + (1+T)\, d(r)^{-\theta}\right]^{-\frac{\alpha_j}{\theta}} & \text{if} \quad \alpha_j \leq \frac{1}{2}, \\ \gamma\, w\, \left[1 + (1+T)\, d(r)^{-\theta}\right]^{-\frac{1-\alpha_j}{\theta}} \left[(1+T) + d(r)^{-\theta}\right]^{-\frac{\alpha_j}{\theta}} & \text{if} \quad \alpha_j \geq \frac{1}{2}. \end{cases} \tag{9}$$

The expressions in (9) show that the marginal cost of final good j is affected by the internal transport cost of intermediate inputs that are *not* sourced locally via the term $d(r)$. In that regard, as Lemma 1 shows, final good producers locate in the region that is *main* home of the intermediate sector that they use more intensively so as to minimize the quantity of inputs they must source from the *other* region. However, while this agglomeration pattern allows economizing on transportation cost, it does not do so equally across all sectors. In particular, locating in region A is especially beneficial for final sectors whose α_j is very close to zero, whereas locating in B is so for those whose α_j is very close to one. Instead, when α_j lies near 0.5, locating production in either A or B will not drastically cut down on total transportation costs, since those sectors ultimately rely on a balanced combination of intermediate varieties from sector 0 and 1. As a result, internal transport costs will tend to affect more severely those sectors that use a relatively balanced combination of inputs. Yet, an interesting flip side of this argument is that improvements in r will end up benefiting relatively more exactly those sectors with intermediate values of α_j. The following lemma states this result more formally.

Lemma 2 *Consider two generic values of the internal road network density: $r_1 < r_2$. Then,*

1. *$c_j(r_1)/c_j(r_2) > 1$ for all $\alpha_j \in [0, 1]$.*

2. *The ratio $c_j(r_1)/c_j(r_2)$ is strictly increasing in α_j for all $\alpha_j \in \left[0, \frac{1}{2}\right)$ and strictly decreasing in α_j for all $\alpha_j \in \left(\frac{1}{2}, 1\right]$. Moreover, the highest value of $c_j(r_1)/c_j(r_2)$ is reached at $\alpha_j = \frac{1}{2}$.*

Lemma 2 shows that larger values of r lead to lower costs of production for all final goods. However, the drop in c_j resulting from a higher r turns out to be proportionally greater in sectors with values of α_j closer to $\frac{1}{2}$. This result will prove key in the next section where we allow for international trade, and it will imply that the density of the internal road network will become a source of comparative advantage. In particular, countries with denser road networks will tend to enjoy a comparative advantage in final sectors that rely on a relatively balanced mix of intermediate inputs (i.e., those characterized by values of α_j around $\frac{1}{2}$).

4 Two-Country Model

We now incorporate our previous framework into a Dornbusch-Fischer-Samuelson (1977) world economy with two countries: H and F. We initially keep most of the main features of the environment in Section 3 essentially intact. Some of these specific features are subsequently relaxed, so as to show how the main results extend to more general setups.

Geography and Population: Country H and country F are both characterized by an identical internal geography, comprising two separate regions: A_H and B_H in H; A_F and B_F in F. To ease notation, whenever creating no confusion, we will skip the use of country subscripts for their respective regions, and simply refer them as region A or B in country $C = H, F$. Both countries are populated by a mass L of individuals. Each individual is endowed with one unit of labor that is supplied inelastically to local firms. There is free labor mobility within countries. In contrast, labor is *immobile* between countries. We let w_H and w_F denote the wage in H and in F, respectively. Henceforth, we set $w_F = 1$, and use $\omega \equiv w_H/w_F$ to denote the relative wage of H with respect to F. All individuals in the world economy share the same preferences, given by (2). All markets are subject to perfect competition everywhere.

Final Goods Sectors: The technologies to produce final goods are identical in both countries, and are given by the two-level CES nested production functions in (3).[17] All final goods are internationally tradeable. To keep the analysis simple and notation light, we initially assume that international trade of final goods is costless. (This simplifying assumption is next dropped in Section 4.2 to show how the results carry through with costly international trade.)

Intermediate Sectors: Unlike for final goods, we assume that intermediate goods are all non-tradeable internationally.[18] The technology in the intermediate sectors is directly extrapolated

[17]Section 4.3.1 relaxes this assumption, introducing country-specific technologies to produce final goods.

[18]Section 4.3.2 relaxes this assumption, allowing also intermediate goods to be traded internationally.

from Section 3.1.2: producing one unit of variety i of sector S in region $R_C = A_C, B_C$ (with $C = H, F$) requires $1/Z_{R_C,S}(i)$ units of labor, where each $Z_{R_C,S}(i)$ is independently drawn from a Fréchet distribution like (1). Regarding each $T_{R_C,S} \in \{A_C, B_C\} \times \{0,1\}$, we assume that they are given in both H and F by Assumption 1.[19]

Internal Transportation Cost: Shipping intermediate goods between C's regions entails an iceberg cost $d(r_C)$, where r_C is the density of C's internal road network, with $C = H, F$. The properties of function $d(\cdot)$ are still given by Assumption 2. Throughout the paper, we let H's internal road network be denser than F's. Namely,

Assumption 3 $r_H > r_F$.

Assumption 3 will have two main implications in the context of our world economy model. First (and more importantly in our context), it will confer a source of comparative advantage to H in the types of final goods that rely more strongly on efficient internal transportation of inputs. Second, the fact that H will be able to ship inputs internally at lower cost than F will in turn grant a source of aggregate absolute advantage by H over F.[20]

4.1 Equilibrium and Specialization with Frictionless International Trade

Costless international trade of final goods (coupled with perfect competition) implies that consumers will buy each final good j from whoever can produce it at the lowest marginal cost. Seeking to minimize marginal costs, final good producers in H and in F will locate in region A of their country when $\alpha_j \leq 0.5$, and in region B of their country when $\alpha_j \geq 0.5$. Should final good j be produced *somewhere* in country C, it would be thus sold in both H and in F at price given by

$$P_j^C = \begin{cases} \gamma\, w_C \,\left[(1+T) + d(r_C)^{-\theta}\right]^{-\frac{1-\alpha_j}{\theta}} \left[1 + (1+T)\, d(r_C)^{-\theta}\right]^{-\frac{\alpha_j}{\theta}} & \text{if} \quad \alpha_j \leq \frac{1}{2} \\ \gamma\, w_C \,\left[1 + (1+T)\, d(r_C)^{-\theta}\right]^{-\frac{1-\alpha_j}{\theta}} \left[(1+T) + d(r_C)^{-\theta}\right]^{-\frac{\alpha_j}{\theta}} & \text{if} \quad \alpha_j \geq \frac{1}{2}, \end{cases} \tag{10}$$

with $C = H, F$, and where the superscript in P_j^C denotes the country of production of j.

[19] More rigorously, (1) should be now written $F_{R_C,S}(Z) = \exp\left(-T_{R_C,S}\, Z^{-\theta}\right)$, with $C = H, F$. Similarly, Assumption 1 should now state: $T_{A_C,0} = T_{B_C,1} = 1 + T$, with $T > 0$, and $T_{A_C,1} = T_{B_C,0} = 1$.

[20] Like in Section 3, we keep assuming costless internal transportation of final goods. In Appendix D.2, we show how the main results of the model go through when internal transportation of final goods is also costly.

To ease notation, we will let henceforth $d_C \equiv d(r_C)$, and denote by $\lambda > 1$ the ratio of the internal iceberg trade cost in F relative to H (where $\lambda > 1$ follows from Assumption 3). Namely,

$$\lambda \equiv \frac{d_F}{d_H} > 1. \tag{11}$$

When international trade is costless, each final good j will be actively produced in only one single country. Recalling that $w_F = 1$ and $w_H = \omega$, and using (11), it follows from (10) that H will produce final good j (and F will import it from there) when

$$\omega < \begin{cases} \left[\dfrac{(1+T)d_H^\theta + 1}{(1+T)d_H^\theta + \lambda^{-\theta}}\right]^{(1-\alpha_j)/\theta} \left[\dfrac{(1+T) + d_H^\theta}{(1+T)\lambda^{-\theta} + d_H^\theta}\right]^{\alpha_j/\theta} & \text{if } \alpha_j \leq 0.5, \\ \left[\dfrac{(1+T) + d_H^\theta}{(1+T)\lambda^{-\theta} + d_H^\theta}\right]^{(1-\alpha_j)/\theta} \left[\dfrac{(1+T)d_H^\theta + 1}{(1+T)d_H^\theta + \lambda^{-\theta}}\right]^{\alpha_j/\theta} & \text{if } \alpha_j \geq 0.5. \end{cases} \tag{12}$$

Conversely, when (12) fails to hold, j will be produced in F (and H will import it form there).

The expression in (12) valid for $\alpha_j \leq 0.5$ is strictly increasing in α_j, whereas the one valid for $\alpha_j \geq 0.5$ is strictly decreasing in α_j. As a consequence, we can define two thresholds

$$\underline{\alpha}(\omega) \equiv \frac{\ln\left(\omega^\theta\right) - \ln\left[\frac{(1+T)d_H^\theta+1}{(1+T)d_H^\theta+\lambda^{-\theta}}\right]}{\ln\left[\frac{(1+T)+d_H^\theta}{(1+T)\lambda^{-\theta}+d_H^\theta} \times \frac{(1+T)d_H^\theta+\lambda^{-\theta}}{(1+T)d_H^\theta+1}\right]} \quad \text{and} \quad \overline{\alpha}(\omega) \equiv \frac{\ln\left[\frac{(1+T)+d_H^\theta}{(1+T)\lambda^{-\theta}+d_H^\theta}\right] - \ln\left(\omega^\theta\right)}{\ln\left[\frac{(1+T)+d_H^\theta}{(1+T)\lambda^{-\theta}+d_H^\theta} \times \frac{(1+T)d_H^\theta+\lambda^{-\theta}}{(1+T)d_H^\theta+1}\right]}, \tag{13}$$

such that (12) holds true for all $\underline{\alpha}(\omega) < \alpha_j < \overline{\alpha}(\omega)$.[21] It may also be noticed from (13) that $\underline{\alpha}(\omega) = 1 - \overline{\alpha}(\omega)$. Hence, $\underline{\alpha}(\omega)$ and $\underline{\alpha}(\omega)$ will lie symmetrically around the midpoint $\alpha_j = \frac{1}{2}$.

In equilibrium, the total (world) spending on final goods produced in each country must equal the total labor income of each country. In a two-country setup, this condition can be restated as a trade balance equilibrium for either H or F. The utility function (2) implies that consumers allocate identical expenditure shares across all final goods in the optimum. As a result, in equilibrium:

$$[\underline{\alpha}(\omega) + (1 - \overline{\alpha}(\omega))]\,\omega = \overline{\alpha}(\omega) - \underline{\alpha}(\omega)$$

where $\underline{\alpha}(\omega)$ and $\overline{\alpha}(\omega)$ are given in (13). Bearing in mind that $\underline{\alpha}(\omega) = 1 - \overline{\alpha}(\omega)$, the trade balance equilibrium boils down to:

$$2\underline{\alpha}(\omega^*)(1 + \omega^*) = 1, \tag{14}$$

where ω^* denotes the equilibrium relative wage between H and F.

From (14), together with (13), we can obtain our first result concerning the equilibrium relative wage, ω^*.

[21] Note that, in equilibrium, it will necessarily be the case that $0 < \underline{\alpha}(\omega) < \overline{\alpha}(\omega) < 1$.

Proposition 1 *In equilibrium, the wage in H is strictly greater than in F (i.e., $\omega^* > 1$). Furthermore, ω^* is strictly increasing in λ (i.e., $\partial\omega^*/\partial\lambda > 0$), converges to 1 as r_F approaches r_H (i.e.,$\lim_{\lambda\to 1}\omega^*(\lambda) = 1$), and $\underline{\omega}(\lambda) < \omega^*(\lambda) < \overline{\omega}(\lambda)$ where*

$$\underline{\omega}(\lambda) \equiv \left[\frac{(1+T)d_H^{\theta}+1}{(1+T)d_H^{\theta}+\lambda^{-\theta}}\right]^{\frac{1}{\theta}} \quad \text{and} \quad \overline{\omega}(\lambda) \equiv \left[\frac{(1+T)+d_H^{\theta}}{(1+T)\lambda^{-\theta}+d_H^{\theta}} \times \frac{(1+T)d_H^{\theta}+1}{(1+T)d_H^{\theta}+\lambda^{-\theta}}\right]^{\frac{1}{2\theta}}.$$

The result $\omega^* > 1$ is a straightforward implication of the fact that Assumption 3 conveys an aggregate advantage by H over F. Hence, in equilibrium, ω must rise above one, in order to allow F to be able to export some goods to H. Notice that since labor is the only non-reproducible input in our model, wages are also equal to income per head in each country.

Consumers in both H and F will buy good j from the producer who can sell it in each market at the lower price. Hence, by using (12), (13) and (14), we can next derive the equilibrium patterns of trade and specialization.

Proposition 2 *When international trade of final goods is frictionless and $r_H > r_F$, country H specializes in final goods with α_j located around the mid-point of the unit interval, while F specializes in final goods located on both ends of the unit interval. More precisely,*

- *H produces (and exports to F) all the final goods with $\alpha_j \in [\underline{\alpha}(\omega^*), 1-\underline{\alpha}(\omega^*)]$, where $\underline{\alpha}(\omega^*) \in \left(0, \frac{1}{2}\right)$ stems from (13) and (14).*
- *F produces (and exports to H) all the final goods with $\alpha_j \in [0, \underline{\alpha}(\omega^*)]$ and $\alpha_j \in [1-\underline{\alpha}(\omega^*), 1]$.*

The pattern of specialization described by Proposition 2 represents the main insight of the model. Owing to the geographically diffuse distribution of input sources, the density of the internal road network becomes a source of comparative advantage across industries differing in their α_j. Final goods with midway values of α_j require a relatively even use of all intermediate varieties. This, in turn, means that a large share of their inputs will *necessarily* have to be transported along the road network. On the other hand, firms producing final goods with either *high* or *low* values of α_j are able to source a large share of their intermediate inputs from the *same* location as where their firms are placed. As such, sectors in the intermediate range of α_j require stronger use of the internal transport network than those whose α_j lies on the upper and lower spectrum of the unit interval. Thus, when H has a denser road network than F, the former then ends up specializing in the final goods with intermediate values of α_j, while the latter in those with more extreme values of α_j.

4.2 Equilibrium and Specialization under Costly International Trade

We now extend the results of Section 4.1 to a context in which shipping final goods between H and F is costly. We assume that $\tau > 1$ units of final good j must be shipped from the exporter of j in order for the importer to receive one unit of j.[22] Costly international trade changes the previous results in two ways. Firstly, consumer prices will no longer be the identical in H and F. Secondly, not all final goods will be turn out to be traded internationally, and thus there will be incomplete international specialization across the set of final goods.

For goods that are traded internationally, the price paid by importers will now incorporate the iceberg trade cost τ. Hence, when j is produced *only* by country C, it will be sold in country M at a price equal to $P_{j,M}^{C} = [\mathbb{I}_{M \neq C} \cdot \tau + (1 - \mathbb{I}_{M \neq C})] \cdot c_j^C$, where $\mathbb{I}_{M \neq C}$ is an indicator function that is equal to 1 when $M \neq C$ and 0 otherwise, and c_j^C denotes the marginal cost of producing j in C. On the other hand, if j remains untraded, the price in H (resp. F) will be $P_{j,H}^{H} = c_j^H$ (resp. $P_{j,F}^{F} = c_j^F$). Country H will then be exporting good j to F when $\tau c_j^H < c_j^F$, while H will be importing j from F when $c_j^H > \tau c_j^F$. Positive international trade requires then that τ is not too large. The following lemma stipulates the conditions under which there is positive trade between H and F in equilibrium.

Lemma 3 *When international trade of final goods entails an iceberg trade cost $\tau > 1$, the equilibrium of the two-country model will feature positive international trade if and only if $\tau < \widehat{\tau}$, where $\widehat{\tau} > 1$ and is given by:*

$$\widehat{\tau} = \left[\frac{(1+T) + d_H^{\theta}}{(1+T)\,\lambda^{-\theta} + d_H^{\theta}} \times \frac{(1+T)\, d_H^{\theta} + \lambda^{-\theta}}{(1+T)\, d_H^{\theta} + 1} \right]^{\frac{1}{4\theta}} . \qquad (15)$$

The presence of $\tau > 1$ will naturally hamper international trade. When τ is greater than the threshold $\widehat{\tau}$ any scope for trade between H and F vanishes away. Notice that, given that both H and F share the same technologies to produce final goods and the same Fréchet distribution to generate the labor productivities (i.e., the $Z(i)'s$), the only source of comparative advantage in the model is given by the fact that $\lambda > 1$ (i.e., by Assumption 3). In that respect, it is

[22] Notice that this implicitly assumes that the international trade cost between *any* pair of foreign regions is equal to τ. In other words, we disregard the possibility that one of the two regions is closer to international markets than the other one, like in models that study the division of industries between (export-oriented) coastal regions and (local-market-oriented) interior regions – e.g., Coşar and Fajgelbaum (2016). Allowing the internal geography of our model to feature a coastal and an interior region will certainly have implications in terms of regional specialization. However, this will not alter *per se* the main prediction of our model regarding patterns of *international* trade and specialization. Namely, H will still have a comparative advantage in final goods with broader intermediate input bases, while F will have it in final goods with narrower input bases.

interesting to note from (15) that the threshold $\widehat{\tau}$ is strictly increasing in λ (while $\widehat{\tau}$ approaches one as $\lambda \to 1$). In other words, the scope for international trade increases as the disparity between r_H and r_F widens.

Given that we are interested in studying the equilibrium patterns of international trade between H and F, we concentrate henceforth on the cases that verify $\tau < \widehat{\tau}$. Under such restriction, we extend below Proposition 2 to a context with costly international trade.

Proposition 3 *Let $r_H > r_F$, and assume trading final goods between H and F entails an iceberg cost $\tau > 1$. Assume also that $\tau < \widehat{\tau}$, where $\widehat{\tau}$ is given by (15). Then, there exist $\underline{\alpha}_F$ and $\underline{\alpha}_H$, where $0 < \underline{\alpha}_F < \underline{\alpha}_H < \frac{1}{2}$, such that, in equilibrium:*

- *H exports to F all the final goods whose $\alpha_j \in [\underline{\alpha}_H, 1 - \underline{\alpha}_H]$.*
- *F exports to H all the final goods whose $\alpha_j \in [0, \underline{\alpha}_F]$ and $\alpha_j \in [1 - \underline{\alpha}_F, 1]$.*
- *All the final goods whose $\alpha_j \in (\underline{\alpha}_F, \underline{\alpha}_H)$ and $\alpha_j \in (1 - \underline{\alpha}_H, 1 - \underline{\alpha}_F)$ end up being sourced from national producers in both H and F.*

Proposition 3 shows that the main result of Section 4.1 carries through (in qualitatively terms) when international trade of final goods is costly: H exports goods with α_j located around the mid-point of the unit interval, while F exports those with α_j lying towards both ends of the spectrum of the unit interval. In other words, H specializes in goods with broad intermediate input bases, and F in goods with narrow input bases. The main difference with respect to Proposition 2 is that now we also have two subsets of goods that are produced by *both* countries and remain *untraded* in equilibrium. The reason for the subsets $(\underline{\alpha}_F, \underline{\alpha}_H)$ and $(1 - \underline{\alpha}_H, 1 - \underline{\alpha}_F)$ to remain untraded is that for them the comparative advantage stemming from the fact that $r_H > r_F$ is not strong enough to overcome the trade cost τ.[23]

4.3 Extensions: Robustness of the Specialization Predictions

We now relax two of the simplifying assumptions embedded in the previous analysis, and show how the main features of the patterns of specialization continue to hold under more general contexts. In the first part, we let final goods technologies to possibly differ across countries. In the second, we allow for international trade of intermediate goods alongside that of final goods.

[23] One last aspect to highlight here is that, when the equilibrium with costly international trade features positive trade, $\omega^* > 1$ will still hold true as in Proposition 1. Furthermore, $\omega^* > 1$ will hold true *even* in certain cases where $\tau > \widehat{\tau}$, and no trade takes place in equilibrium between H and F. The reason why $\omega^* > 1$ may hold even in the absence of trade, is that this is needed in order to drive imports by F from H to zero. Nevertheless, for values of τ sufficiently above $\widehat{\tau}$, the model will give rise to an equilibrium with $\omega^* = 1$ and zero trade.

4.3.1 Country-Specific Technologies of Final Goods

So far we have assumed that both country H and F have access to the same technologies to produce final goods. We now relax this assumption, replacing (3), by the following country-specific production functions:

$$Y_{j,H} = \zeta_{j,H} Y_j \quad \text{and} \quad Y_{j,F} = \zeta_{j,F} Y_j, \tag{16}$$

where Y_j is given by (3) for $j \in [0,1]$. We assume that the country specific parameters $\zeta_{j,H}$ and $\zeta_{j,F}$ are independently drawn from a Fréchet distribution function with location parameter equal to 1 and shape parameter $\vartheta > 1$:

$$F_C(\zeta) = \exp\left(-\zeta^{-\vartheta}\right). \tag{17}$$

Given the country-specific production functions (16), the marginal cost of producing final good j in country C can now be written as

$$c_j^C = \frac{1}{\zeta_{j,C}} \chi_{j,C}, \quad \text{with } C = H, F.$$

where $\chi_{j,C}$ denotes the intermediate inputs cost used for final good j in country C. Recall that $\chi_{j,C} = \gamma w_C \left[(1+T) + d_C^{-\theta}\right]^{-(1-\alpha_j)/\theta} \left[(1+T)\, d_C^{-\theta} + 1\right]^{-\alpha_j/\theta}$ when $\alpha_j \leq 0.5$, and $\chi_{j,C} = \gamma w_C \left[(1+T)\, d_C^{-\theta} + 1\right]^{-(1-\alpha_j)/\theta} \left[(1+T) + d_C^{-\theta}\right]^{-\alpha_j/\theta}$ when $\alpha_j \geq 0.5$. Using $w_F = 1$, $\omega = w_H/w_F$, and $d_F = \lambda d_H$, we then can observe that

$$\frac{\chi_{j,H}}{\chi_{j,F}} = \begin{cases} \omega \left[\dfrac{(1+T) + (\lambda d_H)^{-\theta}}{(1+T) + d_H^{-\theta}}\right]^{(1-\alpha_j)/\theta} \left[\dfrac{(1+T)\,(\lambda d_H)^{-\theta} + 1}{(1+T) d_H^{-\theta} + 1}\right]^{\alpha_j/\theta} & \text{if } \alpha_j \leq 0.5, \\[2ex] \omega \left[\dfrac{(1+T)\,(\lambda d_H)^{-\theta} + 1}{(1+T) d_H^{-\theta} + 1}\right]^{(1-\alpha_j)/\theta} \left[\dfrac{(1+T) + (\lambda d_H)^{-\theta}}{(1+T) + d_H^{-\theta}}\right]^{\alpha_j/\theta} & \text{if } \alpha_j \geq 0.5. \end{cases} \tag{18}$$

Differentiating (18) with respect to α_j, the following result obtains:

$$\begin{aligned} \frac{\partial\left(\chi_{j,H}/\chi_{j,F}\right)}{\partial \alpha_j} &< 0, \quad \text{for all } 0 \leq \alpha_j < \frac{1}{2}; \\ \frac{\partial\left(\chi_{j,H}/\chi_{j,F}\right)}{\partial \alpha_j} &> 0, \quad \text{for all } \frac{1}{2} < \alpha_j \leq 1. \end{aligned} \tag{19}$$

The fact that $\zeta_{j,H}$ and $\zeta_{j,F}$ are all independently drawn from (17) implies that, for any final good $j \in [0,1]$, the probability that either of the two countries ends up exporting this good to the other country is *never* zero. Using the properties of (17), it follows that the probability that country C exports final good j to the other country (i.e., $-C$), denoted by $\pi_C(j)$, equals

$$\pi_C(j) = \frac{1}{1 + \left(\dfrac{\chi_{j,-C}}{\tau\, \chi_{j,C}}\right)^{-\vartheta}}. \tag{20}$$

Bearing in mind (19) and (20), we can state now the following proposition linking the probability that either H or F ends up exporting a final good with a given value of α_j.

Proposition 4 *Let $\pi_C(\alpha_j)$ denote the probability that country $C = H, F$ exports the final good with Cobb-Douglas weight parameter $\alpha_j \in [0, 1]$ to the other country. Then:*

1. *$\partial\pi_H/\partial\alpha_j > 0$ for all $0 \leq \alpha_j < \frac{1}{2}$, and $\partial\pi_H/\partial\alpha_j < 0$ for all $\frac{1}{2} < \alpha_j \leq 1$.*
2. *$\partial\pi_F/\partial\alpha_j < 0$ for all $0 \leq \alpha_j < \frac{1}{2}$, and $\partial\pi_F/\partial\alpha_j > 0$ for all $\frac{1}{2} < \alpha_j \leq 1$.*
3. *The ratio $\pi_H(\alpha_j)/\pi_F(\alpha_j)$ reaches its highest value at $\alpha_j = \frac{1}{2}$. Moreover, in equilibrium, $\pi_H\left(\alpha_j = \frac{1}{2}\right) > \pi_F\left(\alpha_j = \frac{1}{2}\right)$.*

The above proposition extends our previous result concerning the patterns of specialization to a context where H and F may also differ in their productivity to transform intermediate good bundles into final goods. Although, in this context, we no longer can obtain a deterministic pattern of specialization for each j, Proposition 4 shows that the ratio $\pi_H(\alpha_j)/\pi_F(\alpha_j)$ will be strictly increasing in α_j whenever $\alpha_j < \frac{1}{2}$, and it will be strictly decreasing in α_j when this parameter is above $\frac{1}{2}$. In other words, H's exports will tend to be overrepresented by final goods with intermediate values of α_j, whereas F's exports will be overrepresented by final goods with values of α_j close to both extremes of the unit interval. Furthermore, the third result in Proposition 4 implies that the probability that H will export goods with α_j sufficiently close to $\frac{1}{2}$ is larger than the probability that those goods end up being exported by F.

4.3.2 International Trade of Intermediate Inputs

We now allow international trade of intermediate input varieties. We assume that trading intermediate goods between H and F entails an iceberg cost $\iota > 1$. Throughout this subsection, we also assume that the international transportation cost of intermediate goods is greater than its internal transportation, even for the country with more costly internal transportation:

$\iota > d_F = \lambda d_H$.[24] We maintain the assumption that international trade of final goods is subject to an iceberg cost $\tau > 1$. We also keep assuming that the technologies for final goods are country-specific, and given by (16) together with (17).

The possibility to import intermediate goods will imply that the marginal cost of producing final good j will encompass now, not only prices from national inputs, but also prices from imported inputs. Applying again the results in Eaton and Kortum (2002) to this framework, we can obtain that the marginal cost of producing final good j in H, given the relative wage ω, is given by $c_j^H = \chi_{j,H}/\zeta_{j,H}$ where now

$$\chi_{j,H} = \begin{cases} \gamma \left[\frac{(1+T)+d_H^{-\theta}}{\omega^\theta} + \frac{2+T}{\iota^\theta} \right]^{-\frac{1-\alpha_j}{\theta}} \left[\frac{(1+T)\,d_H^{-\theta}+1}{\omega^\theta} + \frac{2+T}{\iota^\theta} \right]^{-\frac{\alpha_j}{\theta}} & \text{if } \alpha_j \leq \frac{1}{2}, \\ \gamma \left[\frac{(1+T)\,d_H^{-\theta}+1}{\omega^\theta} + \frac{2+T}{\iota^\theta} \right]^{-\frac{1-\alpha_j}{\theta}} \left[\frac{(1+T)+d_H^{-\theta}}{\omega^\theta} + \frac{2+T}{\iota^\theta} \right]^{-\frac{\alpha_j}{\theta}} & \text{if } \alpha_j \geq \frac{1}{2}. \end{cases} \tag{21}$$

Similarly, the marginal cost of producing j in F is given by $c_j^F = \chi_{j,F}/\zeta_{j,F}$ where now

$$\chi_{j,F} = \begin{cases} \gamma \left[(1+T) + \frac{1}{(\lambda d_H)^\theta} + \frac{2+T}{(\iota\omega)^\theta} \right]^{-\frac{1-\alpha_j}{\theta}} \left[\frac{1+T}{(\lambda d_H)^\theta} + 1 + \frac{2+T}{(\iota\omega)^\theta} \right]^{-\frac{\alpha_j}{\theta}} & \text{if } \alpha_j \leq \frac{1}{2}, \\ \gamma \left[\frac{1+T}{(\lambda d_H)^\theta} + 1 + \frac{2+T}{(\iota\omega)^\theta} \right]^{-\frac{1-\alpha_j}{\theta}} \left[(1+T) + \frac{1}{(\lambda d_H)^\theta} + \frac{2+T}{(\iota\omega)^\theta} \right]^{-\frac{\alpha_j}{\theta}} & \text{if } \alpha_j \geq \frac{1}{2}. \end{cases} \tag{22}$$

Comparing (21) vis-a-vis (22), it can be observed that the ratio $\chi_{j,H}/\chi_{j,F}$ is (again) strictly decreasing in α_j for all $\alpha_j \in \left[0, \frac{1}{2}\right)$, and (again) strictly increasing in α_j for all $\alpha_j \in \left(\frac{1}{2}, 1\right]$. (The formal proof of this result is provided at the end of Appendix A.) Notice now that the probability that country C exports final good j to the other country (i.e., $\pi_C(j)$), is still given by the expression in (20), with $\chi_{j,C}$ and $\chi_{j,-C}$ given in this case by (21)-(22) as they correspond to each case. As a consequence, an analogous result at that one in Proposition 4 also holds true when intermediate inputs can be traded internationally alongside final goods.

[24] The reason for this assumption is that, given that final goods are produced only by combining intermediate inputs, a higher international transportation cost of inputs relative to the internal transportation cost ensures (in particular, it is a *sufficient condition* though not a *necessary condition*) that an increase in the relative wage of H over F (i.e., a higher ω) will have a stronger impact on the marginal cost of final goods produced in H than in those produced in F. It is important to bear in mind that the model assumes that the internal transport cost inputs is not added to the international transport cost of them, which in general is a rather unrealistic assumption (unless we are thinking of a coastal region, which happens to be also closest to the foreign source).

5 Empirical Predictions: From the Theory to the Data

In this section, we first describe how we attempt to bring to the data the main variables of interest present in the model. Next, we explain how we approach the data on trade flows to seek for evidence consistent with the main predictions of the model.

5.1 Main Variables of Interest

Input Narrowness

The first task is coming up with a measure of the breadth of the set of intermediate inputs used by each industry. The model is quite stylized to allow a direct match between its technological environment and real world data on inputs and outputs by sectors. In particular, production functions tend to rely on more than only two intermediate sectors. In addition, the distinction between intermediate and final goods is not so clear-cut as assumed by the model, as many goods satisfy both roles. Despite these shortcomings, we can still use the model as a guide to construct measures of narrowness of industries' inputs base.

In the model, (final) sector j allocates a fraction $1 - \alpha_j$ of their total spending in intermediate goods from (intermediate) sector 0 and the remainder α_j in intermediate goods from (intermediate) sector 1. This means that industries with very low or very high values of α_j source most of their inputs from only one intermediate sector, and thus exhibit a *narrow* intermediate input base. Conversely, industries with values of α_j around one half rely quite heavily on both intermediate sectors, and thus display a *wide* intermediate input base.

We measure *narrowness* of the input base of sector j by the Gini coefficient of their expenditure shares across both inputs ($Gini_j$).[25] By using the fact that expenditure shares on input from sector 0 and 1 are given, respectively, by $1 - \alpha_j$ and α_j, we can observe that:

$$Gini_j = \begin{cases} \frac{1}{2} - \alpha_j & \text{if } 0 \leq \alpha_j \leq \frac{1}{2} \\ \alpha_j - \frac{1}{2} & \text{if } \frac{1}{2} \leq \alpha_j \leq 1. \end{cases} \tag{23}$$

To construct a measure of input narrowness analogous to that one in (23), but based on the available real world data, we resort to the input-output (IO) matrix of the US in 2007 from the Bureau of Economic Analysis (BEA).[26] The IO matrix comprises 389 sectors/industries.

[25] Imbs and Wacziarg (2003) have previously used the Gini coefficient to measure the degree of concentration of labor and value added across different economic sectors. We apply the same methodology as them, but in this case we use it to measure the degree of narrowness or concentration of the intermediate input bases of different industries.

[26] This data is publicly available from https://www.bea.gov/industry/io_annual.htm.

Although the IO matrix exhibits the same number of sectors producing intermediate goods as those producing final output, we restrict the set of final goods to those also present in the international trade data (see footnote 28 below). Thus, we index by $k = 1, 2, ..., K$ each of the sectors present in the IO matrix and also in the trade data, and by $n = 1, 2, ..., N$ each of the sectors selling intermediate inputs.

We let $X_{k,n} \geq 0$ denote the total value of intermediate good n purchased by sector k. Defining $S_{k,n} \equiv X_{k,n} / \sum_{n=1}^{N} X_{k,n} \geq 0$ as the share of n over the total value of intermediates purchased by k, we can compute the Gini coefficients analogously to those in (23). Namely,

$$Gini_k = \frac{2 \times \sum_{n=1}^{N} n \times S_{k,n}}{N \times \sum_{n=1}^{N} S_{k,n}} - \frac{N+1}{N}, \tag{24}$$

where the argument $\sum_{n=1}^{N} n \times S_{k,l}$ in the numerator of $Gini_k$ is ordering intermediates in non-decreasing order (i.e., $S_{k,n} \leq S_{k,n+1}$).

In the empirical analysis in Section 5.3, we use $Gini_k$ to measure the degree of narrowness of the input base of sector k. Large values of $Gini_k$ are the result of sector k sourcing most of their intermediate inputs from relatively few sectors. Conversely, small values of $Gini_k$ tend to occur when the distribution of $S_{k,n}$ is quite evenly spread across a large number of intermediates.[27] Notice finally the link between $Gini_k$ in (24) and $Gini_j$ in (23): the former boils down to the latter when $N = 2$, and $S_{k,n} = \alpha_{k,n}$ with $\alpha_{k,1} + \alpha_{k,2} = 1$.

Export Specialization

To measure the degree of export specialization by sectors we use the data on trade flows from COMTRADE compiled by Gaulier and Zignago (2010). We use only trade flows in year 2014. The data are categorized following the Harmonized System (HS) 6-digit classification, with 5,192 products. We map the trade flows data based on the HS 6-digit classification to the BEA industry codes using the concordance table between the 2002 IO matrix commodity codes and the HS 10-digit classification from the BEA website (after grouping the HS 10-digit codes into HS 6-digit products).[28] In the cases in which an HS-6 product maps into more than one

[27] In the extreme (unequal) case in which $S_{k,n'} = 1$ for some n' and $S_{k,n} = 0$ for all $n \neq n'$, (24) yields $Gini_k = 2 - [(N+1)/N]$, which approaches 1 as $N \to \infty$. On the other hand, in the case when $S_{k,n} = S_k > 0$ for all $n = 1, ..., N$, we would have $Gini_k = 0$.

[28] When mapping the HS 6-digit products into the BEA 2002 codes, several of the original industries in the Input-Output matrix are lost due to lack of export data on them (these are mainly non-tradeable services). In the end, after merging the COMTRADE export data with the IO matrix information, we are left with data on trade flows and input narrowness for 294 industries as coded by the BEA 2002 classification.

BEA code, we assign trade flows proportionally to each of the BEA sectors which it maps into.[29] Lastly, the IO industry codes of the 2002 classification are matched to those of the 2007 classification, which are the ones actually used in the computation of the Gini coefficients.[30]

Road Network

The last key variable in our model is the length of the road network of country c (r_c). We take the road network length by countries from the data on roadways from the CIA World Factbook. Roadways are defined as 'total length of the road network, including paved and unpaved portions'. The year of the data point for each country varies, ranging from year 2000 to 2016, with the median year of the sample being 2010. (See details of this variable in Table A.9 in Appendix C.) When defining our empirical counterpart of the variable r_c, we divide the length of the road network by the total area of the country: $r_c \equiv roadways_c/area_c$. In some of the robustness checks, we use also two additional measures of transport density: waterways density (defined as $waterways_c/area_c$) and railway density (defined as $railways_c/area_c$). The data on length of waterways and railways are also taken from the CIA World Factbook.

5.2 Road Density and Patterns of Specialization: Testing the predictions of the model

The two-country model presented in Section 4 predicts that the country with the denser road network (i.e., country H) will tend to export goods with intermediate values of α_j, while the country with the sparser road network (i.e., country F) will tend to specialize in goods with either high or low values of α_j. Conceptually, this prediction can be interpreted as stating that countries with denser road networks will tend to exhibit a comparative advantage in the types of industries that rely on a broader (or more diverse) set of intermediate inputs.

From an empirical viewpoint, if road network length differences across countries shaped

[29] There are 526 HS-6 products that map into two BEA Input-Output industry codes, 96 products that map into three IO codes, 33 products that map into four IO codes, and 11 products that map into five or more IO codes. (We excluded the 11 products that map into five or more IO codes.) None of the regression results in Section 5.3 are significantly altered when all the HS-6 products that map into more than one BEA Input-Output industry code are dropped from the sample.

[30] Unfortunately, we are not aware of any correspondence table between BEA 2007 codes and the HS codes (only BEA 2002 codes are matched to the 10-digit HS commodities via the correspondence table in https://www.bea.gov/industry/benchmark-input-output-data). As a consequence of this, we link the industry input narrowness measures and total exports at the industry level via the BEA 2002 codes.

somehow their patterns of specialization as our model predicts, we should then observe the following: economies with a greater r_c will tend to export relatively more of the goods produced in industries with a smaller value of $Gini_k$ vis-a-vis economies with smaller r_c. We test this prediction using the following regression:

$$\ln(Expo_{c,k}) = \beta \cdot (r_c \times Gini_k) + \boldsymbol{\chi} \cdot \boldsymbol{\Delta}_{k,c} + \varsigma_c + \kappa_k + \upsilon_{c,k}. \tag{25}$$

In the regression equation (25) the dependent variable is given by the logarithm of the total value of exports in industry k by country c to all other countries in the world in year 2014. The term $(r_c \times Gini_k)$ interacts the measure of input narrowness defined in (24) with the measure of road density (i.e., length of roadways per square kilometer). $\boldsymbol{\Delta}_{k,c}$ denotes a vector of additional covariates that may possibly influence specialization across countries in industries differing in terms of the degree of input narrowness. ς_c and κ_k denote country fixed effects and industry fixed effects, respectively, and $\upsilon_{c,k}$ represents an error term.[31]

The main coefficient of interest in (25) is β. If countries with a denser road network (i.e., countries with a greater r_c) indeed tend to exhibit a comparative advantage in industries that require a wider set intermediate inputs (i.e., industries with a smaller $Gini_k$), then the data should deliver a negative estimate of β.

5.3 Empirical Results

Table I displays the first set of estimation results corresponding to (25). Column (1) includes only our main variable of interest (i.e., $r_c \times Gini_k$), together with the exporter and industry dummies. The correlation is negative and highly significant, suggesting that countries with denser road networks tend to export relatively more of the final goods whose production process requires a broader set of intermediates (i.e., those exhibiting a lower $Gini_k$).

In the remaining columns of Table I, we sequentially incorporate additional interaction terms that may also influence the patterns of specialization across industries with different levels of input narrowness. Column (2) adds an interaction term between $Gini_k$ and an index of *Rule of Law*, taken from World Governance Indicators. The rationale behind including this term lies on the argument in Levchenko (2007) and Nunn (2007), who show that countries with

[31] The interaction-term approach followed by regression (25) is analogous to that in used by Rajan and Zingales (1998) to assess whether countries with deeper financial sectors grow proportionally more in industries that rely more strongly on external finance; a similar apparoach is also used by Manova (2013) but with focus placed on bilateral exports. Similar regression equations with interaction terms between country-level and industry-level variables can also be found in Romalis (2004), Nunn (2007), Levchenko (2007), Chor (2010), among others.

better contract enforcement institutions display a comparative advantage in industries that are heavily dependent on relationship-specific investments. Within our context, industries that need to source a broader set of intermediates may benefit relatively more from a sound legal environment, as they need to establish relationships with a greater number of input providers. Given that countries with better institutions tend to be also richer and invest more in basic infrastructure, omitting this term could lead to an overestimation (in absolute value) of the correlation coefficient of interest in (25). The regression in column (2) yields indeed a negative and significant coefficient associated with the interaction term between rule of law and $Gini_k$, consistent with the previous literature on institutions and specialization. In addition, the magnitude of $\widehat{\beta}$ falls relative to column (1), but it remains negative and significant.

Another possible source of omitted variable bias is linked to financial markets. There is a large body of literature that sustains that financial markets are instrumental to opening new sectors and increasing the variety of industries in the economy (e.g., Greenwood and Jovanovic, 1990; Saint-Paul, 1992; Acemoglu and Zilibotti, 1997). We could then expect that countries with more developed financial markets would also be better able to specialize in industries that require a wider input base. To deal with this concern, in column (3) we interact the Gini coefficients with an indicator of financial development: the log ratio of private credit to GDP. (This indicator is taken from the World Bank Indicators database, and averaged during years 2005-2014.) The effect of financial development interacted with $Gini_k$ is significant, and it carries a sign consistent with the past literature on growth and diversification. Yet, the estimate of β still remains negative and significant.

Column (4) adds an interaction term between $Gini_k$ and log GDP per capita. This term would control for the possibility that richer economies may be better able to produce goods with lower $Gini_k$, given that richer economies tend to exhibit a more diversified productive structure than poorer ones. As we can see, the estimate of β remains essentially unaltered.

In column (5), following the approach by Romalis (2004), we introduce two additional regressors to control for specialization driven by factor endowments: *i*) an interaction term between capital intensity of industry k and the (log) stock of physical capital per worker in country c; *ii*) an interaction term between the skill intensity of industry k and the (log) stock of human capital in country c. The measures of capital and skill intensity at the industry level are constructed from the NBER-CES Manufacturing Industry database, for year 2011.[32]

[32] Capital intensity is computed as the total stock of physical capital per worker by industry. Skill intensity is measured by the average wage by industry. (See Becker, Gray and Marvakov (2013) for details on the NBER-CES Manufacturing Industry database.) Both the measure of physical capital per worker and the index of human capital are drawn from the Penn Tables database. (The human capital index is based on the average years of

TABLE I
Export Specialization across Industries with Different Levels of Input Narrowness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Road Density x Input Narrowness	-4.084***	-2.240***	-2.382***	-2.356***	-1.852***	-1.845***	-2.439***
	(0.305)	(0.334)	(0.363)	(0.364)	(0.379)	(0.385)	(0.453)
Rule of Law x Input Narrowness		-5.414***	-2.639***	-2.435***	-3.132***	-3.045***	-2.999***
		(0.441)	(0.648)	(0.702)	(0.762)	(0.766)	(0.765)
log (Priv Credit/GDP) x Input Narrowness			-3.743***	-3.547***	-1.704*	-1.681*	-1.500*
			(0.760)	(0.822)	(0.913)	(0.915)	(0.918)
log GDP per capita x Input Narrowness				-0.417	0.905	1.306**	0.822
				(0.602)	(0.656)	(0.654)	(0.658)
Capital Intensity x log $(K/L)_c$					0.010**		0.010**
					(0.004)		(0.004)
Skill Intensity x log H_c					0.008***		0.008***
					(0.001)		(0.001)
(Pop) Density x Input Narrowness							-0.197**
							(0.101)
Road Dens x (Pop) Dens x Input Nwness							0.064**
							(0.027)
Observations	42,578	41,947	40,692	40,692	31,892	31,892	31,892
R-squared	0.765	0.764	0.764	0.764	0.794	0.793	0.794
Number of Countries	166	163	157	157	134	134	134
Number of Industries	294	294	294	294	259	259	259

Robust standard errors reported in parentheses. All regressions include country and industry fixed effects. The dependent variable is the logarithm of exports in industry k by country c in 2014. Rule of law is taken from the World Governance Indicators (World Bank) for year 2014. Private credit over GDP is taken from the World Bank Indicators, and it is averaged for years 2005-2014. GDP per capita, stock of physical capital, and the human capital index are taken from Penn Tables, all for year 2014. Measures of physical capital and skill intensity by industry are taken from the NBER-CES Manufacturing Industry Database, and corresponds to year 2011. *** p<0.01, ** p<0.05, * p<0.1

Some industries are lost from the sample when we introduce the industry factor intensity measures, since the NBER dataset contains information only for manufacturing industries. For comparability, in column (6) we display the results of the regression in column (4), but using the restricted sample. The coefficients associated to the factor intensities carry the expected sign, while the estimates of β remain negative and significant. Furthermore, the estimated coefficients are of similar magnitude in both columns.

Finally, the last column of Table I addresses the possibility of a differential effect of the road network on the pattern of specialization depending on the population density of the economy. One could expect that more densely populated countries may display also a greater concentration of activities in fewer locations. Hence, all else equal, more densely populated countries may need to resort less strongly on a vast road network than sparsely populated countries. Column (7) assesses this possibility by introducing an interaction term between population density and $Gini_k$, and a *triple* interaction term which also includes r_c. If road network length is especially important for specialization in economies that are *less* densely populated, then the triple interaction term should carry a positive estimate. As can be observed, this is indeed the case. Moreover, the estimate of β after introducing the triple interaction term

schooling from Barro & Lee (2013) and an assumed rate of return of education based on Mincer estimates.)

TABLE II
High-Income and Low-Income Subsamples

	(1)	(2)	(3)	(4)	(5)	(6)
Road Density x Input Narrowness	-2.118***	-4.935***	-1.636***	-4.495***	-1.635***	-4.477***
	(0.354)	(1.482)	(0.373)	(1.712)	(0.375)	(1.713)
Rule of Law x Input Narrowness	-0.673	-4.075**	-1.393	-2.865	-1.170	-2.810
	(1.005)	(1.739)	(1.094)	(2.025)	(1.091)	(2.029)
log (Priv Cred/GDP) x Input Narrowness	-4.882***	-2.001	-4.640***	0.758	-4.712***	0.739
	(1.086)	(1.319)	(1.193)	(1.428)	(1.190)	(1.428)
log GDP per capita x Input Narrowness	-0.418	1.723	2.002	2.087*	2.357	2.459**
	(1.541)	(1.069)	(1.638)	(1.112)	(1.650)	(1.111)
Capital Intensity x log $(K/L)_c$			0.010	0.014*		
			(0.012)	(0.009)		
Skill Intensity x log H_c			0.014***	-0.015***		
			(0.002)	(0.002)		
Observations	22,087	18,605	17,137	14,755	17,137	14,755
R-squared	0.791	0.632	0.808	0.652	0.807	0.651
Countries Sample (high/low income)	High	Low	High	Low	High	Low
Number of Countries	79	78	68	66	68	66
Number of Industries	294	294	259	259	259	259

Robust standard errors reported in parentheses. All regressions include country and industry fixed effects. The dependent variable is log ($Expo_{c,k}$) in the year 2014. The high-income sample comprises countries with GDP per capita above the sample median, and the low-income sample the countries with GDP per capita below it. The median income of the sample lies between that of Iraq ($12,095 PPP) and South Africa ($12,128 PPP) in 2014. Rule of law is taken from World Governance Indicators for year 2014. Private credit over GDP is taken from the World Bank Indicators, averaged for years 2005-2014. Physical capital and skill intensity by industry are taken from the NBER-CES Manufacturing Industry Database, and corresponds to year 2011. *** p<0.01, ** p<0.05, * p<0.1

is still negative and highly significant.

Next, Table II displays some of the regressions previously presented in Table I, but now splitting the sample of countries in two subsamples, according to whether their income is above or below the median. The odd-numbered columns show the results for the subsample of 'high-income countries', while the even-numbered columns do that for the 'low-income countries'. The results show that the effect of road density on pattern of specialization holds true both for richer and poorer countries. In addition to that, the effect seems to be consistently greater in magnitude for the subsample of economies whose income is below the median.[33]

Additional Robustness checks

Further robustness checks are provided in Appendix B. Tables A.2 and A.3 change the measure of transport density. In Table A.2, we show the results of a set of regressions substituting r_c in (25) by railway density, computed as the total railway network length of country c per

[33] This difference in magnitude could suggest the presence of some sort of decreasing marginal effect of road density, since richer economies tend to exhibit denser road networks than poorer ones (see Figure 1 later on).

square kilometer. In Table A.3, we expand our measure transport network length to include (in addition to roadways) also the total length of internal railways and waterways. All the results in Table A.2 and A.3 follow a similar pattern as those shown in Table I.

Table A.4 shows that all the previous results are also robust to: *i*) excluding very small countries (both in terms of area and population), *ii*) excluding very large countries (in terms of area), *iii*) controlling for the effect of area and population (in both cases interacted with the measure of input narrowness), *iv*) including the interaction between total GDP and input narrowness, and *v*) excluding from the sample those countries whose road networks were measured before year 2010 (which is the median year in the sample). The rationale for these additional robustness checks is the following. In the case of very small countries, on the one hand, they may find it easier to link together geographic locations while, on the other hand, they may be less able to provide sufficient opportunities for input diversity. Next, regarding very large countries in terms of area, the concern could be that some of those countries may contain very large swathes of uninhabitable land, which may end up turning our measure of road density somewhat imprecise in those cases. Controlling for the size of the country (both in terms of area and population) takes into account the possibility that larger countries may face more opportunities for input diversity, regardless of the density of their internal transport network. Similarly, including total GDP can control for the possibility that there exist minimum size requirements to open up some sectors in the economy. Finally, restricting the sample to countries whose road networks were measured after 2010 helps in harmonizing the data year on trade flows and road density, and shows that the found effects are not contaminated by countries whose data on road networks is relatively older.[34]

Lastly, Table A.5 shows the results of some of the regressions of Table I, but computing the variable 'road density' using only the total length of *paved* roads. All the main results hold through. Moreover, the estimate for the main interaction term is in absolute terms greater when using only paved roads than in the regression that relies on all types of roads. This is in fact the type of result one should expect to see if paved roads provide better transportation than unpaved roads do.

[34] For this same harmonization purpose, column (7) of Table A.4 shows also the results when using only data on road density for years 2010 and 2011. The estimates remain essentially unaffected as well in this case.

6 Endogeneity and Alternative Interpretations

The previous section presented a robust association between road density in country c and its degree of specialization in industries that rely on a wide set of inputs. While those results are certainly consistent with the main predictions of the model, they cannot be taken as hard evidence of its underlying mechanism. Two separate issues deserve further discussion and analysis. First, the correlation found in the previous regressions could as well be the result of road infrastructure responding to transport needs stemming from industry specialization (i.e., reverse causation). Second, our interpretation of a lower value of $Gini_k$ as reflecting greater need of industry k for the local transport infrastructure is debatable, as previous authors have looked at that variable as capturing a different feature: the degree of product complexity of industry k. In the next two subsections we aim to address these two points more explicitly.

6.1 Endogeneity and Reverse Causation

Our model has taken r_c as exogenously given. The length of a country's road network is however the result of investment choices in infrastructure, and hence it will respond to a host of economic variables and incentives. The exogeneity of r_c represents thus a critical assumption that warrants further discussion in case the previous empirical results are intended to be interpreted as evidence of a *causal* effect from road density to specialization.

Endogenizing countries' road networks can easily lead to a model where β in (25) can be confounding an effect from road density to specialization, together with reverse causality from the latter to the former. For example, suppose that for some reason the final good production functions differ between H and F, and that H turns out to be relatively more productive in the final sectors that tend to rely on a broader set of intermediate inputs. If countries were able to invest in expanding their road networks, we could well expect r_H to be larger than r_F simply because the incentives to do so are greater in H than in F. From an empirical viewpoint, this reasoning means that β could end up capturing (at least partially) an effect going from patterns of specialization to road density.

The rest of this subsection provides some further support for the notion that the density of the internal transport network is instrumental to specialization in industries with wider input bases. First, we show that a similar correlation to that one found in Section 5.3 arises when the density of the transport network is measured by the density of internal waterways. Next, we show that the results in Section 5.3 remain true when we instrument the density of a country's road network with topographical measures of terrain roughness.

6.1.1 Patterns of Specialization and Waterways Density

This first part intends to provide some further evidence consistent with the main mechanism of the model, by relying on a measure of countries' transport network that is *less* sensitive to reverse causality concerns than r_c is. We measure now the internal transport network of an economy by the density of their waterways network. We draw the data on waterways from the CIA World Factbook, and define waterways density as waterways length per square km.[35] Arguably, while countries can still expand their waterways by investing in creating canals or improving the navigability of some rivers and bodies of water, the scope for this is far more limited than in the case of roads.

One additional aspect we investigate here is the possibility that waterways impact specialization *heterogeneously* at different stages of development. For various reasons, richer economies tend to have much denser road networks than poorer ones. In particular, poorer economies may find it harder to undertake the necessary investment to build a sufficiently developed road infrastructure. On the other hand, while the presence of waterways may have influenced patterns of development before railroads and roads became more widespread worldwide, waterways are no longer a mode of transportation that seems to be associated with economies' current level of development. In fact, a quick look at simple cross-country correlations in Figure 1 shows that income per head and road density display a clear positive correlation, while the association between income per head and waterways density is rather weak.[36]

Table III displays the results of a regression equation analogous to (25), but where r_c is replaced by a measure of *waterways density*. The table shows the results for three different countries sample: entire sample, high-income countries, and low-income countries.

The regressions based on the whole set of countries tend to yield an estimate that is negative. However, this aggregate result masks important heterogeneities in the impact of waterways density on specialization for richer versus poorer economies. Column (2) shows that waterways density carries no impact at all in the subsample of above-median income economies. By contrast, column (3) exhibits a negative and highly significant coefficient. This suggests that, in the case of poorer economies, those that enjoy a denser network of waterways tend to export relatively more in industries that require a wider intermediate input base.

[35] The CIA World Factbook measures waterways as the total length of navigable rivers, canals and other inland bodies of water.

[36] One possible interpretation of Figure 1 is that, as economies grow richer, roadways tend to gradually overshadow waterways for internal transportation. From this perspective, we could then expect waterways to represent an important determinant of specialization in poorer economies, but losing preeminence in richer economies where roadways can more easily make up for an insufficiently dense internal waterway network.

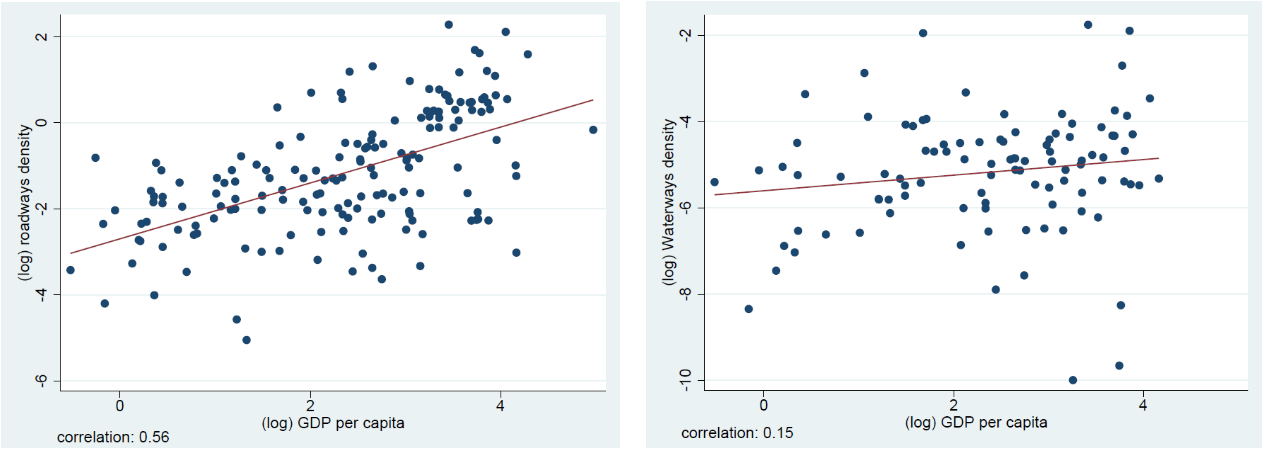

Figure 1: Roadways and waterways density against GDP per head

Columns (4)-(6) re-run the regressions in columns (1)-(3) but also including the original interaction term $r_c \times Gini_k$. The results for the impact of waterways density on specialization in (5) and (6) follow a very similar pattern as those in (2) and (3). Moreover, the regressions also show that the coefficient for road density remains negative and significant.[37]

Finally, Table A.6 in Appendix B shows the results a of set of regressions that include a triple interaction term between r_c, $Gini_k$ and log income, rather than splitting the sample of countries according to income. All the results remain qualitatively consistent with those in Table III.

6.1.2 Instrumental Variables: Terrain Roughness and Roadway Density

This second part intends to address more directly the concern of reverse causation from industry specialization to road density. To do so we instrument r_c with measures of terrain roughness in country c. The idea is drawn from Ramcharan (2009), who shows that countries with rougher terrain surface tend to exhibit less dense road networks.[38] In the context of our paper, if the roughness of the terrain affects the density of the internal road network, but it does not exert a systematic impact on specialization across industries with varying degrees of input narrowness

[37] Note that the results in columns (5) and (6) of Table III are not directly comparable to those in columns (3) and (4) of Table II due to the loss of some countries in the samples of Table III.

[38] Ramcharan (2009) studies the spatial concentration of economic activities within countries, and how this is affected by their surface topography. The author argues that countries with rougher topography tend to display stronger spatial concentration of economic activity, and that this is partly of explained by the poorer land transportation associated with rougher terrain.

TABLE III

Waterways Density as Measure of Transport Network

	(1)	(2)	(3)	(4)	(5)	(6)
Road Density x Input Narrowness	-0.262*	0.016	-1.120***	-0.137	0.194	-1.100***
	(0.148)	(0.140)	(0.347)	(0.170)	(0.172)	(0.347)
Rule of Law x Input Narrowness	-4.205***	-4.727***	-7.389***	-3.576***	-4.257***	-4.533*
	(0.838)	(1.546)	(2.511)	(0.902)	(1.561)	(2.811)
log (Priv Cred/GDP) x Input Narrowness	-1.724*	-2.403*	0.484	-1.857*	-2.505*	0.624
	(1.077)	(1.377)	(1.664)	(1.085)	(1.382)	(1.663)
log GDP per capita x Input Narrowness	0.457	5.220*	0.283	0.620	5.752**	0.830
	(0.867)	(2.944)	(1.256)	(0.874)	(2.962)	(1.268)
Capital Intensity x log $(K/L)_c$	0.007	-0.014	0.021*	0.008	-0.013	0.021*
	(0.006)	(0.014)	(0.012)	(0.006)	(0.014)	(0.012)
Skill Intensity x log H_c	0.007***	0.020***	-0.018***	0.007***	0.020***	-0.018***
	(0.001)	(0.003)	(0.002)	(0.001)	(0.003)	(0.002)
Road Density x Input Narrowness				-1.381**	-1.426**	-7.249***
				(0.711)	(0.703)	(2.386)
Observations	22,357	11,750	10,607	22,357	11,750	10,607
R-squared	0.811	0.798	0.712	0.811	0.798	0.712
Countries Sample	All	High	Low	All	High	Low
Number of Countries	93	46	47	93	46	47
Number of Industries	259	259	259	259	259	259

Robust standard errors in parentheses. All regressions include country and industry fixed effects. The dependent variable is log($Expo_{c,k}$) in year 2014. Waterways is taken from the CIA World Factbook. It comprises total length of navigable rivers, canals and other inland water bodies. Waterway density equals internal waterways per square km. Columns (2) and (5) include only those countries with above median income in the sample. Columns (3) and (6) include only those countries whose income lies below the median income in the sample. ***p<0.01, **p<0.05, *p<0.1

via other alternative channels, then it can serve as a valid instrument for r_c.[39]

In Table IV we show the results of the two-stage least square (2SLS) regressions using three alternative measures of terrain roughness: *i*) the difference between the maximum and minimum land elevation in country c, taken from the CIA Factbook; *ii*) the standard deviation of elevation of country c measured at 30" degree grids (approx. 1km cells), taken from Ramcharan (2009); *iii*) the percentage of mountainous terrain, taken from Fearon and Laitin (2003). Table A.7 in

[39] Notice that a violation of the exclusion restriction in this context requires more than simply terrain roughness having an impact on industry specialization. For the exclusion restriction to be violated, it must be the case that terrain roughness affects specialization across industries in a way that is also correlated with their degree of input narrowness (besides the effect mediated by the impact of terrain roughness on road density). For example, the exclusion restriction may be threatened if economies with rougher terrain tend to also display more heterogeneous climatic conditions and land configurations, and this allows them to enjoy a more diverse productive structure. Conversely, it may be that rougher terrain reduces the share of inhabitable land, curtailing productive heterogeneity, and thereby possibly leading to a violation of the exclusion restriction via a negative impact of roughness on specialization in industries with wide input breadth. While we cannot test the validity of the exclusion restriction, the results in Table A.8 in Appendix B (see also the discussion therein) are in principle encouraging about how concerned we should remain about a the possibility of a direct impact of terrain roughness, besides that one mediated by its effect on road density.

Appendix B shows that all three measures of terrain roughness are negatively correlated with road density, even after controlling for several other country-level variables.[40]

Columns (1) and (2) display the results of the 2SLS regressions based on the difference between the maximum and minimum land elevation in c as instrument for r_c. Arguably, this variable may be seen as a relatively crude measure for terrain roughness. However, it has the upside of being available for the exact same samples as those in Section 5.3. As a result, columns (1) and (2) of Panel A can be directly compared to their respective OLS counterparts in columns (4) and (5) in Table I. Next, columns (4) and (6) display the estimation results of the 2SLS regressions in which the instrument for road density is based on the standard deviation of land elevation. Since for some of the countries in the original sample this information is missing, columns (3) and (5) show their respective OLS estimates for the corresponding sample. Finally, columns (8) and (10) report the 2SLS results when using the percentage of mountainous terrain as instrumental variable. Again, in the sake of comparability, columns (7) and (9) report the OLS estimates for the corresponding country samples.

The main message to draw from Table IV is that the 2SLS regressions consistently yield a negative and significant estimate for our coefficient of interest. These results reinforce the support for the hypothesis that the density of the internal transport network is an important determinant of specialization in industries with wide input bases, by exploiting the variation in the internal road network across countries predicted by their degrees of terrain roughness.

One additional point to note is that the 2SLS estimates for β tend to be consistently greater in absolute magnitude than their OLS counterparts. This would in principle run against the direction of the bias that would stem from the reverse causality concern discussed in the second paragraph of Section 6.1 (i.e., the notion that economies specializing in industries that rely on a wide variety of inputs may tend to invest more in transport infrastructure). One possible reason behind these results is that the instrument may also be alleviating some degree of measurement error in our indicator of road density. In that respect, recall that r_c is computed using the total length of roads by country. This disregards the fact that different roads may differ substantially in terms of quality and width, and it is also summing up together paved and unpaved roads. Furthermore, the total length or the road network is not taking into account the possibility of a

[40] All the measures of terrain roughness used here aim at capturing large-scale terrain irregularities. In a sense, these seem to be the types of irregularities that can most severely hinder internal transport networks. Other papers, e.g. Nunn and Puga (2012), have resorted to the methodology developed by Riley et al. (1999) so as to measure small-scale terrain irregularities. While small-scale terrain roughness measures seem more appropriate for capturing the presence of small geographic formations that may provide natural sources of protection to certain groups of people, they may not represent the main source of obstruction to dense transport networks.

very inefficient lay out of the network. All these issues could end up reducing the precision with which r_c captures the notion that a denser road network allows cheaper internal transportation of inputs. Therefore, when instrumenting r_c, we may not only be dealing with problems of endogeneity, but also with the fact that in some cases our measure of road density may be quite imprecisely gauging the efficiency of the internal road network.

6.2 Alternative Interpretations of the Input Breadth Measures

The analysis in Section 5 was based on the notion that the degree of input breadth of industry k can serve as an indicator for how reliant this industry is on the internal transport network. The need to source a large variety of inputs can certainly make a particular sector heavily dependent on efficient transportation; however, it can also mean that the sector is highly sensitive to sound contract enforcement. Indeed, Blanchard and Kremer (1997) and Levchenko (2007) have previously used input-output data to compute diversification indices for intermediate input purchases across industries, and use them to proxy the degree of complexity of sectors: sectors with more diversified (i.e., less concentrated) input bases are considered to be more complex. In their analysis, more complex sectors require better contract enforcement to work efficiently. From this viewpoint, countries with better functioning institutions should exhibit a comparative advantage in industries with broad intermediate input bases. Levchenko (2007) shows that this is indeed the case for US imports: the US imports a higher share of goods with greater diversity of intermediate inputs from countries with better rule of law.

The previous regressions have conditioned on the interaction between rule of law in country c and the Gini coefficient for intermediate inputs in industry k. The estimate of β remained consistently negative and significant, regardless of the introduction of this additional control. In that regard, our results seem to suggest that *both* institutions and local transport networks are instrumental and complementary to the growth of industries with wide input bases. This section will attempt to further strengthen this argument

Countries with better institutions are in general richer, and also exhibit a denser transportation infrastructure network. If $Gini_k \times r_c$ in (25) were *not* capturing any type of impact related to how transport-intensive industry k is, but *only* the effect of rule of law in country c through its correlation with r_c, then the correlation found in Table I should arise more prominently for industries that are relatively more dependent on judicial quality. The regressions reported in Table V show this is actually not the case in the data.

Columns (1) and (2) in Table V show the results of the simple correlation reported initially in column (1) of Table I, after splitting the set of industries in two subsamples: low contract

TABLE IV

Two-Stage Least Squares Regressions using Terrain Roughness as Instrument for Roadway Density

	PANEL A: SECOND-STAGE RESULTS (AND OLS COMPARISONS)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	2SLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
Road Density x Gini	-5.006***	-3.188*	-3.364***	-6.199**	-2.204***	-6.166*	-3.347***	-5.638*	-2.000***	-5.645*
	(1.513)	(1.670)	(0.509)	(2.800)	(0.544)	(3.573)	(0.506)	(3.258)	(0.515)	(3.355)
Rule of Law x Gini	-1.260	-2.469**	-2.285***	-1.306	-2.886***	-1.264	-2.082***	-1.108	-3.147***	-1.543
	(0.943)	(1.100)	(0.720)	(1.187)	(0.786)	(1.635)	(0.737)	(1.564)	(0.790)	(1.661)
Fin Dev x Gini	-3.340***	-1.678*	-4.063***	-3.978***	-2.605***	-2.849***	-3.975***	-4.060***	-1.507	-1.719*
	(0.838)	(0.917)	(0.878)	(0.890)	(0.956)	(0.963)	(0.870)	(0.874)	(0.938)	(0.948)
log GDP pc x Gini	-0.088	1.077*	0.161	0.620	1.170*	1.857**	0.056	0.437	0.926	1.562*
	(0.623)	(0.677)	(0.667)	(0.781)	(0.714)	(0.908)	(0.641)	(0.820)	(0.676)	(0.868)
K Intens x log $(K/L)_c$		0.010**			0.009**	0.009**			0.008*	0.008*
		(0.004)			(0.0045)	(0.0045)			(0.004)	(0.004)
Skill Intens x log H_c		0.008***			0.007***	0.007***			0.008***	0.008***
		(0.001)			(0.001)	(0.001)			(0.001)	(0.001)
Observations	40,692	31,892	35,988	35,988	29,229	29,229	36,544	36,544	30,067	30,067
R-squared	0.763	0.794	0.764	0.764	0.794	0.793	0.757	0.757	0.795	0.795
Number of Countries	157	134	135	135	121	121	138	138	126	126
Number of Industries	294	259	294	294	259	259	294	294	259	259
	PANEL B: FIRST-STAGE RESULTS (Dep. Variable: Road Density x Input Narrowness)									
Terrain Roughness x Gini	-0.013***	-0.013***		-0.452***		-0.392***		-0.007***		-0.007***
	(0.0004)	(0.001)		(0.017)		(0.019)		(0.0002)		(0.0003)
Rule of Law x Gini	0.344***	0.371***		0.290***		0.350***		0.398***		0.413***
	(0.011)	(0.012)		(0.011)		(0.012)		(0.011)		(0.012)
Fin Dev x Gini	0.109***	0.059***		0.066***		-0.029***		-0.004		-0.032***
	(0.011)	(0.013)		(0.011)		(0.011)		(0.009)		(0.010)
log GDP pc x Gini	0.144***	0.164***		0.165***		0.184***		0.152***		0.163***
	(0.007)	(0.010)		(0.009)		(0.010)		(0.008)		(0.009)
K Intens x log $(K/L)_c$		0.005				0.010				0.008
		(0.008)				(0.007)				(0.007)
Skill Intens x log H_c		-0.010				-0.017				-0.010
		(0.015)				(0.013)				(0.014)
F-Stat: P-Value	0.00	0.00		0.00		0.00		0.00		0.00

Robust standard errors in parentheses. All regressions include country and industry fixed effects. The dependent variable in Panel A is the log($Expo_{c,k}$) in year 2014. 'Terrain Roughness' is measured as follows: *i*) in columns (1) and (2) by the diffference between the maximum and the minimum elevation in country *c* (source: CIA World Factbook); *ii*) in columns (4) and (6) by the std. deviation of elevation in country c computed at the 30" degree resolution (source: Ramcharan, 2009); *iii*) in columns (8) and (10) by the percentage of mountainous in terrain in country c (source: Fearon and Laitin, 2003). *** p<0.01, ** p<0.05, * p<0.1

TABLE V

Regressions on Industry Subsamples: Effects at Different Leves of Contract Intensity

	(1)	(2)	(3)
Road Density x Input Narrowness	-3.443***	-1.929***	-2.070***
	(0.652)	(0.483)	(0.440)
Rule of Law x Contract Intensity			0.441***
			(0.057)
Rule of Law x Input Narrowness			-1.974**
			(0.926)
log (Priv Credit/GDP) x Input Narrowness			-0.533
			(1.082)
log GDP x Input Narrowness			0.897
			(0.772)
Capital Intensity x log $(K/L)_c$			0.007***
			(0.002)
Skill Intensity x log H_c			0.008*
			(0.005)
Observations	13,179	14,231	20,823
R-squared	0.704	0.820	0.790
Contract Intensity	low	high	all
Number of Countries	166	166	134
Number of Industries	91	91	163

Robust standard errors in parentheses. All regressions include country and industry fixed effects. The dependent variable is log($Expo_{c,k}$) in year 2014. Contract intensity by industry measures are taken from Nunn (2007). The original measures are coded according to the NAICS 1997 classification and matched to the BEA codes. Nunn (2007) measures contract intensity in k as the proportion of inputs of k classified as differentiated by Rauch (1999). Rule of law is from the World Governance Indicators (year 2014). Private credit over GDP is from the World Bank Indicators (averaged for 2005-14). GDP per capita, total GDP, the stock of physical capital (per capita), and the human capital index (H_c) are all taken from the Penn Tables (year 2014). Capital and skill intensity by industry are from the NBER-CES Manuf. Industry Database (for year 2011). *** p<0.01, ** p<0.05, * p<0.1

intensity vs. high contract intensity. To do so, we take the measure of contract intensity by industries from Nunn (2007), and split the sample of industries according to whether they rank below or above the median value of contract intensity.[41] If the $Gini_k$ were simply proxying for how sensitive to efficient contract enforcement industry k is, then the estimate in column (1) should turn out to be significantly milder than that one in column (2). The regressions show, however, that the negative correlation is significant in both subsamples, and moreover the magnitude of the estimates are not significantly different from one another.

Finally, to complement our previous results, column (3) displays the outcomes of a regression analogous to that one in column (5) of Table I, but including the interaction term between rule of law in country c and contract intensity of industry k (namely, 'Rule of Law $\times$ Contract Intensity'). Consistently with the previous results in the literature, the regression shows that

[41] Nunn (2007) reports contract intensity measures for 222 industries coded according to NAICS 1997. We lose some of the original industries in Table I when matching the NAICS 1997 codes to those of BEA 2002.

countries with better institutions exhibit a comparative advantage in the industries with higher contract intensity. Moreover, the regression results are in line with those in column (5) of Table I, and support the prediction that countries with denser road networks export relatively more in sectors that require a broader variety of intermediate inputs.

7 Concluding Remarks

We proposed a simple trade model where the density of the internal transport network represents a key factor in shaping comparative advantage and specialization. The underlying mechanism rests on the idea that shipping intermediate inputs across spatially diffuse locations is costly. As a consequence, industries that rely on a wider set of intermediate inputs become heavier users of the internal transport network. The model shows that countries with denser internal transport infrastructures exhibit a comparative advantage in industries that rely on a broader variety of intermediate goods.

Drawing on intermediate goods transactions from the US input-output matrix to measure industries' input breadth, we have also shown that the patterns of specialization predicted by the model are broadly consistent with the trade flows observed in the data. Several caveats apply nonetheless to the empirical evidence. First, patterns of specialization could be influencing investment in road infrastructure, and thus be behind the correlation found in the data. In that respect, the fact that the same correlation appears when substituting roadway density by waterway density seems reassuring, as waterways are harder to expand in response to increased transport needs. Furthermore, our empirical results also remain in line with the predictions of the model when we instrument the density of countries' road networks with indicators of roughness of the terrain. Second, the measure of input breadth by industries could alternatively be capturing a stronger need for contract enforcement when the input base is wider. We showed however that the correlation predicted by our model is still present when the confounding effect of institutional quality (by country) and judiciary intensity (by industry) is also taken into account. Finally, our measure of road density is a relatively imprecise way to capture the efficiency in connectedness of different locations. Road networks not only differ in length, but also in terms of width, quality, etc. In addition, our measure of road density fails to account for inefficiencies in the layout of internal roads. It would be certainly desirable to use of a more detailed and refined measure of internal road networks by countries. Yet, it is hard envision a clear reason why the above sources of measurement error could end up systematically biasing the previous empirical results in the direction predicted by the model.

As a final remark, the paper has abstracted from analyzing the possibility of a heterogenous impact of transport networks on the internal geographic configuration of industries with different degrees of input breadth. Access to international markets could indeed create a force leading to an agglomeration of industries with wider input bases along "coastal" regions (i.e., regions closer to international markets). Furthermore, this force pushing certain industries towards coastal regions could be especially important in countries lacking a sufficiently dense internal transport network. We leave this question open for future research.

Appendix A: Proofs

Proof of Lemma 1. The proof follows from noting that, since $T > 0$ and $d(r)^{-\theta} < 1$, (7) and (8) imply that $c_{j,A} < c_{j,B}$ when $0 \leq \alpha_j < \frac{1}{2}$, $c_{j,A} > c_{j,B}$ when $\frac{1}{2} < \alpha_j \leq 1$, and $c_{j,A} = c_{j,B}$ when $\alpha_j = \frac{1}{2}$. ■

Proof of Lemma 2. The first part of the lemma follows immediately from (9), bearing in mind that Assumption 2 means $d(r_1)^{-\theta} < d(r_2)^{-\theta}$.

To prove the second part, apply logs on the ratio $c_j(r_1)/c_j(r_2)$, which using the expressions in (9) leads to:

$$\ln\left(\frac{c_j(r_1)}{c_j(r_2)}\right) = \begin{cases} \frac{1}{\theta}\left[(1-\alpha_j)\ln\left(\frac{(1+T)+d(r_2)^{-\theta}}{(1+T)+d(r_1)^{-\theta}}\right) + \alpha_j \ln\left(\frac{1+(1+T)d(r_2)^{-\theta}}{1+(1+T)d(r_1)^{-\theta}}\right)\right] & \text{if } \alpha_j \in \left[0, \frac{1}{2}\right], \\ \frac{1}{\theta}\left[(1-\alpha_j)\ln\left(\frac{1+(1+T)d(r_2)^{-\theta}}{1+(1+T)d(r_1)^{-\theta}}\right) + \alpha_j \ln\left(\frac{(1+T)+d(r_2)^{-\theta}}{(1+T)+d(r_1)^{-\theta}}\right)\right] & \text{if } \alpha_j \in \left[\frac{1}{2}, 1\right]. \end{cases} \tag{26}$$

Consider first the case when $\alpha_j \leq 0.5$. Differentiating the relevant expression in (26) with respect to α_j, it follows that

$$\frac{\partial\left(\ln(c_j(r_1)/c_j(r_2)\right)}{\partial \alpha_j} > 0 \quad \Leftrightarrow \quad \frac{1+(1+T)\,d(r_2)^{-\theta}}{(1+T)+d(r_2)^{-\theta}} > \frac{1+(1+T)\,d(r_1)^{-\theta}}{(1+T)+d(r_1)^{-\theta}},$$

which holds holds true since $T > 0$ and $d(r_2)^{-\theta} > d(r_1)^{-\theta}$. Similarly, considering the case when $\alpha_j \geq 0.5$, and differentiating the relevant part of (26) with respect to α_j, it follows now that

$$\frac{\partial\left(\ln(c_j(r_1)/c_j(r_2)\right)}{\partial \alpha_j} < 0 \quad \Leftrightarrow \quad \frac{1+(1+T)\,d(r_2)^{-\theta}}{(1+T)+d(r_2)^{-\theta}} > \frac{1+(1+T)\,d(r_1)^{-\theta}}{(1+T)+d(r_1)^{-\theta}}.$$

Therefore, the ratio $c_j(r_1)/c_j(r_2)$ is strictly increasing in α_j whenever $\alpha_j < 0.5$, while it is strictly decreasing in α_j whenever $\alpha_j > 0.5$. In addition, by continuity, the ratio $c_j(r_1)/c_j(r_2)$ must reach a global maximum at $\alpha_j = 0.5$. ■

Proof of Proposition 1. Suppose firstly that $\omega \leq 1$. Then, given that Assumption 3 means that $\lambda > 1$, it follows that the expressions in (12) would hold true for all $\alpha_j \in [0, 1]$. But that would then mean that F is running a trade deficit with H, and hence $\omega \leq 1$ cannot possibly hold in equilibrium. Next, to prove that $\partial\omega^*/\partial\lambda > 0$, note from (13) that $\underline{\alpha}'_\lambda(\cdot) < 0$ and $\underline{\alpha}'_\omega(\cdot) > 0$. Therefore, differentiating condition (14) with respect to ω^* and λ, the result $\partial\omega^*/\partial\lambda > 0$ obtains. Thirdly, the result that $\lim_{\lambda\to 1}\omega^*(\lambda) = 1$ follows from (12). Finally, notice that since F must be exporting some goods to H, in equilibrium, it must be that $\underline{\alpha}(\omega^*) > 0$, which requires that

$$\omega^* > \left[\frac{(1+T)d_H^\theta + 1}{(1+T)d_H^\theta + \lambda^{-\theta}}\right]^{\frac{1}{\theta}}.$$

On the other hand, in equilibrium, H must be exporting some goods to F, which entails that $\underline{\alpha}(\omega^*) < \frac{1}{2}$, from where it follows that

$$\omega^* < \left[\frac{(1+T) + d_H^\theta}{(1+T)\lambda^{-\theta} + d_H^\theta} \times \frac{(1+T)d_H^\theta + 1}{(1+T)d_H^\theta + \lambda^{-\theta}} \right]^{\frac{1}{2\theta}}.$$

■

Proof of Proposition 2. First of all, note that the expression in (12) corresponding to the cases when $\alpha_j < 0.5$ is strictly increasing in α_j, while the one corresponding to the cases when $\alpha_j > 0.5$ is strictly decreasing in α_j. In addition, note that both expressions in (12) are equal to one another in the special case when $\alpha_j = 0.5$. As a result, given the equilibrium relative wage ω^* satisfying the properties stated in Proposition 1, there must exist $\underline{\alpha} \in \left(0, \frac{1}{2}\right)$ and $\overline{\alpha} \in \left(\frac{1}{2}, 1\right)$, such that (12) holds true for all $\alpha_j \in (\underline{\alpha}, \overline{\alpha})$ and fails to hold true when $\alpha_j \notin (\underline{\alpha}, \overline{\alpha})$. Using again the expressions in (12), the equations for $\underline{\alpha}$ and $\overline{\alpha}$ obtain when solving it for the case of equality, and they are given by (13). Finally, the fact that $\overline{\alpha}(\omega) = 1 - \underline{\alpha}(\omega)$ follows straightforwardly from (13). ■

Proof of Lemma 3. In the presence of the iceberg trade cost for final goods $\tau > 1$, H will export good j to F when

$$\omega < \begin{cases} \frac{1}{\tau} \left[\frac{(1+T)d_H^\theta + 1}{(1+T)d_H^\theta + \lambda^{-\theta}} \right]^{(1-\alpha_j)/\theta} \left[\frac{(1+T) + d_H^\theta}{(1+T)\lambda^{-\theta} + d_H^\theta} \right]^{\alpha_j/\theta} & \text{if } \alpha_j \leq 0.5, \\ \frac{1}{\tau} \left[\frac{(1+T) + d_H^\theta}{(1+T)\lambda^{-\theta} + d_H^\theta} \right]^{(1-\alpha_j)/\theta} \left[\frac{(1+T)d_H^\theta + 1}{(1+T)d_H^\theta + \lambda^{-\theta}} \right]^{\alpha_j/\theta} & \text{if } \alpha_j \geq 0.5. \end{cases} \tag{27}$$

On the other hand, H will import good j to F when

$$\omega > \begin{cases} \tau \left[\frac{(1+T)d_H^\theta + 1}{(1+T)d_H^\theta + \lambda^{-\theta}} \right]^{(1-\alpha_j)/\theta} \left[\frac{(1+T) + d_H^\theta}{(1+T)\lambda^{-\theta} + d_H^\theta} \right]^{\alpha_j/\theta} & \text{if } \alpha_j \leq 0.5, \\ \tau \left[\frac{(1+T) + d_H^\theta}{(1+T)\lambda^{-\theta} + d_H^\theta} \right]^{(1-\alpha_j)/\theta} \left[\frac{(1+T)d_H^\theta + 1}{(1+T)d_H^\theta + \lambda^{-\theta}} \right]^{\alpha_j/\theta} & \text{if } \alpha_j \geq 0.5. \end{cases} \tag{28}$$

From (27), and defining

$$\underline{\alpha}_H \equiv \frac{\theta \left[\ln(\omega) + \ln(\tau)\right] - \ln \left[\frac{(1+T)d_H^\theta + 1}{(1+T)d_H^\theta + \lambda^{-\theta}} \right]}{\ln \left[\frac{(1+T) + d_H^\theta}{(1+T)\lambda^{-\theta} + d_H^\theta} \times \frac{(1+T)d_H^\theta + \lambda^{-\theta}}{(1+T)d_H^\theta + 1} \right]}, \tag{29}$$

it follows that, given the value of ω, and provided $\underline{\alpha}_H < \frac{1}{2}$, H will export to F the goods with $\alpha_j \in [\underline{\alpha}_H, 1 - \underline{\alpha}_H]$.

On the other hand, from (28), and defining

$$\underline{\alpha}_F \equiv \frac{\theta\left[\ln(\omega)-\ln(\tau)\right]-\ln\left[\frac{(1+T)d_H^\theta+1}{(1+T)d_H^\theta+\lambda^{-\theta}}\right]}{\ln\left[\frac{(1+T)+d_H^\theta}{(1+T)\lambda^{-\theta}+d_H^\theta}\times\frac{(1+T)d_H^\theta+\lambda^{-\theta}}{(1+T)d_H^\theta+1}\right]}, \tag{30}$$

it follows that, given the value of ω, and provided $\underline{\alpha}_F > 0$, F will export to H the goods with $\alpha_j \in [0, \underline{\alpha}_F]$ and $\alpha_j \in [1-\underline{\alpha}_F, 1]$.

Now, from the expression in (30), we can observe that $\underline{\alpha}_F > 0$ if and only if

$$\omega > \underline{\omega} \equiv \tau\left[\frac{(1+T)d_H^\theta+1}{(1+T)d_H^\theta+\lambda^{-\theta}}\right]^{\frac{1}{\theta}}. \tag{31}$$

Next, from (29), we can observe that $\underline{\alpha}_H < \frac{1}{2}$ if and only if

$$\omega < \overline{\omega} \equiv \frac{1}{\tau}\left[\frac{(1+T)d_H^\theta+1}{(1+T)d_H^\theta+\lambda^{-\theta}}\times\frac{(1+T)+d_H^\theta}{(1+T)\lambda^{-\theta}+d_H^\theta}\right]^{\frac{1}{2\theta}}. \tag{32}$$

As a result, an equilibrium with positive international trade can arise if and only if, $\underline{\omega} < \overline{\omega}$, which using (31) and (32) leads to $\tau < \widehat{\tau}$, given by (15). ■

Proof of Proposition 3. Notice that when $\tau < \widehat{\tau}$ holds true, if $\omega \leq \underline{\omega}$, F will not be able to export any good to H, but exports by H to F will actually be positive. On the other hand, when $\tau < \widehat{\tau}$ holds true, if $\omega \geq \overline{\omega}$, then H will not be able to export any good to F, but exports by F to H will actually be positive. As a result, when $\tau < \widehat{\tau}$ holds true, in equilibrium, it must be the case that $\underline{\omega} < \omega^* < \overline{\omega}$. Then, from the Proof of Lemma 3, it follows that $0 < \underline{\alpha}_F < \underline{\alpha}_H < \frac{1}{2}$, and that H will export to F all j whose $\alpha_j \in (\underline{\alpha}_H, 1-\underline{\alpha}_H)$, while F will import to F all j whose $0 \leq \alpha_j < \underline{\alpha}_F$ and $1-\underline{\alpha}_F < \alpha_j \leq 1$. ■

Proof of Proposition 4. To prove the first two results, note that using (20), we can write:

$$\pi_H(j) = \frac{\tau^{-\vartheta}}{\tau^{-\vartheta}+\left(\chi_F(j)/\chi_H(j)\right)^{-\vartheta}} \quad \text{and} \quad \pi_F(j) = \frac{\tau^{-\vartheta}}{\tau^{-\vartheta}+\left(\chi_H(j)/\chi_F(j)\right)^{-\vartheta}}. \tag{33}$$

Therefore, using (33) together with (19), the results concerning $\partial\pi_H/\partial\alpha_j$ and $\partial\pi_F/\partial\alpha_j$ below and above $\alpha_j = \frac{1}{2}$ immediately obtain.

To prove the last result, note that the trade balance equilibrium requires

$$\int_0^1 \pi_H(j)\,dj = \omega^* \int_0^1 \pi_F(j)\,dj. \tag{34}$$

Next, given that in equilibrium $\omega^* > 1$, from (34) it follows that there must necessarily exist a positive mass of j for which $\pi_H(j) > \pi_F(j)$. As a result, given that the ratio $\pi_H(j)/\pi_F(j)$ reaches a its highest value at $\alpha_j = \frac{1}{2}$, it must be that $\pi_H(\alpha_j = \frac{1}{2}) > \pi_F(\alpha_j = \frac{1}{2})$. ■

Proof of Result in Section 4.3.2. Consider first the case in which $\alpha_j \leq \frac{1}{2}$. Using (21) and (22), we can write $\ln\left(\chi_{j,H}/\chi_{j,F}\right) = \theta^{-1}\left[\ln(\Theta) - \alpha_j \ln(\Theta \times \Psi)\right]$, where

$$\Theta \equiv \frac{\left[(1+T) + (\lambda d_H)^{-\theta}\right]\omega^{\theta} + (2+T)\,\iota^{-\theta}}{\left[(1+T) + d_H^{-\theta}\right] + (2+T)\,\iota^{-\theta}\omega^{\theta}} \tag{35}$$

and

$$\Psi \equiv \frac{\left[(1+T)\,d_H^{-\theta} + 1\right] + (2+T)\,\iota^{-\theta}\omega^{\theta}}{\left[(1+T)\,(\lambda d_H)^{-\theta} + 1\right]\omega^{\theta} + (2+T)\,\iota^{-\theta}}. \tag{36}$$

Let us now define $\Lambda \equiv \Theta \times \Psi$. Thus, we can say that $\partial(\chi_{j,H}/\chi_{j,F})/\partial\alpha_j < 0 \Longleftrightarrow \Lambda > 1$. Notice now that $\lim_{\lambda\to 1}(\Lambda) > 1$, since in equilibrium $\omega > 1$. Furthermore, differentiating Λ with respect to λ, we can observe that $\partial\Lambda/\partial\lambda > 0$. Therefore, it follows that $\Lambda > 1$ for all $\lambda > 1$, proving that $\partial(\chi_{j,H}/\chi_{j,F})/\partial\alpha_j < 0$ for all $\alpha_j \leq \frac{1}{2}$.

Consider now the case in which $\alpha_j \geq \frac{1}{2}$. Using (21) and (22), we can write $\ln\left(\chi_{j,H}/\chi_{j,F}\right) = \theta^{-1}\left[\ln\left(\Psi^{-1}\right) + \alpha_j \ln(\Theta \times \Psi)\right]$, where Θ is given by (35) and Ψ by (36). In this case, the proof follows then from the fact that $\partial(\chi_{j,H}/\chi_{j,F})/\partial\alpha_j > 0 \Longleftrightarrow \Lambda > 1$. ■

Appendix B: Additional Empirical Results

B.1 Complementary Results for Section 5.3

Table A.1 shows the results of a set of regressions that include only the interaction term between road density and a number of measures of input narrowness. Column (1) of Table A.1 reproduces the same regression as in column (1) of Table I in the main text where input narrowness is measured by the Gini coefficient. Next, columns (2) - (4) of Table A.1 display the results of this simple correlation when input narrowness is measured by three alternative measures: the Herfindahl index, the coefficient of variation, and the log-variance of industry k's intermediates expenditure shares. As it can be observed, the estimate of β is negative and highly significant under all these alternative measures as well.[42]

TABLE A.1

Export Specialization across Industries with Different Levels of Input Narrowness

	(1)	(2)	(3)	(4)
Road Density x Input Narrowness	-4.084***	-0.710***	-0.034***	-0.122***
	(0.305)	(0.141)	(0.004)	(0.012)
Observations	42,578	42,578	42,578	42,578
R-squared	0.765	0.764	0.764	0.764
Country FE	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes
Number of Countries	166	166	166	166
Number of Industries	294	294	294	294
Narrowness Measure	Gini	Herf	Coef Var	Log Var

Robust standard errors reported in parentheses. The dependent variable is the log of total exports in industry k by country c in year 2014.
*** p<0.01, ** p<0.05, * p<0.1

Table A.2 and A.3 show the results of a set of regressions, analogous to some of those previously presented in Section 5.3, but where the measure of the density of the transport network is either changed or expanded. In Table A.2, road density is replaced by railway density, as our main measure of depth of local transport network. All the results follow a similar pattern as those in Section 5.3. In Table A.3, we expand the measure of transport

[42]Note that the magnitudes of the different estimates of β in Table A.1 are not directly comparable to one another, as their effects should be computed at the relevant values of each of the alternative measures of input narrowness. To get an idea of the actual magnitudes, computing the effect of an increase in one standard deviation of road density on exports of industries in the 90th percentile vs. those in the 10th percentile of the input narrowness, the increase in exports when using the Gini is 33.8% higher for the former than for the latter. When using the Herfindahl index the increase differential is 16.9% higher for the 90th relative to the 10th percentile; when using the coefficient of variation this gap is 24.0%; and when using the log-var is 28.4%.

network density to include, in addition to roadways, also railways and waterways. In this case, the density of the transport network of country c is measured as the sum of total kilometers of roadways, railways and waterways, divided by the area of the country. Again, all the results follow a similar pattern as those in Section 5.3.[43]

Table A.4 provides additional robustness checks, by restricting the samples of the regressions reported in column (5) of Table I on a number of dimensions, and also by adding a few additional covariates to that regression. Column (1) restricts the sample to countries with area greater than 10,000 sq km, while column (2) restricts the sample to countries with population larger than half million inhabitants. Results remain qualitatively unaltered when we exclude small countries (either in size or population) from the sample. Column (3) excludes very large countries in terms of their size. In particular, we drop from the sample countries whose area is larger than 3,000,000 sq km.[44] The rationale for this additional robustness check is to account for the possibility that results may be affected by the fact that some very large countries may also have large swaths of uninhabitable land. Next, column (4) uses again the entire sample of countries, but adds interactions terms between countries log area and $Gini_k$, and between log population and $Gini_k$. This would control for the possibility that larger countries may offer more opportunity for input diversity than smaller countries. Column (5) includes an interaction term log GDP and $Gini_k$, in case the aggregate size of the economy may have some impact on specialization in sectors with different degrees of input narrowness. Finally, one additional issue that may raise some concern is the fact that the roads data come from a broad range of years, from 2000 to 2015, while the export data is only from year 2014. In order to assess whether such heterogeneity in the data creates serious problems to our estimates, we restrict the sample of countries in columns (6) and (7), so as to have a more temporally homogenous one. In particular, column (6) excludes from the sample countries whose road network was measured before year 2010, which corresponds to the median year in the sample (see details in Table A.9 in Appendix C). Even more stringently, column (7) restricts the sample of countries to those whose roads were measured either in year 2010 or 2011. Again, all our estimates remain qualitatively unaltered in both subsamples.

[43]Note that for many countries we do not have information on railways or waterways, while we do have information on roadways. For additional comparison, in columns (4), (5) and (6), we also include countries with missing information on either railways or waterways (or in both), replacing the missing values by zeros.

[44]This excludes the following seven countries from the sample: Russia, Canada, United States, China, Brazil, Australia and India. The results are robust to setting the area-threshold for exclusion on alternative levels, such as at 7,000,000 km^2 (which would leave India within the sample), or at 2,000,000 km^2 (which would additionally remove Argentina, Algeria, Congo, and Saudi Arabia from the sample).

TABLE A.2

Transport Density measured by Railway Density

	(1)	(2)	(3)	(4)
Railway Density x Input Narrowness	-2.281***	-1.189***	-0.903***	-0.885***
	(0.133)	(0.159)	(0.162)	(0.163)
Rule of Law x Input Narrowness		-1.567**	-2.408***	-2.359***
		(0.766)	(0.824)	(0.828)
log (Priv Credit/GDP) x Input Narrowness		-4.328***	-2.048**	-2.031**
		(0.936)	(1.006)	(1.007)
log GDP per capita x Input Narrowness		-0.442	0.574	0.979
		(0.746)	(0.782)	(0.775)
Capital Intensity x log $(K/L)_c$			0.009*	
			(0.005)	
Skill Intensity x log H_c			0.007***	
			(0.001)	
Observations	33,099	32,153	26,444	26,444
R-squared	0.754	0.756	0.794	0.793
Number of Countries	122	118	109	109
Number of Industries	294	294	259	259

Robust standard errors in parentheses. All regressions include country and industry fixed effects. Railway density equals the total length of the railway network in km, divided by the area measured in sq km. Data of railway network length is taken from the CIA factbook. *** p<0.01, ** p<0.05, * p<0.1

TABLE A.3

Transport density measured by sum of roadways, railways and waterways per square km

	(1)	(2)	(3)	(4)	(5)	(6)
Transp. (road + rail + waterway) Density x Input Nwness	-7.268***	-2.680***	-1.519***	-4.108***	-2.376***	-1.881***
	(0.519)	(0.598)	(0.587)	(0.302)	(0.361)	(0.375)
Rule of Law x Input Narrowness		-2.381***	-2.892***		-2.394***	-3.091***
		(0.900)	(0.936)		(0.703)	(0.764)
log (Priv Credit/GDP) x Input Narrowness		-4.304***	-2.487**		-3.537***	-1.698*
		(1.072)	(1.114)		(0.822)	(0.913)
log GDP per capita x Input Narrowness		-0.872	-0.054		-0.410	0.911
		(0.874)	(0.933)		(0.602)	(0.656)
Capital Intensity x log $(K/L)_c$			0.009			0.010**
			(0.006)			(0.004)
Skill Intensity x log H_c			0.007***			0.008***
			(0.002)			(0.001)
Observations	25,180	24,234	21,014	42,578	40,692	31,892
R-squared	0.768	0.769	0.805	0.765	0.764	0.794
Number of Countries	92	88	86	166	157	134
Number of Industries	294	294	259	294	294	259

Robust standard errors in parentheses. All regressions include country and industry fixed effects. Columns (1) to (3) only include observations where information on all transport measures (i.e., roadway, railway and waterway) is available. Columns (4) to (6) also include observations where information on either railway or waterway (or both) are missing, replacing the missing values by zero. *** p<0.01, ** p<0.05, * p<0.1

TABLE A.4

Additional Robustness Checks: area, population and year of roads in sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Road Density x Input Narrowness	-2.589***	-1.767***	-2.034***	-1.584***	-1.733***	-1.748***	-1.831***
	(0.579)	(0.399)	(0.514)	(0.485)	(0.512)	(0.511)	(0.667)
Rule of Law x Input Narrowness	-2.869***	-3.929***	-3.172***	-3.363***	-3.350***	-2.888***	-2.615**
	(0.804)	(0.833)	(0.788)	(0.769)	(0.770)	(0.842)	(1.186)
log (Priv Credit/GDP) x Input Nwness	-1.687*	-0.793	-1.611*	-1.422*	-1.600*	-2.123**	-3.288**
	(0.950)	(0.961)	(0.928)	(0.922)	(0.940)	(1.099)	(1.484)
log GDP per capita x Input Narrowness	0.889	0.797	0.983	0.736	5.417	0.207	1.076
	(0.676)	(0.677)	(0.663)	(0.665)	(5.263)	(0.821)	(1.234)
Capital Intensity x log $(K/L)_c$	0.009***	0.011**	0.009**	0.010**	0.010**	0.017***	0.023***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.007)
Skill Intensity x log H_c	0.007***	0.006***	0.007***	0.008***	0.008***	0.012***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
log area x Input Narrowness				0.606	0.572		
				(0.397)	(0.398)		
log population x Input Narrowness				-1.107**	3.599		
				(0.453)	(5.310)		
log GDP x Input Narrowness					-4.666		
					(5.229)		
Observations	29,760	30,085	30,817	31,892	31,892	24,467	13,119
R-squared	0.796	0.780	0.794	0.794	0.794	0.792	0.788
Number of Countries	125	127	129	134	134	101	55
Number of Industries	259	259	259	259	259	294	294
Sample (countries)	> 10 000 km^2	> 500,000 pop	< 3 million km^2	all	all	year roads 2010+	year roads 2010-11

Robust standard errors in parentheses. All regressions include country and industry fixed effects. Column (1) excludes countries whose area is smaller than 10,000 km^2. Column (2) excludes countries whose population is below 500,000 inhabitants. Column (3) excludes countries with area larger than 3,000,000 km^2. Column (6) excludes countries for which the measure of road density in the dataset was recorded before 2010, and column (7) includes only countries whose road length is measured in either year 2010 or 2011. *** p<0.01, ** p<0.05, * p<0.1

TABLE A.5

Additional Robustness Checks: Paved Roads Density

	(1)	(2)	(3)	(4)
(Paved) Road Density x Input Narrowness	-2.401***		-1.830***	
	(0.427)		(0.440)	
Road Density x Input Narrowness		-2.231***		-1.760***
		(0.374)		(0.391)
Rule of Law x Input Narrowness	-2.436***	-2.423***	-3.046***	-2.960***
	(0.765)	(0.764)	(0.810)	(0.812)
log (Priv Credit/GDP) x Input Narrowness	-3.694***	-3.689***	-1.660*	-1.724*
	(0.921)	(0.920)	(1.004)	(1.003)
log GDP per capita x Input Narrowness	-0.350	-0.317	0.786	0.822
	(0.639)	(0.640)	(0.696)	(0.696)
Physical K intesity x log (K/L)			0.011**	0.011**
			(0.005)	(0.005)
Skill intesity x H			0.006***	0.006***
			(0.001)	(0.001)
Observations	35,843	35,843	27,983	27,983
R-squared	0.750	0.750	0.750	0.780
Number of Countries	139	139	118	118
Number of Industries	294	294	294	259

Robust standard errors reported in parentheses. All regressions include country fixed effects and industry fixed effects. The interaction term (Paved) Road Density x Input Narrowness is computed using the share of roads that are defined as paved by the International Road Federation and the World Road Statistics. Countries whose share of paved roads data is dated before 2000 were dropped from the sample. *** p<0.01, ** p<0.05, * p<0.1

Lastly, Table A.5 shows the results corresponding to the previous regressions in column (4) and (5) of Table I, but computing the variable 'road density' using only the total length of *paved* roads.[45] Note that in Table A.5 some of the original countries in the regressions reported in Table I were lost (this is due to lack of sufficiently updated data on share of paved roads for some of the original countries). For that reason, in the sake of comparability, Table A.5 reports in columns (2) and (4) the regressions of the relevant restricted sample of countries, in which 'road density' is computed with both paved and unpaved roads, as previously done in the main text. Besides the fact that results still hold through, an interesting observation is that the estimate for the interaction term using *only* paved roads is (in absolute terms) quantitatively greater than that one obtained when it includes *both* paved and unpaved of roads. This is indeed the kind of behavior one should expect to see if paved roads provide better quality transportation infrastructure than unpaved roads.

B.2 Complementary Results for Section 6.1.1

TABLE A.6

Additional Robustness Checks: Waterways Density

	(1)	(2)	(3)	(4)
Waterways Density x Input Narrowness	-1.634***	-1.921***	-2.037***	-2.155***
	(0.584)	(0.606)	(0.592)	(0.614)
Waterways Density x Input Narrowness x log Income	0.486***	0.554***	0.738***	0.702***
	(0.174)	(0.180)	(0.183)	(0.189)
Rule of Law x Input Narrowness	-4.968***	-4.731***	-3.420***	-3.811***
	(0.822)	(0.853)	(0.875)	(0.904)
log (Priv Credit/GDP) x Input Narrowness	-2.745**	-1.037	-2.793***	-1.076
	(1.071)	(1.112)	(1.071)	(1.113)
log GDP per capita x Input Narrowness	-1.174	-0.235	-1.052	-0.145
	(0.854)	(0.905)	(0.854)	(0.906)
Capital Intensity x log $(K/L)_c$		0.007		0.007
		(0.006)		(0.006)
Skill Intensity x log H_c		0.007***		0.007***
		(0.001)		(0.001)
Road Density x Input Narrowness			-3.903***	-2.328***
			(0.757)	(0.744)
Observations	25,975	22,357	25,975	22,357
R-squared	0.775	0.811	0.775	0.811
Number of Countries	96	93	96	93
Number of Industries	294	259	294	259

Robust standard errors in parentheses. All regressions include country and industry fixed effects. The dependent variable is log($Expo_{c,k}$) in year 2014. Waterways data id taken from the CIA World Factbook, and comprises total length of navigable rivers, canals and other inland water bodies. Waterway density equals internal waterways per square km. ***p<0.01, **p<0.05, *p<0.1

[45] The information on the share of paved roads by country is drawn from the *World Road Statistics* by the *International Road Federation*.

TABLE A.7

Terrain Roughness and Road Density

	Dependent Variable: Roads per Km^2					
	(1)	(2)	(3)	(4)	(5)	(6)
Elevation Difference	-0.020***			-0.012**		
	(0.006)			(0.005)		
Std. Dev. Elevation		-0.591***			-0.363**	
		(0.184)			(0.153)	
% Mountainous			-0.835***			-0.534**
			(0.296)			(0.232)
Ruggedness						
Rule of Law				0.371***	0.329***	0.404***
				(0.117)	(0.112)	(0.118)
Financial Development				0.045	-0.038	-0.046
				(0.133)	(0.114)	(0.098)
log Income				-0.091	-0.275	-0.167
				(0.219)	(0.204)	(0.164)
Human Capital Index				-0.028	-0.076	-0.099
				(0.202)	(0.232)	(0.237)
log $(K/L)_c$				0.257	0.468***	0.364***
				(0.188)	(0.146)	(0.134)
Observations	166	140	142	134	122	126
R-squared	0.100	0.051	0.037	0.306	0.366	0.383

Robust standard reported errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.6 shows the results of regressions analogous to those in Table III in the main text, but instead of splitting the sample of countries according to their income per head, it introduces a triple interaction term between waterway density, the degree of input narrowness and the (log) income per head of countries.

The results on Table A.6 are consistent with those of Table III. In particular, we can see that the triple interaction term carries always a positive and significant coefficient. This suggests that the positive effect of waterways density on specialization in industries with broad input bases tends to be lower for richer economies than it is for lower-income countries.

B.3 Complementary Results for Section 6.1.2

Table A.7 shows in columns (1) - (3) the simple correlation between the different used measures of terrain roughness in country c and road density in c. In all three cases the simple correlation between the variables is negative and highly significant. Next, in columns (4) - (6), we add some additional country-level controls that may be affecting road density (and which are used in the regressions in the main text interacted with industry-level variables). As it can be readily seen, the partial correlation between the three measures of terrain roughness and road density always remains negative and highly significant.

TABLE A.8

Direct Effect of Terrain Roughness

	(1)	(2)	(3)	(4)	(5)	(6)
Road Density x Input Narrowness	-2.180***	-1.771***	-3.252***	-2.086***	-3.277***	-1.884***
	(0.380)	(0.394)	(0.521)	(0.554)	(0.516)	(0.526)
Rule of Law x Input Narrowness	-2.233***	-2.995***	-2.160***	-2.693***	-2.047***	-3.095***
	(0.707)	(0.777)	(0.729)	(0.802)	(0.739)	(0.792)
Fin Dev x Input Narrowness	-3.648***	-1.761**	-4.174***	-2.728***	-4.051***	-1.600*
	(0.822)	(0.910)	(0.877)	(0.953)	(0.872)	(0.937)
log GDP per capita x Input Narrowness	-0.494	0.844	0.133	1.106	0.078	0.948
	(0.605)	(0.665)	(0.669)	(0.720)	(0.641)	(0.675)
Capital Intensity x log $(K/L)_c$		0.010**		0.009**		0.008*
		(0.004)		(0.005)		(0.004)
Skill Intensity x log H_c		0.008***		0.007***		0.008***
		(0.001)		(0.001)		(0.001)
Terrain Roughness x Input Narrowness	***0.376****	***0.186***	***1.331***	***1.601***	***0.015***	***0.027***
	(0.210)	***(0.228)***	***(1.293)***	***(1.427)***	***(0.022)***	***(0.025)***
Measure of Terrain Roughness	Elevation Difference (CIA Wolrd Factbook)		Std. Dev. Elevation (Ramcharan, 2009)		% Mountainous Terrain (Fearon and Laitin, 2003)	
Observations	40,692	31,892	35,988	29,229	36,544	30,067
R-squared	0.764	0.794	0.764	0.794	0.757	0.795
Number of Countries	157	134	137	122	138	126
Number of Industries	294	259	294	259	294	259

Robust standard errors in parentheses. All regressions include country and industry fixed effects. 'Terrain Roughness' is measured in column (1) and (2) by the difference between the max and min elevation within country *c* (source: CIA Factbook), in columns (3) and (4) by the std dev of elevation in country c at the 30" resolution (source: Ramcharan, 2009), and in columns (8) and (10) by the percentage of mountainous in terrain in country *c* (source: Fearon and Laitin, 2003).
*** p<0.01, ** p<0.05, * p<0.1

Table A.8 displays the results of a set of regressions that simultaneously include together as independent variables $r_c \times Gini_k$ and terrain roughness interacted with $Gini_k$. The rationale for these regressions is to try to get some sense of whether, after controlling for the effect of the internal roadway network, terrain roughness may still display a systematic impact on specialization across industries with different degrees of input narrowness. As it can be observed, once the regressions control for the effect of road density, the measures of terrain roughness tend not to exhibit a significant effect on industry specialization.[46] Table A.8 does not represent any sort of test about the validity of the exclusion restriction in the regressions in Table IV. (In fact, there is no way to test the validity of the exclusion restriction in the context of our paper.) Yet, those results are comforting, in the sense that they somehow tame the concerns that the instrument may be capturing some direct effect of topography on specialization in industries with broad input bases, besides the effect mediated through its impact on road density.

[46] The only exception is column (1), where the coefficient is positive and significant at 10%. This estimate would imply that terrain roughness is associated with lower specialization in industries with wide input bases, even after controlling for the impact of road density. Notice, however, that the significance disappears in column (2), after we control for the effect of factor endowments.

Appendix C: Further Data Details

TABLE A.9

country	roads (km)	year (roads)	roads/sq km	gdp pc (2014)	country	roads (km)	year (roads)	roads/sq km	gdp pc (2014)
Angola	51,429	2001	0.04125	7,968	Germany	645,000	2010	1.80661	45,961
Albania	18,000	2002	0.62613	10,664	Djibouti	3,065	2000	0.13211	3,200
UAE	4,080	2008	0.04880	64,398	Dominica	1,512	2010	2.01332	10,188
Argentina	231,374	2004	0.08322	20,222	Denmark	74,497	2016	1.72871	44,924
Armenia	7,792	2013	0.26198	8,586	Dominican Rep.	19,705	2002	0.40487	12,511
Antigua & Barbuda	1,170	2011	2.64108	21,002	Algeria	113,655	2010	0.04772	12,812
Australia	823,217	2011	0.10634	43,071	Ecuador	43,670	2007	0.15401	10,968
Austria	133,597	2016	1.59289	47,744	Egypt	137,430	2010	0.13723	9,909
Azerbaijan	52,942	2006	0.61134	15,887	Spain	683,175	2011	1.35183	33,864
Burundi	12,322	2004	0.44276	772	Estonia	58,412	2011	1.29150	28,538
Belgium	154,012	2010	5.04494	43,668	Ethiopia	110,414	2015	0.09999	1,323
Benin	16,000	2006	0.14207	1,922	Finland	454,000	2012	1.34262	40,401
Burkina Faso	15,272	2010	0.05570	1,565	Fiji	3,440	2011	0.18825	7,909
Bangladesh	21,269	2010	0.14326	2,885	France	1,028,446	2010	1.59746	39,374
Bulgaria	19,512	2011	0.17598	17,462	Gabon	9,170	2007	0.03426	14,161
Bahrain	4,122	2010	5.42368	41,626	United Kingdom	394,428	2009	1.61910	40,242
Bahamas, The	2,700	2011	0.19452	23,452	Georgia	19,109	2010	0.27416	9,362
Bosnia and Herz.	22,926	2010	0.44780	10,028	Ghana	109,515	2009	0.45912	3,570
Belarus	86,392	2010	0.41615	20,290	Guinea	44,348	2003	0.18038	1,429
Belize	2,870	2011	0.12497	8,393	Gambia, The	3,740	2011	0.33097	1,544
Bermuda	447	2010	8.27778	57,531	Guinea-Bissau	3,455	2002	0.09564	1,251
Bolivia	80,488	2010	0.07327	6,013	Equatorial Guinea	2,880	2000	0.10267	40,133
Brazil	1,580,964	2010	0.18565	14,871	Greece	116,960	2010	0.88635	25,990
Barbados	1,600	2011	3.72093	14,220	Grenada	1,127	2001	3.27616	11,155
Bhutan	10,578	2013	0.27551	6,880	Guatemala	17,332	2015	0.15917	6,851
Central African Rep.	20,278	2010	0.03255	594	Hong Kong	2,100	2015	1.89531	51,808
Canada	1,042,300	2011	0.10439	42,352	Honduras	14,742	2012	0.13152	4,424
Switzerland	71,464	2011	1.73133	58,469	Croatia	26,958	2015	0.47634	21,675
Chile	77,764	2010	0.10285	21,581	Haiti	4,266	2009	0.15373	1,562
China	4,106,387	2011	0.42788	12,473	Hungary	203,601	2014	2.18860	25,758
Cote d'Ivoire	81,996	2007	0.25428	3,352	Indonesia	496,607	2011	0.26075	9,707
Cameroon	51,350	2011	0.10801	2,682	India	4,699,024	2015	1.42946	5,224
Congo, Dem. Rep.	153,497	2004	0.06546	1,217	Ireland	96,036	2014	1.36661	48,767
Congo, Rep.	17,000	2006	0.04971	4,426	Iran	198,866	2010	0.12066	15,547
Colombia	204,855	2015	0.17987	12,599	Iraq	59,623	2012	0.13603	12,096
Comoros	880	2002	0.39374	1,460	Iceland	12,890	2012	0.12515	42,876
Cabo Verde	1,350	2013	0.33474	6,290	Israel	18,566	2011	0.89389	33,270
Costa Rica	39,018	2010	0.76356	14,186	Italy	487,700	2007	1.61844	35,807
Curacao	550	*N.A.*	1.23874	25,965	Jamaica	22,121	2011	2.01265	7,449
Cayman Islands	785	2007	2.97348	51,465	Jordan	7,203	2011	0.08062	10,456
Cyprus	20,006	2011	2.16258	28,602	Japan	1,218,772	2015	3.22499	35,358
Czech Republic	130,661	2011	1.65673	31,856	Kazakhstan	97,418	2012	0.03575	23,450

TABLE A.9 (cont.)

country	roads (km)	year (roads)	roads/sq km	gdp pc (2014)	country	roads (km)	year (roads)	roads/sq km	gdp pc (2014)
Kenya	160,878	2013	0.27720	2,769	Paraguay	32,059	2010	0.07882	8,284
Kyrgyzstan	34,000	2007	0.17004	3,359	Qatar	9,830	2010	0.84844	144,340
Cambodia	44,709	2010	0.24696	2,995	Romania	84,185	2012	0.35314	20,817
Korea, South	104,983	2009	1.05278	35,104	Russia	1,283,387	2012	0.07506	24,039
Kuwait	6,608	2010	0.37086	63,886	Rwanda	4,700	2012	0.17845	1,565
Laos	39,586	2009	0.16717	5,544	Saudi Arabia	221,372	2006	0.10298	48,025
Lebanon	6,970	2005	0.67019	13,999	Sudan	11,900	2000	0.00639	3,781
Liberia	10,600	2000	0.09518	838	Senegal	15,000	2015	0.07625	2,247
Sri Lanka	114,093	2010	1.73896	10,342	Singapore	3,425	2012	4.91392	72,583
Lithuania	84,166	2012	1.28891	28,208	Sierra Leone	11,300	2002	0.15751	1,419
Latvia	72,440	2013	1.12155	23,679	El Salvador	6,918	2010	0.32879	7,843
Morocco	58,395	2010	0.13077	7,163	Serbia	44,248	2010	0.57113	13,441
Moldova	9,352	2012	0.27627	4,811	Sao Tome & Princ.	320	2000	0.33195	3,239
Madagascar	37,476	2010	0.06384	1,237	Suriname	4,304	2003	0.02627	15,655
Maldives	88	2013	0.29530	14,391	Slovakia	54,869	2012	1.11898	28,609
Mexico	377,660	2012	0.19225	15,853	Slovenia	38,985	2012	1.92300	30,488
Macedonia	14,182	2014	0.55155	13,151	Sweden	579,564	2010	1.28708	44,598
Mali	22,474	2009	0.01812	1,434	Seychelles	526	2015	1.15604	25,822
Malta	3,096	2008	9.79747	31,644	Syria	69,873	2010	0.37732	4,200
Burma	34,377	2010	0.05081	5,344	Turks and Caicos	121	2003	0.12764	20,853
Montenegro	7,762	2010	0.56198	14,567	Chad	40,000	2011	0.03115	2,013
Mongolia	49,249	2013	0.03149	11,526	Togo	11,652	2007	0.20520	1,384
Mozambique	30,331	2009	0.03794	1,137	Thailand	180,053	2006	0.35090	13,967
Mauritania	10,628	2010	0.01031	3,409	Tajikistan	27,767	2000	0.19269	2,747
Mauritius	2,149	2012	1.05343	17,942	Turkmenistan	58,592	2002	0.12004	20,953
Malawi	15,450	2011	0.13040	949	Trinidad & Tobago	9,592	2015	1.87051	31,196
Malaysia	144,403	2010	0.43779	23,158	Tunisia	19,418	2010	0.11868	10,365
Niger	18,949	2010	0.01496	852	Turkey	385,754	2012	0.49231	19,236
Nigeria	193,200	2004	0.20914	5,501	Tanzania	86,472	2010	0.09128	2,213
Nicaragua	23,897	2014	0.18330	4,453	Uganda	20,000	2011	0.08297	1,839
Netherlands	138,641	2014	3.33729	47,240	Ukraine	169,694	2012	0.28116	10,335
Norway	93,870	2013	0.28990	64,274	Uruguay	77,732	2010	0.44112	20,396
Nepal	10,844	2010	0.07368	2,173	United States	6,586,610	2012	0.66981	52,292
New Zealand	94,902	2012	0.35301	34,735	Uzbekistan	86,496	2000	0.19333	8,195
Oman	60,230	2012	0.19460	38,527	Venezuela	96,189	2014	0.10546	14,134
Pakistan	263,942	2014	0.33155	4,646	British Virgin Isl.	200	2007	1.32450	26,976
Panama	15,137	2010	0.20070	19,702	Vietnam	195,468	2013	0.59016	5,353
Peru	140,672	2012	0.10945	10,993	Yemen	71,300	2005	0.13505	3,355
Philippines	216,387	2014	0.72129	6,659	South Africa	747,014	2014	0.61276	12,128
Poland	412,035	2012	1.31773	25,156	Zambia	40,454	2005	0.05375	3,726
Portugal	82,900	2008	0.90021	28,476	Zimbabwe	97,267	2002	0.24892	1,869

Appendix D: Internal Transport Cost of Final Goods

D.1 Closed Economy Case

This section briefly shows the implications of extending the model in Section 3 to add an iceberg cost for internal transportation final goods. We assume now that shipping final goods between regions A and B entails an iceberg cost $\delta(r)$, where $\delta(r) > 1$ and $\delta'(r) < 0$, for all $r \in \mathbb{R}_+$.

One particularly simplifying feature of the closed-economy model with costless internal transportation of final goods is that it straightforwardly implies that, in equilibrium, the wage must be necessarily be identical in A and B (otherwise all individuals will prefer to live in the region with the higher wage). As the next lemma formally shows, this result remains unaltered when we extend the model in Section 3 to encompass an iceberg cost term $\delta(r) > 1$ for final goods. The reason for regional equality to remain valid lies in the (implicit) technological symmetry of regions coupled with the logarithmic utility function (2). Those two features entail that, in order to equalize the utility of individuals living in A and B, their earnings should be equal.

Lemma 4 *Let Assumptions 1 and 2 hold true, preferences be given by (2), and final goods technologies be the same in regions A and B. Then, in equilibrium, the wage per unit of labor in regions A and B must be identical, no matter the value of the iceberg trade cost of final goods between regions A and B (that is, for any $\delta(r) \geq 1$).*

Proof. Let w_R denote the wage in region R, and take $w_A = 1$ and $w_B = w$ (i.e., take w_A as the numeraire). The marginal costs of producing final good j in region A and B are thus given, respectively, by

$$c_{j,A} = \gamma \left[(1+T) + \frac{1}{(wd(r))^{\theta}} \right]^{-\frac{1-\alpha_j}{\theta}} \left[1 + \frac{1+T}{(wd(r))^{\theta}} \right]^{-\frac{\alpha_j}{\theta}} \tag{37}$$

and

$$c_{j,B} = \gamma \left[\frac{1}{w^{\theta}} + \frac{1+T}{d(r)^{\theta}} \right]^{-\frac{1-\alpha_j}{\theta}} \left[\frac{1+T}{w^{\theta}} + \frac{1}{d(r)^{\theta}} \right]^{-\frac{\alpha_j}{\theta}}. \tag{38}$$

Prices of final goods faced by individuals living in each of the two regions are given by

$$P_j^A = \begin{cases} c_{j,A} & \text{if } \alpha_j \leq \widetilde{\alpha}_B(w) \\ \delta(r) c_{j,B} & \text{if } \alpha_j > \widetilde{\alpha}_B(w) \end{cases} \quad \text{and} \quad P_j^B = \begin{cases} \delta(r) c_{j,B} & \text{if } \alpha_j < \widetilde{\alpha}_A(w) \\ c_{j,B} & \text{if } \alpha_j \geq \widetilde{\alpha}_A(w) \end{cases}, \tag{39}$$

where $c_{j,A}$ is given by (37), $c_{j,B}$ is given by (38), and

$$\widetilde{\alpha}_A(w) \equiv \frac{\ln\left[\frac{1+(1+T)(wd(r))^\theta}{d(r)^\theta+(1+T)w^\theta}\right] - \ln\left(\delta(r)^\theta\right)}{\ln\left[\frac{(1+T)d(r)^\theta+w^\theta}{(1+T)+(wd(r))^\theta}\frac{1+(1+T)(wd(r))^\theta}{d(r)^\theta+(1+T)w^\theta}\right]} \quad \text{and} \quad \widetilde{\alpha}_B(w) \equiv \frac{\ln\left(\delta(r)^\theta\right) + \ln\left[\frac{1+(1+T)(wd(r))^\theta}{d(r)^\theta+(1+T)w^\theta}\right]}{\ln\left[\frac{(1+T)d(r)^\theta+w^\theta}{(1+T)+(wd(r))^\theta}\frac{1+(1+T)(wd(r))^\theta}{d(r)^\theta+(1+T)w^\theta}\right]}. \tag{40}$$

For the sake of the proof, let henceforth both $\widetilde{\alpha}_A(w) \in (0,1)$ and $\widetilde{\alpha}_B(w) \in (0,1)$.[47] The utility obtained by an individual in region A is given by:

$$U^A(w) = \frac{1}{\gamma}\left[\int_0^{\widetilde{\alpha}_B(w)} \left(1+T+\frac{1}{(wd(r))^\theta}\right)^{\frac{1-\alpha_j}{\theta}} \left(1+\frac{1+T}{(wd(r))^{-\theta}}\right)^{\frac{\alpha_j}{\theta}} d\alpha_j \; + \right. \tag{41}$$

$$\left. \frac{1}{\delta(r)}\int_{\widetilde{\alpha}_B(w)}^1 \left(\frac{1}{w^\theta}+\frac{1+T}{d(r)^\theta}\right)^{\frac{1-\alpha_j}{\theta}} \left(\frac{1+T}{w^\theta}+\frac{1}{d(r)^\theta}\right)^{\frac{\alpha_j}{\theta}} d\alpha_j\right].$$

The utility obtained by an individual in region B is:

$$U^B(w) = \frac{w}{\gamma}\left[\frac{1}{\delta(r)}\int_0^{\widetilde{\alpha}_A(w)} \left(1+T+\frac{1}{(wd(r))^\theta}\right)^{\frac{1-\alpha_j}{\theta}} \left(1+\frac{1+T}{(wd(r))^{\theta}}\right)^{\frac{\alpha_j}{\theta}} d\alpha_j \; + \right. \tag{42}$$

$$\left. \int_{\widetilde{\alpha}_B(w)}^1 \left(\frac{1}{w^\theta}+\frac{1+T}{d(r)^\theta}\right)^{\frac{1-\alpha_j}{\theta}} \left(\frac{1+T}{w^\theta}+\frac{1}{d(r)^\theta}\right)^{\frac{\alpha_j}{\theta}} d\alpha_j\right].$$

Free mobility of individuals between regions implies that, in equilibrium, $U^A = U^B$. We first show that $w = 1$ is an equilibrium. Next, we show that it the equilibrium is unique.

Letting $w = 1$, the expressions in (40) boil down to (in order to lighten up notation we skip henceforth the dependence of d and δ with respect to r):

$$\widetilde{\alpha}_A(1) = \frac{1}{2} - \frac{\ln\left(\delta^\theta\right)}{2}\left[\ln\left(\frac{(1+T)d^\theta+1}{(1+T)+d^\theta}\right)\right]^{-1} \quad \text{and} \quad \widetilde{\alpha}_B(1) = \frac{1}{2} + \frac{\ln\left(\delta^\theta\right)}{2}\left[\ln\left(\frac{(1+T)d^\theta+1}{(1+T)+d^\theta}\right)\right]^{-1},$$

from where it can be observed that $\widetilde{\alpha}_B(1) = 1 - \widetilde{\alpha}_A(1)$. Thus, plugging $w = 1$ into (41) yields:

$$U^A(1) = \frac{1}{\gamma}\left[\int_0^{\widetilde{\alpha}_B(1)} \left[(1+T)+d^{-\theta}\right]^{(1-\alpha_j)/\theta} \left[1+(1+T)\,d^{-\theta}\right]^{\alpha_j/\theta} d\alpha_j + \right. \tag{43}$$

$$\left. \frac{1}{\delta}\int_{\widetilde{\alpha}_B(1)}^1 \left[1+(1+T)\,d^{-\theta}\right]^{(1-\alpha_j)/\theta} \left[(1+T)+d^{-\theta}\right]^{\alpha_j/\theta} d\alpha_j\right].$$

[47] Extending the proof to cases in which either $\widetilde{\alpha}_A(w)$ or $\widetilde{\alpha}_B(w)$, or both, could fall outside the unit interval is quite straightforward, hence we skip it in the sake of brevity.

On the other hand, plugging $w=1$ into (42), and defining $\beta_j = 1-\alpha_j$, we can write $U^B(1)$ as:

$$U^B(1) = \frac{1}{\gamma}\left[\int_0^{1-\widetilde{\alpha}_A(1)} \left[(1+T)+d^{-\theta}\right]^{\beta_j/\theta}\left[1+(1+T)\,d^{-\theta}\right]^{(1-\beta_j)/\theta} d\beta_j + \frac{1}{\delta}\int_{1-\widetilde{\alpha}_A(1)}^{1} \left[1+(1+T)\,d^{-\theta}\right]^{\beta_j/\theta}\left[(1+T)+d^{-\theta}\right]^{(1-\beta_j)/\theta} d\alpha_j\right]. \tag{44}$$

Finally, recalling that $\widetilde{\alpha}_B(1) = 1-\widetilde{\alpha}_A(1)$, from (43) and (44) it follows that $U^A = U^B$ when $w=1$, and thus $w=1$ is an equilibrium.

To prove that this equilibrium is unique, note that from (41) it follows that $dU^A(w)/dw < 0$, while (42) could be re-written as

$$U^B(w) = \frac{1}{\gamma}\left[\frac{1}{\delta}\int_0^{\widetilde{\alpha}_A(w)} \left[(1+T)\,w^\theta + d^{-\theta}\right]^{(1-\alpha_j)/\theta}\left[w^\theta+(1+T)\,d^{-\theta}\right]^{\alpha_j/\theta} d\alpha_j + \int_{\widetilde{\alpha}_B(w)}^{1} \left[w^\theta+(1+T)\,d^{-\theta}\right]^{(1-\alpha_j)/\theta}\left[(1+T)+w^\theta d^{-\theta}\right]^{\alpha_j/\theta} d\alpha_j\right],$$

from where we can observe that $dU^B(w)/dw > 0$.[48] As a result of this, for any $w>1$ we would have $U^B(w) > U^A(w)$, whereas for any $w<1$ we would have $U^A(w) > U^B(w)$, implying that $w=1$ is the unique equilibrium value of the relative wage between regions A and B. ∎

The result in Lemma 4 implies all the expressions regarding to the cost of production of final goods, namely (4) - (8), will still hold true in the presence of costly internal transportation of final goods. The difference with respect to Section 3 is that the price at which final good j produced in region R will be sold in region $-R$ will now incorporate the iceberg trade cost, and thus be equal to $\delta(r)\,c_{j,R}$. On the other hand, the price of locally produced goods will still remain equal to $c_{j,R}$. The price gap between locally produced goods and goods sourced from region $-R$ will mean that some goods will end up being produced in *both* regions, and sold *only* locally. The following lemma states this result, extending the previous result in Lemma 1 in the main text.

[48] Note that the effect of $d\widetilde{\alpha}_A(w)/dw$ on $dU^A(w)/dw$ is actually zero, as its effect will perfectly cancel out when considering the two separate definite integrals in (41) and the definition of $\widetilde{\alpha}_A(w)$ in (40). The same argument applies to the effect of $d\widetilde{\alpha}_B(w)/dw$ on $dU^B(w)/dw$, which thus also cancels itself out and turns out to be zero.

Lemma 5 *i) When* $\delta < \left[\frac{(1+T)d^\theta+1}{(1+T)+d^\theta}\right]^{\frac{1}{\theta}}$*, there exists a cut-off level* $\widetilde{\alpha}_A$ *given by*

$$\widetilde{\alpha}_A \equiv \frac{1}{2}\left(1 - \frac{\ln\left(\delta^\theta\right)}{\ln\left[\frac{(1+T)d^\theta+1}{(1+T)+d^\theta}\right]}\right) \tag{45}$$

such that, $0 < \widetilde{\alpha}_A < \frac{1}{2}$*, and i) all producers of final goods for which* $\alpha_j < \widetilde{\alpha}_A$ *will locate in region A, and sell their output to consumers in both A and B; ii) all producers of final goods for which* $\alpha_j > 1 - \widetilde{\alpha}_A$ *will locate in region B, and sell their output to consumers in both A and B; iii) the producers of the final goods for which* $\alpha_j \in [\widetilde{\alpha}_A, 1 - \widetilde{\alpha}_A]$ *will locate in both regions, and will only sell their output to local consumers.*
ii) When $\delta \geq \left[\frac{(1+T)d^\theta+1}{(1+T)+d^\theta}\right]^{\frac{1}{\theta}}$*, all final goods consumed in A are produced by firms located in A, whereas all final goods consumed in B are produced by firms located in B.*

Proof. Consider first the case in which $\delta < \left[\frac{(1+T)d^\theta+1}{(1+T)+d^\theta}\right]^{\frac{1}{\theta}}$. A firm will locate in A and cater to consumers in both A and B when $\delta(r)c_{j,A} < c_{j,B}$. Using the expressions in (7) and (8), this leads to

$$\delta(r)c_{j,A} < c_{j,B} \quad \Leftrightarrow \quad \delta^{\frac{\theta}{1-2\alpha_j}} > \frac{(1+T)d^\theta+1}{(1+T)+d^\theta}.$$

Noting that $\delta^{\theta/(1-2\alpha_j)}$ is increasing in α_j when $\alpha_j < 0.5$, it then follows that for all $\alpha_j < \widetilde{\alpha}_A$, where $\widetilde{\alpha}_A$ is given in (45), we have $\delta(r)c_{j,A} < c_{j,B}$. Next, a firm will locate in B and cater to consumers in both A and B when $c_{j,A} > \delta(r)c_{j,B}$. Using again (7) and (8), this leads to

$$c_{j,A} > \delta(r)c_{j,B} \quad \Leftrightarrow \quad \delta^{\frac{\theta}{2\alpha_j-1}} > \frac{(1+T)d^\theta+1}{(1+T)+d^\theta}.$$

Now, noting that $\delta^{\theta/(2\alpha_j-1)}$ is decreasing in α_j when $\alpha_j > 0.5$, it then follows that for all $\alpha_j > 1 - \widetilde{\alpha}_A$, we have $c_{j,A} > \delta(r)c_{j,B}$. Notice now that the above results in turn imply that, for all $\widetilde{\alpha}_A < \alpha_j < 1 - \widetilde{\alpha}_A$, both $\delta(r)c_{j,A} > c_{j,B}$ and $c_{j,A} < \delta(r)c_{j,B}$ simultaneously hold. Thus, all those final goods such that $\alpha_j \in [\widetilde{\alpha}_A, 1 - \widetilde{\alpha}_A]$ will be locally sourced in both regions. Finally, in the case in which $\delta > \left[\frac{(1+T)d^\theta+1}{(1+T)+d^\theta}\right]^{\frac{1}{\theta}}$, it is straightforward to observe that $\delta(r)c_{j,A} > c_{j,B}$ and $c_{j,A} < \delta(r)c_{j,B}$ hold simultaneously true for all $\alpha_j \in [0, 1]$. ■

In the main text, Lemma 2 states that the marginal cost of production of final goods characterized by α_j closer to $\frac{1}{2}$ fall proportionally more when r increases. This result naturally remains still true in the case of a closed economy with costly internal trade of final goods, since the cost of production of final goods does not depend on $\delta(r)$, which only (possibly) affects the prices of final goods faced by consumers.

D.2 Two-Country World Economy Case

We introduce now costly internal trade of final goods into the Dornbusch-Fischer-Samuelson (1977) framework used in Section 4. Like in Section 4.2, we assume that $\tau > 1$ units of final good j must be shipped from the exporter of j in order for the importer to receive one unit of j. We let δ_H and δ_F denote the internal (iceberg) trade cost of final goods in H and in F, respectively. We keep assuming that $\delta_C = \delta(r_C) > 1$ and $\delta'(r_C) < 0$, for all $r \in \mathbb{R}_+$. Note that, since $r_H > r_L$, we have $\delta_H < \delta_F$. Throughout this section we assume that $\delta_F < \tau$.

When we introduce costly internal trade of final goods, consumers in a given region could face three different cases regarding the price paid for final good j in relation to the marginal cost of production of j. Consider a generic consumer from region $R_C \in \{A_C, B_C\}$, where $C = H, F$. (We use now $-R_C$ to denote the *other* region in country C, and $-C$ to denote the *other* country.) Depending on where exactly each j ends up being produced in equilibrium, a consumer from region R_C will face the following prices according to the origin of the goods:

- **Local goods:** If, in equilibrium, final good j is *actively* produced in region R_C, consumers from R_C will end up paying c_j^C for each unit of j.

- **National non-local goods:** If, in equilibrium, final good j is *not* actively produced in region R_C, but is *actively* produced in region $-R_C$ and it is *not* imported from abroad, then consumers from R_C will end up paying $\delta_C \, c_j^C$ for each unit of j.

- **Imported goods:** If, in equilibrium, final good j is *not* actively produced in region R_C, and it is *imported* from abroad, then consumers from R_C will end up paying τc_j^{-C} for each unit of j, regardless of whether or not j is also actively produced in region $-R_C$.

Introducing costly internal trade of final goods yields an interesting new implication, relative to the model in Section 4.2: some goods will end up being imported from abroad by only one of the two regions in the country, while one of the regions may keep producing it to cater to local consumers. The following proposition expands the results in Proposition 3 to a context with costly interregional trade of final goods, and shows formally this result.

Proposition 3 (bis) *Let $r_H > r_F$. Assume trading final goods between H and F entails an iceberg cost $\tau < \widehat{\tau}$, where $\widehat{\tau}$ is given by* (15)*. Assume also that the internal transportation of final goods between regions A_C and B_C of country $C = H, F$ entails an iceberg cost δ_C, where $\delta_H < \delta_F < \tau$. Then, there exist $\underline{\alpha}_F, \underline{\alpha}'_F, \underline{\alpha}'_H$, and $\underline{\alpha}_H$, where $0 < \underline{\alpha}_F < \underline{\alpha}'_F < \underline{\alpha}'_H < \underline{\alpha}_H < \frac{1}{2}$, and where $\underline{\alpha}_F$ is given by (30) and $\underline{\alpha}_H$ by (29), such that, in equilibrium:*

1. *Imports by H from F: Individuals living in A_H will import from F final goods with $\alpha_j < \underline{\alpha}_F$ and with $\alpha_j > 1 - \underline{\alpha}'_F$. Individuals living in B_H will import from F final goods with $\alpha_j < \underline{\alpha}'_F$, and with $\alpha_j > 1 - \underline{\alpha}_F$.*

2. *Imports by F from H: Individuals living in A_F will import from H final goods with $\alpha_j > \underline{\alpha}_H$ and with $\alpha_j < 1 - \underline{\alpha}'_H$. Individuals living in B_F will import from H final goods with $\alpha_j > \underline{\alpha}'_H$ and with $\alpha_j < 1 - \underline{\alpha}_H$.*

Proof. Notice first that A_H can always produce goods with $\alpha_j < \frac{1}{2}$ at lower marginal cost than B_H, while B_H can always produce goods with $\alpha_j > \frac{1}{2}$ at lower marginal cost than A_H. As a consequence, whenever $\alpha_j < \frac{1}{2}$, the results in Proposition 3 hold true unaltered for consumers in A_H, from where condition $\alpha_j < \underline{\alpha}_F$ obtains for them to import from F. Similarly, when $\alpha_j > \frac{1}{2}$, the results in Proposition 3 hold true unaltered for consumers in B_H, from where $\alpha_j > 1 - \underline{\alpha}_F$ obtains for them to import from F. On the other hand, when $\alpha_j > \frac{1}{2}$, consumers in A_H will import j from F when $\tau c_j^F < \min\{c_j^{A_H}, \delta_H c_j^{B_H}\}$, and given that for $\alpha_j > \frac{1}{2}$ we have $c_j^{A_H} > c_j^{B_H}$, it must be that A_H will import j from F when $\alpha_j > \overline{\beta}_F$, with $\overline{\beta}_F < 1 - \underline{\alpha}_F$. Analogously, when $\alpha_j < \frac{1}{2}$, consumers in B_H will import j from F when $\tau c_j^F < \min\{c_j^{B_H}, \delta_H c_j^{A_H}\}$, and given that for $\alpha_j < \frac{1}{2}$ we have $c_j^{B_H} > c_j^{A_H}$, B_H will import j from F when $\alpha_j < \underline{\beta}_F$, with $\underline{\beta}_F > \underline{\alpha}_F$. Lastly, the fact that $\underline{\beta}_F = \underline{\alpha}'_F = 1 - \overline{\beta}_F$ follows from the symmetry of the regional model.

To prove the result for imports by regions in F, notice that A_F can always produce goods with $\alpha_j < \frac{1}{2}$ at lower marginal cost than B_F, whereas B_F can always produce goods with $\alpha_j > \frac{1}{2}$ at lower marginal cost than A_F. Next, applying an analogous reasoning as above for regions in H, and bearing in mind the symmetry of F's regions, the claimed result obtains.

Finally, to prove that $\underline{\alpha}'_F < \underline{\alpha}'_H$, we follow by contradiction. Suppose then that $\underline{\alpha}'_F \geq \underline{\alpha}'_H$. This would mean that for any $\underline{\alpha}'_H \leq \alpha'_j \leq \underline{\alpha}'_F$, we would have that $\tau c^{A_F} < \min\{\delta_H c^{A_H}, c^{B_H}\}$ (so that region B_H will import the good α'_j from F), together with $\tau c^{A_H} < \min\{\delta_F c^{A_F}, c^{B_F}\}$ (so that region B_F will import the good α'_j from H). Notice now that $\tau c^{A_F} < \min\{\delta_H c^{A_H}, c^{B_H}\}$ implies $\tau c^{A_F} < \delta_H c^{A_H}$ and $\tau c^{A_H} < \min\{\delta_F c^{A_F}, c^{B_F}\}$ implies $\tau c^{A_H} < \delta_F c^{A_F}$. But since $\delta_H < \tau$, it must thus be that $\delta_F c^{A_F} > \tau c^{A_F}$, which contradicts the assumption $\delta_F < \tau$. ■

Unlike the previous results in the main text, Proposition 3 (bis) shows that some goods are imported by only one of the two regions in the country (the region with the higher marginal cost of production of the good). Nevertheless, the main message in Proposition 3 (bis) remains essentially the same as that one put forward by Proposition 3: in equilibrium, H exports goods with α_j located around the mid-point of the unit interval, while F exports those with α_j lying towards both ends of the spectrum of the unit interval.

References

[1] Acemoglu, D., Antràs, P. and Helpman, E. (2007). 'Contracts and Technology Adoption', *American Economic Review*, vol. 97, pp. 916-943.

[2] Acemoglu, D. and Zilibotti, F. (1997). 'Was Prometheus Unbound by Chance? Risk, Diversification and Growth', *Journal of Political Economy*, vol. 105, pp. 709-751.

[3] Agnosteva, D., Anderson, J. and Yotov, Y. (2019). 'Intra-national Trade Costs: Assaying Regional Frictions', *European Economic Review,* vol. 112, pp. 35-50.

[4] Allen, T. and Arkolakis, C. (2014). 'Trade and the Topography of the Spatial Economy', *Quarterly Journal of Economics*, vol. 129, pp. 1085–1140.

[5] Allen, T. and Arkolakis, C. (2019). 'The Welfare Effects of Transportation Infrastructure Improvements', NBER WP No. 25487.

[6] Anderson, J. E., and Van Wincoop, E. (2004). 'Trade costs', *Journal of Economic Literature*, vol. 42, pp. 691–751.

[7] Antràs, P. (2005). 'Incomplete Contracts and the Product Cycle', *American Economic Review,* vol. 95, pp. 1054-1073.

[8] Atkin, D. and Donaldson, D. (2015). 'Who's Getting Globalized? The Size and Implications of Intra-national Trade Costs', Mimeo, MIT.

[9] Barro, R. and Lee, J. (2013). 'A New Data Set of Educational Attainment in the World, 1950-2010', *Journal of Development Economics,* vol 104, pp.184-198.

[10] Beck, T. (2002). 'Financial Development and InternationalTrade. Is There a Link?', *Journal of International Economics*, vol. 57, pp. 107-131.

[11] Becker, B., Chen, J. and Greenberg, D. (2012). 'Financial development, fixed costs, and international trade', *Review of Corporate Finance Studies*, vol. 2, pp. 1-28.

[12] Becker, R., Gray, W. and Marvakov, J. (2013). 'NBER-CES manufacturing industry database: Technical notes', NBER.

[13] Bertinelli, L. and Decrop, J. (2005). 'Geographical agglomeration: Ellison and Glaeser's index applied to the case of Belgian manufacturing industry', *Regional Studies*, vol. 39, pp. 567-583

[14] Blanchard, O. and Kremer, M. (1997). 'Disorganization', *Quarterly Journal of Economics*, Vol. 112, pp. 1091-1126.

[15] Caliendo, L., Parro, F., Rossi-Hansberg, E. and Sarte, P. D. (2018). 'The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy', *Review of Economic Studies*, vol. 85, pp 2042-2096.

[16] Chor, D. (2010). 'Unpacking sources of comparative advantage: A quantitative approach', *Journal of International Economics*, vol. 82, pp. 152-167.

[17] Clague, C. (1991a). 'Relative Efficiency, Self-Containment, and Comparative Costs of Less Developed Countries', *Economic Development and Cultural Change*, vol. 39, pp. 507-530.

[18] Clague, C. (1991b). 'Factor Proportions, Relative Efficiency and Developing Countries' Trade', *Journal of Development Economics*, Vol. 35, pp. 357-380.

[19] Coşar, K. and Fajgelbaum, P. (2016). 'Internal Geography, International Trade, and Regional Specialization', *American Economic Journal: Microeconomics*, vol. 8, pp. 24-56.

[20] Costinot, A. (2009), 'On the Origins of Comparative Advantage', *Journal of International Economics*, vol. 77, pp 255-264.

[21] Cuñat, A. and Melitz, M. (2012). 'Volatility, labor market flexibility, and the pattern of comparative advantage', *Journal of the European Economic Association*, vol. 10, pp 225-254.

[22] Devereux, M., Griffith, R. and Simpson, H. (2004). 'The geographic distribution of production activity in the UK', *Regional Science and Urban Economics*, vol. 34, pp. 533-564.

[23] Donaldson, D. (2018). 'Railroads of the Raj: Estimating the Economic Impact of Transportation Infrastructure', *American Economic Review*, vol. 108, pp. 899-934.

[24] Donaldson, D. and Hornbeck, R. (2016). 'Railroads and American Economic Growth: A "Market Access" Approach', *Quarterly Journal of Economics*, vol. 131, pp 799-858.

[25] Dornbusch, R., Fischer, S. and Samuelson, P. (1977). 'Comparative advantage, trade, and payments in a ricardian model with a continuum of goods', *American Economic Review*, vol. 67.

[26] Duranton, G., Morrow, P. and Turner, M. (2014). 'Roads and Trade: Evidence from the US', *Review of Economic Studies*, vol. 81, pp. 681-724.

[27] Duranton, G. and Overman, H. (2005). 'Testing for Localization Using Micro-Geographic Data', *Review of Economic Studies*, Vol. 72, pp. 1077–1106.

[28] Duranton, G. and Overman, H. (2008). 'Exploring the detailed location patterns of UK manufacturing industries using microgeographic data', *Journal of Regional Science*, vol. 48, pp. 213–243.

[29] Eaton, J. and Kortum, S. (2002). 'Technology, geography, and trade', *Econometrica*, vol. 70, pp. 1741-1779.

[30] Ellison, G., Glaeser, E. and Kerr. W. (2010). 'What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns', *American Economic Review*, vol. 100, pp. 1195-1213.

[31] Ellison, G. and Glaeser, E. (1997). 'Geographic concentration in US manufacturing industries: a dartboard approach', *Journal of Political Economy*, vol. 105, pp. 889-927.

[32] Fajgelbaum, P. and Redding, S. (2014). 'External Integration, Structural Transformation and Economic Development: Evidence from Argentina 1870-1914', NBER WP No. 20217.

[33] Fajgelbaum, P. and Schaal, E. (2017). 'Optimal Transport Networks in Spatial Equilibrium', NBER WP No. 23200.

[34] Fearon, J. and Laitin, D. (2003). 'Ethnicity, Insurgency, and Civil War', *American Political Science Review*, vol. 97, pp. 75–90.

[35] Felbermayr, G. and Tarasov, A. (2015). 'Trade and the Spatial Distribution of Transport Infrastructure', CESifo Working Paper 5634.

[36] Ferguson, S. and Formai, S. (2013) 'Institution-Driven Comparative Advantage and Organizational Choice', J*ournal of International Economics*, vol. 90, pp. 193-200.

[37] Gaulier, G. and Zignago, S. (2010). 'BACI: International Trade Database at the Product-level. The 1994-2007 Version', CEPII Working Paper 2010-23.

[38] Greenwood, J. and Jovanovic, B. (1990). 'Financial Development, Growth, and the Distribution of Income', *Journal of Political Economy*, vol. 98, pp. 1076-1107.

[39] Hillberry, R. and Hummels, D. (2008). 'Trade responses to geographic frictions: A decomposition using micro-data', *European Economic Review*, vol. 52, pp. 527-550.

[40] Imbs, J. and Wacziarg, R. (2003). 'Stages of Diversification', *American Economic Review*, vol. 93, pp. 63-86.

[41] Levchenko, A. (2007). 'Institutional Quality and International Trade', *Review of Economic Studies*, vol. 74, pp. 681-724.

[42] Limao, N. and Venables, A. (2001). 'Infrastructure, geographical disadvantage, transport costs, and trade', *World Bank Economic Review*, vol. 15, pp. 451–479.

[43] Manova, K. (2013). 'Credit Constraints, Heterogeneous Firms, and International Trade', *Review of Economic Studies*, vol. 80, pp. 711-744.

[44] Maurel, F. and Sedillot, B. (1999). 'A measure of the geographic concentration in french manufacturing industries', *Regional Science and Urban Economics*, vol. 29, pp. 575-604.

[45] Matsuyama, K. (2017). 'Geographical Advantage: Home Market Effect in a Multi-Region World', *Research in Economics*, vol. 71, pp. 740-758.

[46] Mesquita Moreira M., Blyde, J., Volpe, C. and Molina, D. (2013). 'Too Far to Export: Domestic transport costs and regional export disparities in latin america and the caribbean', IADB.

[47] Mori, T., Nishikimi, K. and Smith, T. (2005). 'A Divergence Statistic for Industrial Localization', *Review of Economics and Statistics*, vol. 87, pp. 635-651.

[48] Nunn, N. (2007). 'Relationship-specificity, incomplete contracts, and the pattern of trade', *Quarterly Journal of Economics*, vol. 122, pp.569-600.

[49] Nunn, N. and Puga, D. (2012). 'Ruggedness: The blessing of bad geography in Africa', *Review of Economics and Statistics*, vol. 94, pp.20-36.

[50] Rajan, R. and Zingales, L. (1998). 'Financial Dependence and Growth', *American Economic Review*, vol.88, pp. 559-586.

[51] Ramcharan, R. (2009). 'Why an economic core: domestic transport costs', *Journal of Economic Geography,* vol. 9, pp. 559-581.

[52] Ramondo, N., Rodriguez-Clare, A. and Saborio-Rodriguez, M. (2016). 'Trade, Domestic Frictions, and Scale Effects', *American Economic Review*, vol. 106, pp. 3159-84.

[53] Rauch, J. (1999). 'Networks versus markets in international trade', *Journal of International Economics*, vol. 48, pp. 7-35.

[54] Redding, S. (2016). 'Goods Trade, Factor Mobility and Welfare', *Journal of International Economics*, vol. 101, pp. 148-167.

[55] Riley, S., DeGloria, S. and Elliot, R. (1999) 'A Terrain Ruggedness Index That Quantifies Topographic Heterogeneity', *Intermountain Journal of Sciences*, vol. (5), pp. 23–27.

[56] Romalis, J. (2004). 'Factor Proportions and the Structure of Commodity Trade', *American Economic Review*, vol. 94, pp. 67-97.

[57] Saint-Paul, G. (1992), 'Technological Choice, Financial Markets and Economic Development', *European Economic Review*, vol. 36, pp. 763-781.

[58] Santamaria, M. (2018), 'The Gains from Reshaping Infrastructure: Evidence from the Division of Germany', mimeo, Universitat Pompeu Fabra.

[59] Svaleryd, H. and Vlachos, J. (2005) 'Financial markets, the pattern of industrial specialization and comparative advantage: Evidence from OECD countries', *European Economic Review*, vol. 49, pp. 113-144.

[60] Volpe Martincus, C. and Blyde, J. (2013). 'Shaky roads and trembling exports: Assessing the trade effects of domestic infrastructure using a natural experiment', *Journal of International Economics*, vol. 90, pp. 148-161.

[61] Volpe Martincus, C., Carballo, J. and Cusolito, A. (2017). 'Roads, exports and employment: Evidence from a developing country', *Journal of Development Economics*, vol. 125, pp. 21-39.

[62] Yeaple, S. and Golub, S. (2007). 'International productivity differences, infrastructure, and comparative advantage', *Review of International Economics*, vol. 15, pp. 223-242.