Trade Networks in Latin America
Spatial Inefficiencies and Optimal Expansions

Nicole Gorton
Elena Ianchovichina

WORLD BANK GROUP
Latin America and the Caribbean Region
Office of the Chief Economist
November 2021
Abstract

How do trade connectivity issues affect the efficient spatial distribution of economic activity within and across countries in Latin America? This paper uses a spatial general equilibrium framework to construct optimal transport networks and optimal expansions to existing networks in most Latin American countries, as well as within MERCOSUR and the Andean Community. The paper assesses the average annual welfare losses due to inefficient domestic road networks in Latin America at 1.7 percent, ranging from 2.5 percent in Brazil to 0.2 percent in El Salvador. Spatial misallocation of transnational road networks is associated with annual welfare losses of 1.8 percent in MERCOSUR and 1.6 percent in the Andean Community. Optimal investments in improvements and expansions of existing networks can correct these inefficiencies and reduce spatial inequality within countries. These investments correlate relatively well with World Bank road projects because both the model and the World Bank prioritize investments in high population areas. Transnational road improvements benefit the most the least developed country in each trade bloc. The results are robust to changes in data sources and model assumptions.
Trade Networks in Latin America: Spatial Inefficiencies and Optimal Expansions

Nicole Gorton* and Elena Ianchovichina§

Keywords: transport network, spatial equilibrium, road infrastructure, economic geography

JEL Codes: C31, R12, R13, R41, R42, F14

Nicole Gorton is a Ph.D. candidate in economics at University of California, Los Angeles, Tel: (610) 608-8671, E-mail: nicolegorton@g.ucla.edu.

Elena Ianchovichina is a lead economist and the deputy chief economist for the Latin America and the Caribbean Region of the World Bank, 1818 H Street NW, Washington, DC 20433, USA, Tel: +1 202 2803576. E-mail: eianchovichina@worldbank.org.

We would like to thank Pablo Fajgelbaum and Edouard Schaal for sharing the code for their model, Javier Morales Sarriera for his comments on the paper and useful suggestions on road networks data, and Luis Andres, Marek Hanusch, Mathilde Lebrand, William Maloney, Kavita Sethi, and Aiga Stokenberga for their comments on earlier drafts of the paper. The findings, interpretations, and conclusions expressed in this paper are entirely ours. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors or the governments they represent.
1. Introduction

Differences in labor income across regions and municipalities are expected to diminish with development as barriers to factors and goods mobility decline and technology spreads within countries. Yet, spatial inequality has remained high throughout Latin America (Acemoglu and Dell, 2010). One potential reason for this may be insufficient and spatially misallocated road networks, which are an important determinant of trade costs (Limao and Venables, 2001; Atkin and Donaldson, 2015). The findings of transport studies lend support to this hypothesis. Road density in Latin America is comparable to that in Sub-Saharan Africa and lower than elsewhere in the world (Ndulu et al. 2007; World Bank, 2009). While geography and other physical characteristics of Latin American countries may explain the low road density, road occupancy rates are very high and large areas remain inaccessible (Fay et al., 2017).

Railway networks are also underdeveloped, implying that railway services are neither an effective substitute for, nor complement to, road transport (Fay et al., 2017). We explore the extent to which trade connectivity issues affect the efficient spatial distribution of economic activity within and across countries in Latin America either because the existing road infrastructure is spatially misallocated or because it is insufficient. Using data from multiple sources and the general equilibrium spatial framework of Fajgelbaum and Schaal (2020), we construct optimal transport networks and optimal expansions to existing networks in most Latin American countries, as well as within MERCOSUR and the Andean Community. Comparisons between the existing and optimal networks or between existing networks and their optimal expansions allow us to identify inefficiencies in the existing networks and their associated welfare effects.

We find that in several LAC countries, including Argentina, Brazil, and Bolivia, transport networks are relatively inefficient. Misallocation of roads in these countries is associated with annual welfare losses of between 1.9% and 2.5%. Therefore, the model results suggest that inefficiencies in existing road networks generate significant trade frictions in the largest and most populous countries on the continent. At the other end of the spectrum are a few countries with relatively efficient transport networks, including Guatemala, Costa Rica, and El Salvador. These countries will not gain much if the social planner had the power to lift existing networks and place them optimally. As is the case for Europe presented in Fajgelbaum and Schaal (2020), the loss associated with road misallocation in Latin America is equal to 1.1% in simple average terms and 1.7% when countries are weighted according to population, indicating that losses are larger in the region’s more populous countries. In most countries, new road expansions can deliver welfare increases comparable to the gains from replacing existing networks with optimal ones. These gains are largest in the countries with the largest inefficiencies in existing networks, including Argentina, Brazil, and Bolivia. The model-implied optimal investments correlate relatively well with World Bank road investments because both the model and the World Bank prioritize projects in high population areas.

Trade frictions also arise due to inefficiencies in optimal transnational road networks within MERCOSUR and the Andean Community. Spatial misallocation of transnational road networks is associated with average annual welfare losses of 1.8% in MERCOSUR and 1.6% in the Andean Community. These losses can be remedied with road investments that improve and expand the existing road network. In the case of MERCOSUR, expansion yields an annual welfare increase of 2%, while in the case of the Andean Community, the gain is 1.5%. The model suggests improvements in connectivity between the largest cities within MERCOSUR and between the largest cities in each member country. Given the location of these cities, these are improvements mostly along the coastal

---

1 Since it is infeasible to build the optimal road network, its comparison with the existing network provides a sense of the losses due to the misallocation of existing roads.
highways. Most of the investments occur in Brazil (71%) and in Argentina (22%), with the remainder split equally between Uruguay and Paraguay, but the highest welfare gain accrues to the least developed country in the trade bloc – Paraguay (3%). Within the Andean community, half of the infrastructure growth occurs in Colombia, a quarter each in Peru and Ecuador, and only 2% in Bolivia. These investments are again pro-poor benefiting the most the poorest member of the bloc – Bolivia – where welfare increases by slightly less than 5%. Optimal investments improve connectivity between La Paz in Bolivia, along the coast of Peru to Lima and through Quito to Medellin.

The findings in this paper are important because they are informative as to the optimal spatial distribution of road infrastructure projects in Latin America. Policy makers and other stakeholders can use these findings to better target their road investments, with the objective of lowering trade costs and getting a bigger growth boost per dollar spent.

Our paper is related to the literature on the aggregate effects of misallocation (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009), particularly the body of work on spatial resource misallocation due to frictions or government policies (Desmet and Rossi-Hansberg, 2013; Fajgelbaum et al., 2018; and Hsieh and Moretti, 2019). It contributes directly to the large and growing literature on evaluating the economic returns to improving infrastructure systems. Some studies in this literature use historical data to measure the welfare effects of transport infrastructure, including the colonial railways in India (Donaldson, 2010), the railway network in 19th century U.S. (Donaldson and Hornbeck, 2016) and in late 19th century Argentina (Fajgelbaum and Redding, 2014), China’s national trunk highway system (Faber, 2014) and the U.S. highways (Allen and Arkolakis, 2014). Other studies explore the impact of transport infrastructure on different aspects of development. Baum-Snow (2007) assesses the effect of highways on suburbanization in the US, while Sotelo (2016) evaluates the impact of highway investments on agricultural productivity in Peru. More recently, Bird, Lebrand, and Venables (2019) study the effects of the Road Belt Initiative on economic development in Central Asia.

Another part of this literature looks at the role of trade costs in different geographic settings. Eaton and Kortum (2002) and Anderson and Van Wincoop (2003) pioneered an approach that fits standard quantitative trade models to data on the geographic distribution of economic activity across countries; they use this framework to evaluate the economic and welfare effects of exogenous shocks to trade costs. Caliendo et al. (2014) and Ramondo et al. (2016) explore the role of trade frictions within countries in the presence of factor mobility. Many papers study how actual changes in transport costs shape domestic economic activity, including the impact of road expansion on productivity across US industries (e.g. Fernald, 1999) and the impact of US highways on regional economic outcomes (e.g. Chandra and Thompson, 2000 and Duranton et al. 2014).

This paper is most closely linked to applications of the framework of Fajgelbaum and Schaal (2020) to Europe (Fajgelbaum and Schaal, 2020) and Africa (Graff, 2019). Ours is the first study to assess the spatial inefficiencies of existing transport networks and the effects of optimal expansions to existing transport networks for all Latin American countries, except those in the Caribbean,2 using the model of Fajgelbaum and Schaal (2020). Building on their methodology, we develop a new discretization procedure to deal with the large spatial scale of Brazil. This procedure also allows us to study road connections between different groups of Latin American countries.

Misallocation in the transport network is measured by the wedge between the optimal level of investment along each link in the network that the social planner would choose relative to the

---

2 The Caribbean countries are excluded due to their small size and lack of land access to South and Central America.
investment observed in the data. The social planner jointly solves the optimal transport problem and the optimal allocation problem to determine the optimal investment in the transport network, the optimal allocation of production and consumption across locations, and the gross trade flows between locations, given fundamentals including endowments and transport costs. In contrast to the standard models in the literature, transport costs in this model are endogenous as they depend on how much is invested in each link of the road network. The level of investment in each road link and consumption and production in each location are also endogenous and chosen by the social planner to maximize welfare. New road investments reduce transport costs and influence through general equilibrium forces the prices at which goods are produced and sold in each location and the quantities traded between locations. Other things equal, trade flows of a good between a pair of locations is higher when the road infrastructure linking the two locations is better. Trade flows decline with congestion along the link and increase with the price differentials between the two locations.

Like other models with an optimal transport problem (Alder, 2019; Allen and Arkolakis, 2019), the model of Fajgelbaum and Schaal (2020) requires choosing least-cost routes across pairs of locations. However, it offers some distinct advantages compared to other methodologies. First, given the large size of the road networks in many countries in South America, the main benefit is the substantial savings in computation time due to a reduction of the search space in the special convex case with congestion. Such savings are possible because in this case optimal infrastructure investments can be represented as functions of optimal prices, avoiding a direct search in the network space. Second, congestion enables the social planner to focus on the optimization over the transport network itself and conduct a search for the global optimum. In this case, consumption and production in each location are not fixed but respond to the general equilibrium forces in the model and are therefore endogenous. By contrast, in the case of no congestion, which is the typical case in other studies and a special case in Fajgelbaum and Schaal (2020), the optimal transport problem can be solved independently from the general equilibrium outcome by mapping sources with fixed supply to destinations with fixed demand. Only in this special non-convex case without congestion does the global solution in Fajgelbaum and Schaal (2020) match closely with the least cost route optimization solutions in the literature. A third advantage is the flexibility of the framework, which can easily be matched to data on actual transport networks around the world. The model’s fundamentals can be calibrated such that the solution to the planner’s problem matches spatially disaggregated data on population and economic activity given an observed transport network. Counterfactuals can be considered assuming a specific technology to build infrastructure, a specific technology to produce goods, and specific consumer preferences.

The paper is structured as follows. Section 2 discusses the main features of the model, the data used to calibrate the model, and the representation of existing road networks in the model. Section 3 presents the baseline results of optimal road networks and expansions and respectively the associated spatial inefficiencies and gains from optimal expansions. Section 4 explores the robustness of the results. Section 5 presents optimal expansions of trans-national road networks within several groups of countries. We offer concluding remarks in Section 6.

2. Methodology

This section briefly describes the model of Fajgelbaum and Schaal (2020), the data, and the discretized road network representation used to implement the model.

2.1 Model

The social planner solves a triple-nested optimization problem:
\[
\max \max_{l_{jk}} \max_{Q_{jk}} \{c_{j}, h_{j}, D_{jn}, L_{jn}, V_{jn}, X_{jn} \} \sum_{j} L_{j} U(c_{j}, h_{j}), \tag{1}
\]

subject to several constraints including the availability of traded and non-traded goods, the balanced-flows constraint requiring that in each location demand for a good should not exceed its supply net of exports to other locations, the network building constraint, the local factor market clearing conditions, and conditions requiring that consumption, trade flows and factor use are always non-negative. The objective of the social planner is to maximize the sum of population-weighted welfare across locations within each country.\(^3\)

The innermost utility maximization problem in (1) is a standard allocation problem of choosing per-capita consumption of a nontraded good \(h_{j}\) and a composite traded good \(c_{j}\) in location \(j\), bundling together the outputs \(D_{jn}\) of a discrete set of \(N\) tradable sectors. Total labor across each sector \(n\) in each location \(j\) is \(L_{j} = \sum_{n} L_{jn}\). There is a fixed supply \(V_{jn} = (V_{jn}^{1n}, ..., V_{jn}^{Mn})\) of \(M\) primary factors that are immobile across locations but mobile across sectors within each location. Finally, \(X_{jn} = (X_{jn}^{1n}, ..., X_{jn}^{Nn})\) are intermediate inputs, and thus represent the quantity of each sector’s output allocated to producing sector \(n\)’s good in location \(j\).

The second nested problem in (1) is the optimal flow problem. It determines the gross flow of goods, \(Q_{jk}^{n}\), traded along the road link between locations \(j\) and \(k\), regardless of where the goods were produced. Goods transit through the network until they reach the destination where they are consumed or used as an intermediate input. Transporting each unit of good \(n\) from \(j\) to \(k\) requires \(\tau_{jk}^{n}\) units of good \(n\) where the per unit cost is a function of the total quantity of good \(n\) shipped along link \(jk\), \(Q_{jk}^{n}\), and the level of infrastructure on this link, \(I_{jk}\):

\[
\tau_{jk}^{n} = \tau_{jk}(Q_{jk}^{n}, I_{jk}), \tag{2}
\]

where the per-unit cost of shipping is increasing the quantity of good \(n\) shipped, implying decreasing returns or the presence of congestion:

\[
\frac{\partial \tau_{jk}(Q_{jk})}{\partial Q_{jk}} \geq 0. \tag{3}
\]

Congestion reflects the fact that increased shipping activity increases marginal transport costs due to road damage, road accidents, longer travel times, and other potential reasons. In contrast, increased road investments in link \(jk\), which either improve the road surface, widen the road and/or increase the number of road lanes, reduce marginal transport costs:

\[
\frac{\partial \tau_{jk}(Q_{jk})}{\partial I_{jk}} \leq 0. \tag{4}
\]

The transport technology captures variation in geographic characteristics such as ruggedness. If elevation is higher in node \(j\) than \(k\) and it is cheaper to transport goods from \(j\) to \(k\), then \(\tau_{jk} < \tau_{kj}\). The optimal flow problem combines (i) an optimal transport problem of how to map production sources to destinations and (ii) a least-cost route problem with congestion. These two problems must be solved jointly because determining the least-cost routes requires information about the flows and the supply and demand of each good, which are endogenously determined in the solution to the

\(^3\)We focus on this objective function because this paper studies optimal national and transnational transport networks. This framework could also be used to consider transport networks that minimize inequality in a country, or weigh the utility of some locations more heavily than others, but these cases are beyond the scope of this paper.
allocation problem. The transport technology in (2) can be represented as a constant-elasticity transport function:

$$\tau_{jk}(Q, I) = \delta_{jk}^T \frac{Q^\beta}{I^\gamma}. \quad (5)$$

Parameter $\beta$ assumes non-negative values. If $\beta > 0$, the marginal cost of shipping is increasing in the shipped quantity. If $\beta = 0$, the marginal cost of shipping is invariant to the quantity shipped as in the standard iceberg case. Parameter $\gamma$ represents the elasticity of the per-unit transport cost to infrastructure investment. Finally, the scalar $\delta_{jk}^T$ captures the geographic frictions that affect per-unit transport costs, given the quantity shipped and infrastructure $I$.

Finally, the outer-most nested problem in (1) is the optimal network problem. Its solution determines the optimal infrastructure investment $I_{jk}$ in each link $jk$ in the road network. The building of infrastructure requires a resource (i.e. asphalt and/or other materials) whose supply $K$ is fixed in the aggregate within a country. Resource $K$ cannot be used for any other purpose. This assumption implies that the opportunity cost of building infrastructure in any location is only foregoing infrastructure elsewhere. Building infrastructure $I_{jk}$ in link $jk$ requires an investment of $\delta_{jk}^l I_{jk}$ units of $K$. Then the network building constraint is given as:

$$\sum_j \sum_{k \in N(j)} \delta_{jk}^l I_{jk} \leq K, \quad (6)$$

where $N(j)$ is the set of neighboring locations to location $j$ and $j \notin N(j)$. From location $j$, which can be interpreted as a county, goods can be shipped to locations other than $N(j)$, but in this case they must transit through a sequence of connected locations. Thus, total spending on the existing road network is measured as $K$: the total cost of investment across all links in the network. Investment takes place in a link $jk$ when there is some minimum infrastructure $I_{jk}$ in this road link. This investment may be limited by an upper bound, $\bar{I}_{jk}$, imposed by geographic constraints on the capacity to build on the link. Given (5), the optimal level of infrastructure in a link is:

$$I_{jk} = \min \{ \max(I_{jk}, \bar{I}_{jk}), \bar{I}_{jk} \},$$

where $I_{jk}$ is the optimal infrastructure in link $jk$ in the unconstrained optimal network problem (i.e. $I_{jk} = 0, \bar{I}_{jk} = \infty$):

$$I_{jk}^* = \left[ \frac{\gamma}{\bar{P}_K \delta_{jk}^T} \left( \sum_n P_j^n (Q_{jk}^n (P_j^n, P_K^n))^{1+\beta} \right) \right]^{\frac{1}{1+\gamma}}. \quad (7)$$

The network problem is tractable because in this setup optimal infrastructure in a link is only a function of prices in each location. Given the relatively small set of prices, the model is solved link by link, instead of searching in the very large space of all networks. Optimal infrastructure increases with gross flows shipped along link $jk$ and their prices at origin. It decreases with the price of building material, $P_K$, as it increases the cost of building infrastructure; it increases with $\delta_{jk}^T$ reflecting that infrastructure investments offset geographic frictions; and decreases with the marginal cost of building infrastructure $\delta_{jk}^l$. Note that infrastructure affects only trade of goods, not the movement of people.

Fajgelbaum and Schaal (2020) show how, using a no arbitrage condition governing prices across space from the first order conditions of the social planner’s problem, gross flows of each commodity shipped along link $jk$ can be expressed as follows:
\[ Q_{jk}^n = \left[ \frac{1}{1+\beta} \frac{I_{jk}}{\delta_{jk}} \max \left( \frac{p_{kn}}{p_{jn}} - 1, 0 \right) \right]^{\frac{1}{\gamma}} \]

All else equal, as noted by Fajgelbaum and Schaal (2020), trade flows of a commodity between a pair of locations will be higher when infrastructure quality is higher. Similarly, trade flows will be higher when there is more to gain from trade; that is, when the price of the commodity in the destination is much higher than the price in the origin. This result significantly simplifies the data required to implement the model since intranational trade data, from which we could observe \( Q_{jk}^n \), can be hard to obtain in many settings. Given this result, computing trade flows between locations for each commodity requires a measure of geographic trade frictions \( \delta_{jk} \), the price of the commodity in the destination \( k \) relative to the origin \( j \), \( \frac{p_{kn}}{p_{jn}} \), a measure of the infrastructure linking \( j \) and \( k \), \( I_{jk} \), and parameter values for \( \beta \) and \( \gamma \). Section 2.2 describes how we measure \( I_{jk} \), and Section 2.3 describes our parametrization of \( \delta_{jk} \), \( \beta \), and \( \gamma \). The price ratio is computed using equilibrium prices in the calibrated model, a process we describe in detail in Section 2.3.

### 2.2 Data

To implement the model, we must first construct a grid describing the spatial distribution of economic activity and population within each country and a discretized road network describing existing road network connections between each pair of locations in each country. Following Fajgelbaum and Schaal (2020), we divide each country into 1-arc degree cells (in most cases), 0.5 arc-degree cells or 0.25 arc-degree cells (in some cases), depending on the surface area of the country. Table 1 shows, for each country, the size of cells that comprise that country’s grid. Brazil, however, is too large even for a grid of 1 arc-degree cells as in this case the grid is made up of more than 800 cells. When the number of grid cells exceeds approximately 275 cells, the calibration exercise becomes prohibitively computationally intense. Thus, in the case of Brazil, we construct grid cells based on the boundaries of mesoregions, which are subdivisions of Brazilian states used by the Brazilian Institute of Geography and Statistics.

To measure population and value added within each grid cell, we follow Fajgelbaum and Schaal (2020) and obtain population data from NASA-SEDAC’s Gridded Population of the World (GPW) v.4 and value-added data from Yale’s G-Econ 4.0; both datasets refer to 2005. The GPW population data are reported for 30 arc-second cells (approximately 1 km) and the G-Econ value-added data are reported for 1 arc-degree cells (appr. 100km). In Panel (a) of Figure 1 we show the distribution of population across grid cells in Argentina. Here, brighter colored cells represent areas with relatively large populations while areas with darker colored cells represent areas with smaller populations. In the case of Brazil, since our grid cells are consistent with those used by statistical agencies, we use mesoregion level data on population and income as provided by the Brazilian government. In Panel (a) of Figure 2 we show the distribution of the population across grid cells in Brazil.

---

4 This is of particular concern in our setting where reliable data on within country trade is severely lacking.
5 The countries in our setting are quite a bit larger in size relative to those in Europe, which is why we use 1 arc-degree cells in most cases as compared with 0.5 arc-degree cells used in Fajgelbaum and Schaal (2020).
6 In a robustness check, we instead use population data from WorldPop. Details on this check are provided in Section 4.1. We use data corresponding to 2005 as this is the latest year for which G-Econ data are available, and it is important to use population and value-added data from the same year. Another option would be to use Nighttime lights (NTL) data to measure economic activity in each grid cell. However, since rural commodity producing areas are common in our setting, relying on NTL data to measure economic activity may underestimate true output in these areas.
Figure 1. Discretization of Argentina’s Road Network

(a) Population

(b) Nodes and Edges

(c) Actual Road Network

(d) Discretized Road Network

Note: In Panel (a), brightly colored cells indicate areas with higher levels of population while darker cells indicate areas with lower levels of population. In Panel (b), brightly colored nodes are more populous. In Panel (c), green lines indicate primary roads while red segments indicate non-primary roads along the actual road network; same in Panel (d), along the discretized road network.
Figure 2. Discretization of Brazil’s Road Network

(a) Population
(b) Nodes and Edges
(c) Actual Road Network
(d) Discretized Road Network

Note: In Panel (a), brightly colored cells indicate areas with higher levels of population while darker cells indicate areas with lower levels of population. In Panel (b), brightly colored nodes are more populous. In Panel (c), green lines indicate primary roads while red segments indicate non-primary roads along the actual road network; same in Panel (d), along the discretized road network.

Some countries in our sample have unique geographies that pose challenges to road building; we handle these situations on a case-by-case basis. In Peru, Iquitos is surrounded by a natural reserve area and is inaccessible for road building, while other areas of the country are covered by forest. Southern Chile is similarly covered by natural reserve areas and has extremely difficult-to-build geography. Finally, a large lake prevents road building in southeastern Nicaragua. In each of these cases we adjust the baseline grid to accommodate these unique conditions. In Peru, we measure
unbuildable areas using the University of Maryland’s Global tree cover dataset, and consider grid cells to be unbuildable if at least 80% of the grid cell is covered by tree canopy.\(^7\) Using this methodology, we exclude the Northeastern part of the country from the grid since roads cannot be built through the reserve areas.\(^8\) In the case of Chile, we focus on the Northern part of the country and exclude the southern portion of the country based on the distribution of reserve areas as measured by the Chilean government.\(^9\) In Nicaragua, we obtain data on the locations of inland bodies of water from the International Steering Committee for Global Mapping, accessed via New York University. We use this information to restrict the grid to cells that are not majority water.\(^10\)

2.3 Discretization process

The data described above are used to construct the locations of nodes as well as the links in the model’s graph for each country. The GPW data is used to locate the population centroid of each cell; these nodes are typically very close to a node on the road network. We then define the links between nodes in contiguous cells. Since cells are square-shaped, this includes up to 4 nodes connecting horizontal and vertical neighbors and up to 4 nodes along the diagonals. Panel (b) of Figure 1 shows the nodes and edges that comprise the largest possible graph in Argentina. The color of each node corresponds with its relative population, with brighter nodes indicating more populous areas. In the case of Brazil, where our grid cells are neither square in shape nor of uniform size, we define the edges of the underlying graph as follows: two centroids are connected via an edge if they share a border, and are within 800 km from each other.\(^11\) The additional distance criteria are required to avoid linking mesoregions that are technically neighboring but may be very far apart in space and thus not truly directly connected since the distribution of cell sizes (mesoregion areas) varies widely. This graphical representation of connections between locations within each country serves as each country’s baseline grid, from which we build the discretized road network. Panel (b) of Figure 2 shows the nodes and edges for the graph in Brazil.

We then convert each country’s actual road network into a discretized road network. The actual and discretized road networks of Argentina and Brazil are shown on Panels (c) and (d) of Figures 1 and 2, respectively. Our objective is to measure the quality of road network connections between each grid cell within each country. To this end, we obtain data on road networks for each country in our sample from the Global Roads Inventory Project (GRIP), accessed via the World Bank. For each road segment in each country, GRIP provides data on the type of road (motorway, trunk, primary, secondary, tertiary, local), the type of surface (paved, unpaved, asphalt, ground), and the number of lanes along that road segment. However, the data on road quality in GRIP is not as complete as the data available for European or other developed countries. For example, the number of lanes for each road segment listed by GRIP is highly incomplete; in most cases, information on this attribute is missing. We thus supplement the GRIP road network data with road network data from OpenStreetMap (OSM). For each country in our sample, we calculate the average number of lanes by

---

8 Taking these geographies into account are important for devising policy-relevant results. Without taking into account the geography surrounding Iquitos, for example, the optimal road network connects Iquitos to Lima and other surrounding cities.
9 For details, please see [http://areasprotegidas.mma.gob.cl/areas-protegidas/](http://areasprotegidas.mma.gob.cl/areas-protegidas/).
10 Although the Amazonian region of Brazil poses a similar problem, we do not impose additional restrictions on building in the Amazon because neither the optimal expansion nor the optimal reallocation involve building roads in these forested areas.
11 Our results are robust to using instead 400km in the definition of neighboring cells.
road type in the OSM data and use these values as our measure of the number of lanes for the corresponding country and road type in the GRIP data. For example, we compute the average number of lanes on primary roads in Colombia and assume that all primary roads in Colombia have this number of lanes. One downside to this approach is that there is little variation in the number of lanes as we fail to capture roads with very many or very few lanes. Still, we use GRIP as our primary data source because it has much more complete information on the type of road and road surface than the OSM data. In a robustness check, detailed in Section 4.1, we instead use road speeds to measure road quality, as in Graff (2019), and obtain similar results.

In our baseline results, we use the measure of infrastructure quality in Fajgelbaum and Schaal (2020). We use road network attributes from each country’s road network to construct a measure of infrastructure quality along each road segment in each country’s discretized road network. To do this, we first compute the shortest path along the observed road network between each set of nodes in the network. The optimal path on the road network between \( j \) and \( k \) is \( P(j,k) \). We then measure the average number of lanes and average road type along a link from \( j \) to \( k \), respectively, as:

\[
\text{lanes}_{jk} = \sum_{s \in P(j,k)} \omega_{jk}(s)\text{lanes}(s)
\]

\[
\text{nat}_{jk} = \sum_{s \in P(j,k)} \omega_{jk}(s)\text{nat}(s)
\]

where \( s \in P(j,k) \) is each road network segment in the optimal route between \( j \) and \( k \) and \( \omega_{jk}(s) = \frac{\text{length}(s)}{\sum_{s' \in P(j,k)} \text{length}(s')} \) is the length traveled along segment \( s \) relative to the length of the journey along the road network from \( j \) to \( k \). Then, we measure average infrastructure quality along each link as \( I_{jk}^{\text{obs}} = \text{lanes}_{jk} \times K(1-\text{nat}_{jk}) \). Following Fajgelbaum and Schaal (2020), we set \( K = \frac{1}{5} \), which is in line with the ratio of construction and maintenance costs along trunk roads relative to highways as reported by Doll et al. (2008). Note that for routes traveled completely along national routes, the average infrastructure index will be equal to the average number of lanes along the route. We impose that infrastructure quality is symmetric, so that \( I_{jk}^{\text{obs}} = I_{kj}^{\text{obs}} \). In some cases, the optimal road network route deviates considerably from the cells containing \( j \) or \( k \). We classify these cases as \( P(j,k) = \emptyset \), indicating that there is no direct link between these locations. In panel (c) of Figure 1, we show Argentina’s actual road network. Here, primary roads are colored in green and secondary roads are colored in red. In panel (d), we show the discretized road network where we have converted the nodes and edges as shown in Panel (b) into a discretized version of the actual road network. Figure 2 shows the same set of figures for Brazil, where we take a different approach and focus on connections between mesoregions, to accommodate Brazil’s large size.

Table 1 displays the characteristics of the actual and discretized networks for the countries in our sample.\(^\text{12}\) In every case, the discretized network is much smaller in distance than the actual road network since the length of the discretized network reflects the shortest path between each set of nodes in the discretized network, while the actual road network comprises thousands of segments connecting thousands of nodes throughout the country. In other words, the discretized network summarizes the actual road network’s connections between the population centers of each grid cell.

\(^{12}\) Note that here we are not using any model-generated results. This table shows, in columns 1 and 2, characteristics of the actual road network, computed directly from data and, in columns 3 through 6, of the discretized version of the actual road network, the creation of which we describe above in Section 2.3.
Average infrastructure, while correlated with the average number of lanes per kilometer, is often lower in magnitude than the average number of lanes in each country. This is because we adjust the average infrastructure index to account for travel along primary versus non-primary roads. For example, as shown in Panels (c) and (d) of Figure 1, much of southern Argentina can only be traversed along non-primary roads; thus, average infrastructure quality across discretized edges in Argentina is considerably lower than the average number of lanes, since travel between a considerable number of nodes involves travel along non-primary roads.

Table 1. Road Network Summary Statistics

<table>
<thead>
<tr>
<th>Country</th>
<th>ACTUAL NETWORK</th>
<th>DISCRETIZED NETWORK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Length of Network (KM)</td>
<td>Average Lanes per KM</td>
</tr>
<tr>
<td>Argentina</td>
<td>855,713</td>
<td>1.97</td>
</tr>
<tr>
<td>Bolivia</td>
<td>121,098</td>
<td>1.16</td>
</tr>
<tr>
<td>Brazil</td>
<td>727,041</td>
<td>1.99</td>
</tr>
<tr>
<td>Chile</td>
<td>118,456</td>
<td>2.03</td>
</tr>
<tr>
<td>Colombia</td>
<td>410,003</td>
<td>1.72</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>20,254</td>
<td>1.14</td>
</tr>
<tr>
<td>Ecuador</td>
<td>87,531</td>
<td>1.35</td>
</tr>
<tr>
<td>El Salvador</td>
<td>30,313</td>
<td>2.01</td>
</tr>
<tr>
<td>Guatemala</td>
<td>42,981</td>
<td>1.22</td>
</tr>
<tr>
<td>Mexico</td>
<td>796,553</td>
<td>2.16</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>26,952</td>
<td>1.44</td>
</tr>
<tr>
<td>Panama</td>
<td>11,196</td>
<td>2.09</td>
</tr>
<tr>
<td>Paraguay</td>
<td>49,434</td>
<td>1.54</td>
</tr>
<tr>
<td>Peru</td>
<td>226,267</td>
<td>1.95</td>
</tr>
<tr>
<td>Uruguay</td>
<td>81,089</td>
<td>2.03</td>
</tr>
<tr>
<td>Venezuela, RB</td>
<td>123,567</td>
<td>1.35</td>
</tr>
</tbody>
</table>

Note: The average infrastructure index is the distance-weighted, road-type weighted average number of lanes connecting two different grid cells.

2.4 Assumptions and Calibration

Our baseline results assume fixed labor, such that people cannot reallocate across space as infrastructure changes, and the convex case of the parameterization on congestion as described in Section 2.1.\textsuperscript{13} We set $\beta$, the parameter governing congestion, to 0.13 and $\gamma$, the parameter governing returns to infrastructure, to 0.10, as calibrated by Fajgelbaum and Schaal 2020 to match existing

\textsuperscript{13} In a robustness check, we consider the case where labor is perfectly mobile within countries.
We maintain the assumptions on preferences, production, and the values of parameters as outlined in Fajgelbaum and Schaal (2020). In particular, we assume that individuals have Cobb-Douglas preferences over traded and non-traded goods with the parameter $\alpha$ governing the share of non-traded goods in consumption. Traded goods enter the utility function through a CES aggregator across goods produced in each location with elasticity of substitution $\sigma$. We set $\alpha = 0.4$ and $\sigma$, the demand elasticity, to 5.

Production is a linear function of productivity and labor, where $n$ is location $n$ in country $j$:

$$Y^m_j = z^m_j L^m_j$$

Following Fajgelbaum and Schaal (2020), we assume that each location within a country produces a single tradable good. In each country, we allow for 10 differentiated goods plus a homogenous good. Each differentiated good is produced by a unique producer; we assume that the 10 differentiated producers are located in the 10 most populous grid cells in each country. The homogenous good is produced in the remaining cells. Finally, we assume that geographic trade frictions, which enter the transport technology in equation (5), are $\delta^T_{jk} = \delta^T_0 \text{dist}_{jk}$. In lieu of data on interregional trade in Latin America or any existing estimates of this parameter from the region from which to draw, we use the value of $\delta^T_0$ calibrated by Fajgelbaum and Schaal (2020) to match the level of intra-regional trade in Spain.

Given the observed discretized road network in each country, data on population and value added in the underlying grid of each country, and the parameterization described above, we calibrate each grid cell’s productivity to match observed value added. In Figure 3, which pools all countries in our sample together, we show that with this calibration the model is able to closely match the distribution of income in the data.

Finally, implementing counterfactuals also requires assumptions on the cost of building along each link in the discretized road network. For this, we maintain the functional form and parameterization of the building cost function used in Fajgelbaum and Schaal (2020):

$$\ln\left(\frac{\delta^I^m_{jk}}{\text{dist}_{jk}}\right) = \ln(\delta^I_0) - 0.11 \times 1(\text{dist}_{jk} > 50\text{km}) + 0.12 \times \ln(\text{ruggedness}_{jk}),$$

where ruggedness along an edge between $j$ and $k$ is measured as:

$$\text{ruggedness}_{jk} = \frac{1}{2}(\text{ruggedness}_k + \text{ruggedness}_j),$$

and ruggedness for each grid cell is constructed as the standard deviation of the change in elevation across neighboring grid cells. We use elevation data from the ETOPO1 Global Relief Model, which is provided at a finer granularity than our grid cells, to compute:

---

14 This combination of parameter values, based on empirical estimates, is consistent with the convex case of the parameterization. Thus, we consider the convex case of the parameterization in our baseline results and leave the non-convex case to a robustness check.

15 While we do not explicitly model the presence of ports, the vast majority of major ports are captured as differentiated producers since they are high population areas. As a result, even though the model does not observe imports and exports from outside each country's borders, roads connecting ports are nevertheless prioritized.

16 In a robustness check, available upon request, we vary this parameter by plus 10% and minus 10% for the case of Brazil and do not find significant changes in our results.
\[
    \text{ruggedness}_n = \left( \sum_{i \in J(n)} \sum_{k \in N(i)} (elev_i - elev_k)^2 \right)^{1/2},
\]
where \(J(n)\) is the set of ETOPO1 grid cells contained in country’s discretized grid cell \(n\) and \(N(i)\) is the set of eight neighboring cells to each cell \(i\) in ETOPO1. Ruggedness will be higher in places with large changes in elevation. Given this cost function, the model recognizes that it is more costly to build in places with large changes in elevation (i.e. mountainous areas), and less costly per kilometer to build along longer links.

Note that with this methodology, we measure the cost of building roads along a link \textit{up to scale} in each country. To implement counterfactual analyses, we will set \(K=1\) in the network building constraint of equation (6), so we re-scale each \(\delta_{jk}^4\) to satisfy this constraint. As a result, the cost of building in each link is measured as a share of the initial network size. We do not assign a dollar value cost to building along each link, and thus we cannot perform a cost-benefit analysis of any simulated infrastructure expansions.\(^\text{17}\)

Figure 3. Calibrated Model Fit

![Figure 3. Calibrated Model Fit](image)

Linear regression slope (robust SE): Fixed Labor: 1.012 (.005)

3. Baseline Results

In our main results, we consider two counterfactual scenarios: a reallocation of the existing road network and a 50% expansion of the existing road network. The reallocation counterfactual studies what each country’s road network would look like if a social planner took all the resources used to build the country’s existing road network, as measured by \(K\) in the network building constraint of equation (6), and changed the spatial distribution of the investments made with these resources in

\(^{17}\) Given data on the monetary cost of building along links with different geographic characteristics, this information could be used to construct a cost function, and then a cost-benefit analysis would be feasible.
order to create the optimal network.\footnote{The model does not require that we take a stance on the potential causes of any misallocation. We do not observe sources of inefficiencies and have no strong priors on what may have caused misallocation in our setting.} Considering this counterfactual allows us to understand the extent to which the existing road network differs from the optimal one, holding constant the resources used to build the network. The expansion counterfactual, on the other hand, asks where the social planner would invest in infrastructure if the resources used to build the existing road network, as measured by $K$ in the network building constraint, were to grow by 50%.

Figure 4 shows the location of infrastructure investments in each scenario for Argentina, Bolivia, Brazil, Colombia, Mexico, Paraguay, and Peru. Figures showing results for the remaining countries in our sample are included in the Appendix. In these figures, brighter green, thicker segments represent links with larger levels of infrastructure growth. Infrastructure growth could mean any investments that improve road network connections between nodes; for example, improving road quality, adding new lanes, or creating new roads. In the reallocation scenarios, results for which are shown in the right-hand side panel in each row, we also have red links which represent areas where infrastructure would be reduced if roads were to be reallocated towards the optimal network. Note that in the reallocation case, reduction in infrastructure along a link (i.e., a link colored in red) does \textit{not} indicate that that link should not have been built; instead, given a fixed level of resources, the level of infrastructure along the link is higher than the social planner would allocate to that link. Similarly, green lines reflect links to which a social planner would allocate \textit{more} resources.

In Argentina, under both scenarios, new infrastructure investments are made to enhance roads radiating from Buenos Aires toward urban centers in Entre Ríos and Santa Fe. Some large investments are also made in poorer areas in the North and in provinces east of the province of Buenos Aires. The optimal network in Argentina indicates overinvestment in roads in the sparsely populated South, areas in the far East and North. In Colombia, new investments center on Bogotá, connecting cities in the North, Northwest, and Southwest. The optimal network points to overinvestment in roads east of Bogotá and in the North of the country. In Mexico, the expansion of the network covers the eastern part of Mexico, starting at the border city of Cuidad Juárez, extending to Monterrey and the densely populated core at the center of which is Mexico City. The model suggests overinvestment in trunk road infrastructure in the Western parts of the country and the Yukatan Peninsula. The model suggests optimal expansions along the coast of Peru. Turning to Bolivia, most investments are in the interior of the country, connected La Paz and El Alto to Santa Cruz. In Paraguay, the optimal investments improve connections between La Asunción and the surrounding cities. In Brazil, the optimal network improves connections along the coast, including between São Paolo, Rio De Janeiro, and the surrounding coastal cities, as well as links connecting northern and southern cities, via Brasilia. In Chile, shown in the Appendix, new investments are concentrated in the central and more populous parts of the country, while the optimal network suggests major overinvestment in the North of Chile (Figure A.1, Panel A2). In the following two sections, we study the factors that influence the distribution of infrastructure investments as well as the welfare gains obtained by each country under each scenario.
Figure 4. Counterfactual Results

A1. Peru Expansion

A2. Peru Reallocation

B1. Argentina Expansion

B2. Argentina Reallocation

C1. Colombia Expansion

C2. Colombia Reallocation
Figure 4 (continued). Counterfactual Results

D1. Mexico Expansion

E1. Bolivia Expansion

F1. Paraguay Expansion

G1. Brazil Expansion

D2. Mexico Reallocation

E2. Bolivia Reallocation

F2. Paraguay Reallocation

G2. Brazil Reallocation
3.1 Drivers of Infrastructure Growth

What drives infrastructure expansion in some places within a country as compared to others? To understand the location of the optimal expansion and reallocation, we pool all the countries together and then estimate the following regression:

\[
\text{infrastructure growth}_{ic} = \alpha + \beta_1 \log(\text{population}_{ic}) + \beta_2 \log(\text{initial infrastructure}_{ic}) + \beta_3 \log(\text{income per capita}_{ic}) + \beta_4 \text{Differentiated Producer} + \delta_c + \epsilon_{ic}
\]  

(7)

where an observation is a cell \(i\) in each country \(c\), \(\delta_c\) is a country fixed effect, and standard errors are clustered at the country level. In Table 2 below we show results from estimating equation (7) across all LAC countries. Columns (1) through (3) correspond to the case of the 50% expansion in road investment while columns (4) through (6) correspond to the reallocation case. In columns (3) and (6) we also include an indicator for whether a grid cell is a differentiated producer.

Table 2. Infrastructure Growth Across Grid Cells

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Expansion</td>
<td>Reallocation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infrastructure</td>
<td>-0.152***</td>
<td>-0.219***</td>
<td>-0.225***</td>
<td>-0.210***</td>
<td>-0.394***</td>
<td>-0.406***</td>
</tr>
<tr>
<td></td>
<td>(0.00793)</td>
<td>(0.0105)</td>
<td>(0.00918)</td>
<td>(0.0213)</td>
<td>(0.0246)</td>
<td>(0.0219)</td>
</tr>
<tr>
<td>Population</td>
<td>0.170***</td>
<td>0.172***</td>
<td>0.172***</td>
<td>0.468***</td>
<td>0.473***</td>
<td>0.473***</td>
</tr>
<tr>
<td></td>
<td>(0.0177)</td>
<td>(0.0110)</td>
<td>(0.0110)</td>
<td>(0.0391)</td>
<td>(0.0233)</td>
<td>(0.0233)</td>
</tr>
<tr>
<td>Income per capita</td>
<td>0.111*</td>
<td></td>
<td>0.111*</td>
<td></td>
<td>0.231*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0380)</td>
<td></td>
<td>(0.0380)</td>
<td></td>
<td>(0.0945)</td>
<td></td>
</tr>
<tr>
<td>Differentiated producer</td>
<td>0.132**</td>
<td></td>
<td>0.132**</td>
<td></td>
<td>0.260***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0349)</td>
<td></td>
<td>(0.0349)</td>
<td></td>
<td>(0.0533)</td>
<td></td>
</tr>
<tr>
<td>Country FE</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>N</td>
<td>1470</td>
<td>1470</td>
<td>1470</td>
<td>1470</td>
<td>1470</td>
<td>1470</td>
</tr>
<tr>
<td>R2</td>
<td>0.492</td>
<td>0.660</td>
<td>0.674</td>
<td>0.281</td>
<td>0.580</td>
<td>0.594</td>
</tr>
</tbody>
</table>

Note: Standard errors in parentheses, * \(p < 0.05\), ** \(p < 0.01\), *** \(p < 0.001\).

The results in columns (1) and (4) show that higher levels of initial infrastructure are associated with smaller increases in infrastructure in both counterfactual scenarios. Initial infrastructure alone can explain almost 50% of the variation in infrastructure growth across grid cells under the expansion scenario and almost 30% under the reallocation scenario. Adding in population as a covariate in columns (2) and (5), we find that grid cells with larger populations experience larger gains in infrastructure; population explains an additional 16% of the variation in infrastructure growth under the expansion scenario and an additional 30% under the reallocation scenario. Finally, in columns (3) and (6), we add income per capita and an indicator for whether a grid cell is a differentiated producer. Grid cells with higher per capita incomes, which are more productive, and cells with differentiated producers, also see higher levels of infrastructure growth. Although differentiated producers are allocated based on the most populous grid cells, these areas become relatively more connected even after controlling for population. In both scenarios, these four covariates explain roughly 60%-70% of the variation in infrastructure growth, with most of the variation being explained by the initial level of infrastructure and the population in the cell.
3.2 Welfare Gains

3.2.1 Aggregate Gains

We measure per-capita consumption in each grid cell in the calibrated model, as well as the per-capita consumption in each grid cell resulting from each counterfactual. Given our parametrization of the utility function described in Section 2.3 and measures of population in each grid cell, we can construct aggregate welfare in each country in the initial equilibrium, and again under each counterfactual. Figure 5 below shows average welfare gains, measured as a percentage of annual consumption, that each country could expect to obtain under a 50% expansion of the existing road network. For example, assuming that consumption is 70% of GDP, a 2.2% welfare gain in Argentina would mean an annual increase in the country’s income of about $6.86 billion.\(^{19}\) We find that Brazil, Argentina, Bolivia, Peru, and Mexico have the most to gain from an expansion of the existing road network, while Panama, Uruguay, Costa Rica, and El Salvador have relatively less to gain. On average, under the expansion counterfactual, countries in Latin America would gain about 1.7% of consumption, weighing each country by its total population, or 1.03% giving each country an equal weight.

Figure 6 below shows welfare gains that countries could expect to obtain from a reallocation of the existing road network. The gains are comparable to those obtained by a 50% expansion of the road network. On average, countries in Latin America would gain about 1.75% of consumption weighing each country by its total population, or 1.05% unweighted. The road networks of Brazil, Argentina, Bolivia, and Peru are relatively more misallocated, with gains from a reallocation ranging from 1.3%-2.5%, while El Salvador, Costa Rica, and Panama are relative less misallocated, with gains below 0.5%. These results are comparable in magnitude to the average gain from network reallocation of 1.7% obtained for Europe by Fajgelbaum and Schaal (2020), as well as to the average gain from network reallocation of 1.1% obtained for Africa by Graff (2019). While welfare gains vary across countries, we find that as in the European case larger countries in terms of population tend to gain more. In addition, as in Fajgelbaum and Schaal (2020), welfare gains from optimal networks are substantially larger under the non-convex parametrization. For example, we find that under the misallocation case Colombia's welfare gains rise from 1.2% to 2.3% and under the expansion case, from 1.2% to 2.9%.

---

\(^{19}\) This is computed as follows: $445.4 \text{ billion} \times 0.7 \times \left(\frac{2.2}{100}\right) = $6.86 \text{ billion}$, where $445.4 \text{ billion}$ is Argentina’s GDP (as measured by the World Bank), 0.7 is consumption’s share of GDP, 2.2% is Argentina’s welfare gain.
Figure 5. Welfare Gains from 50% Expansion of Road Network

![Graph showing welfare gains from 50% expansion of road network across different countries.]

Note: Welfare gains are reported as a percentage of annual consumption in that country.

Figure 6. Welfare Gains from Reallocation of Road Network

![Graph showing welfare gains from reallocation of road network across different countries.]

Note: Welfare gains are reported as a percentage of annual consumption in that country.
3.2.2 Gains within Countries

Aggregate welfare gains hide heterogeneity in gains within countries. In Figures 7a-11a, we show the spatial distribution of welfare gains across grid cells within countries for a few large countries in our sample: Argentina, Brazil, Ecuador, Mexico, and Peru. We focus on the 50% expansion case and show a map of each country with grid cells colored in green showing welfare (as measured by consumption) increases and grid cells colored in orange showing welfare declines. Darker shades correspond to larger increases or declines. On each map, we overlay the road segments where the largest expansions occurred.20 Darker, thicker line segments indicate larger investments, while thinner paler lines correspond to smaller investments. Below each map in Figures 7b-11b, we plot the percentage change in consumption against the log of the initial income per capita for each grid cell in the country.

The results in Figures 7b-11b suggest that the areas that gain the most from the expansion are those that had relatively lower levels of per capita income before the expansion. As noted by Fajgelbaum and Schaal (2020), this is consistent with the social planner’s desire to equalize the marginal utility of consumption across locations. In this sense, the expansion reduces inequalities across space. In addition, the areas that gain the most are often not those where the level of investment was highest. For example, much of Argentina’s major optimal road investments in the expansion scenario radiate from Buenos Aires, but Buenos Aires itself does not gain from these investments (Figure 7a); instead, the surrounding areas, including secondary cities located mostly in the Northwest, are the ones that gain the most from improved connectivity to the capital city. Similarly, in Brazil and Peru, the major optimal road investments improve connections to Sao Paolo and Lima and benefit relatively more the surrounding areas to São Paolo and Lima (Figures 8a and 11a). In Mexico, optimal road investments also tend to reduce spatial inequality as they benefit relatively more the poorer areas in the South and the areas located in the center of Mexico, north of the country’s capital (Figures 10a-b). In Ecuador, the major road investments improve connectivity between the two largest cities (Guayaquil and Quito) and benefit the most the areas to the south and east of Guayaquil and the areas around Quito (Figure 9a).

---

20 To preserve readability of the maps, we do not show all expanded edges but only those with the largest investment increases within each country.
Figure 7a. Spatial Distribution of Gains in Argentina

Figure 7b. Correlation of Gains and Initial Income in Argentina
Figure 8a. Spatial Distribution of Gains in Brazil

Figure 8b. Correlation of Gains and Initial Income in Brazil
Figure 9a. Spatial Distribution of Gains in Ecuador

Figure 9b. Correlation of Gains and Initial Income in Ecuador

Linear regression slope (robust SE): -0.23 (0.06)
Figure 10a. Spatial Distribution of Gains in Mexico

Figure 10b. Correlation of Gains and Initial Income in Mexico
Figure 11a. Spatial Distribution of Gains in Peru

Figure 11b. Correlation of Gains and Initial Income in Peru

Linear regression slope (robust SE): -0.058 (.008)
3.3 Infrastructure Growth and World Bank Investments

How do the model's predicted locations with infrastructure growth and the magnitude of new investments compare with those receiving investments from World Bank projects? We use a list of World Bank Transportation projects in each LAC country between 2005 and 2020 obtained from the World Bank Projects API. Each project is associated with at least one set of latitude and longitude coordinates, and some projects are also associated with a lending amount. The spatial distribution of these projects is shown in Figure 12 below. Many of these investments are located close to cities; for example, several investments are clustered around Asunción in Paraguay, Buenos Aires in Argentina, Lima, Peru, and along the Coast of Brazil. In Appendix Figure A.2., we plot the within-country correlation between population and World Bank investments, where both population and World Bank investments are measured for each grid cells within each country. They are well-correlated in all cases, except in the case of Costa Rica.

Figure 12. Distribution of World Bank Transportation Projects

For each grid cell in each country, we compute the correlation between the model-implied level of infrastructure growth in the 50% expansion counterfactual first with the total amount of dollars spent on infrastructure projects in that grid cell since 2005 and then with the total number of projects in that grid cell. Figure 13 shows these correlations across countries. While this is an imperfect
comparison since we do not observe exactly the links that World Bank investment projects are targeting with these projects, some interesting patterns appear. In most countries, the correlation is quite high; for example, in Paraguay, Guatemala, and Colombia. The reason for this is that within most countries there is a high correlation between population and the level of World Bank investment. This relationship is shown in Appendix Figure A.3: countries with a high level of correlation between population and World Bank investments also have a high level of correlation between model-implied infrastructure growth and World Bank investments. Since the model prioritizes investments in high population areas, we find that World Bank and model-implied investments are generally well-correlated. But this is not, for example, the case in Costa Rica, where most World Bank projects are along the coast while the model highlights investment in the interior, near San Jose.

**Figure 13. Correlations Between Road Network Investments and World Bank Projects**

Note: The blue bars show the country-level correlation between the number of World Bank infrastructure projects within each grid cell and the amount of model predicted infrastructure investment under the 50% expansion counterfactual. The red bars show the country-level correlation between amount of spending on World Bank infrastructure projects within each grid cell and the amount of model predicted infrastructure investment, under the 50% expansion counterfactual. Only countries with more than one observed World Bank Project are included.

4. **Robustness**

In this section, we consider the robustness of our results to alternative data sources, measures of infrastructure quality, grid cell sizes, and size of the expansion.

4.1 **Alternative Data Sources**

We undertake three robustness checks. First, instead of using SEDAC’s GPW dataset to measure the population in each grid cell, we use data from WorldPop. Second, following Graff (2019), we measure infrastructure quality along a link using the average speed that a car could travel between nodes in the network, as computed via OpenStreetMaps.21 Finally, we combine both data sources by

---

21 Note that in this robustness check we do not alter the baseline discretized road network graph; rather, we simply change how we measure the quality of infrastructure along a link.
including WorldPop population data and OSM travel speeds simultaneously. Table 3 shows the welfare gains for each of these robustness checks across both scenarios in each country. In general, the numbers are very similar across specifications. In Appendix Table A.1, we show correlations between changes in infrastructure in each grid cell in each country under the baseline case and under each of the three robustness checks. The correlations are very high, suggesting our results are not sensitive to these different sources of data.

<table>
<thead>
<tr>
<th>Country</th>
<th>Expansion</th>
<th>Reallocation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base</td>
<td>WorldPop</td>
</tr>
<tr>
<td>Brazil</td>
<td>2.5</td>
<td>-2.1</td>
</tr>
<tr>
<td>Argentina</td>
<td>2.2</td>
<td>2.2</td>
</tr>
<tr>
<td>Bolivia</td>
<td>1.9</td>
<td>1.5</td>
</tr>
<tr>
<td>Peru</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>Colombia</td>
<td>1.2</td>
<td>0.9</td>
</tr>
<tr>
<td>Venezuela, RB</td>
<td>1.1</td>
<td>1</td>
</tr>
<tr>
<td>Chile</td>
<td>1.1</td>
<td>1.1</td>
</tr>
<tr>
<td>Paraguay</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>Ecuador</td>
<td>0.7</td>
<td>0.8</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>0.5</td>
<td>0.6</td>
</tr>
<tr>
<td>Guatemala</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Uruguay</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Panama</td>
<td>0.4</td>
<td>0.3</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>El Salvador</td>
<td>0.2</td>
<td>0.2</td>
</tr>
</tbody>
</table>

### 4.2 Rural Road Quality

One concern with our approach to measuring infrastructure quality is that the data we use may overstate the quality of roads in rural areas.\(^{22}\) We therefore explore the robustness of our results to this measure by reducing travel speeds in very isolated areas, where road quality may be lower than measured. We focus here on Brazil and Bolivia, which both include large, remote, difficult-to-access areas where this problem may be especially acute. We identify which grid cells are “rural” in each country based on whether population density is below 20 people per square kilometer.\(^{23}\) In Brazil, we find that about 16% of the population lives in areas that we identify as ultra-low density; in Bolivia, that figure is 29%. Given this set of rural grid cells within each country, we use the OSM version of our results as described in Section 4.1 and reduce speeds by 20% along edges with a rural destination and a rural origin, and by 10% if the origin or destination grid cell is rural.

We focus on the expansion counterfactual and find that this change has almost no effect in the case of Brazil. Here, welfare falls by an iota, from 2.48% to 2.46% and the correlation in the location of investments with this rural version and our baseline version is nearly one. Since this change to the quality of rural roads affects only a small share of the population, it is not too surprising our aggregate welfare results do not change much. In the case of Bolivia, where this adjustment to road speeds affects a larger share of the population, we find a small increase in welfare from 1.47% to 1.5%

---

\(^{22}\) In addition, we do not take into account that weather may render some unpaved rural roads unusable during certain times of the year; thus, their true average quality may be lower than presented in the data.

\(^{23}\) Chomitz et al. (2005) designate this threshold as “ultra-low density” in the context of LAC.
and a correlation in investments’ location of 99%. Thus, we are reassured that our results are robust to changes in our measurement of rural road quality.

### 4.3 Grid Cell Sizes

In our main analysis, different countries with vastly different surface areas are represented as grids with cells of different sizes. For very large countries, like Argentina and Mexico, we use 1 km degree grid cells while for very small countries like El Salvador we use 0.25 km degree grid cells.²⁴ Computationally, it is challenging to check the robustness of our results to different grid cell sizes for every country. In the case of large countries, reducing the grid cell size increases the number of cells such that the optimization is no longer tractable. In the case of small countries, increasing the grid cell size reduces the number of cells to a point where cells are so large relative to the size of the country that there may only be a few cells per country. However, it is important to understand how this modeling choice may affect our results.

To study this question, we focus on two countries whose sizes are amenable to different grid cell sizes: El Salvador and Guatemala. In Table 4 below, we compare the welfare gains for these two countries under their baseline grid with welfare gains with a larger grid and find similar results. The placement of investment within each country is also similar, though it is impossible to compare exactly since the changing grid size prevents comparison at the cell or edge level.

<table>
<thead>
<tr>
<th>Country</th>
<th>Baseline Grid</th>
<th>Larger Grid</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cell Size</td>
<td>Number of Cells</td>
</tr>
<tr>
<td>El Salvador</td>
<td>0.25</td>
<td>32</td>
</tr>
<tr>
<td>Guatemala</td>
<td>0.5</td>
<td>44</td>
</tr>
</tbody>
</table>

### 4.4 10% Expansion

In our main analysis, we considered a 50% expansion counterfactual. However, many countries may face competing priorities and limited fiscal space such that a major infrastructure push may not be possible at this time. Thus, in this section we focus on Brazil and consider a 10% expansion. The welfare gain from a 10% expansion is 1.1% of annual consumption. This result highlights the non-linearity of gains from improving infrastructure: though we reduced the size of the expansion by 80%, the welfare gains obtained under this much smaller expansion are 44% of the gains obtained under the larger, 50% expansion. In terms of infrastructure placement, the reduction in investment in the 10% case as compared with the 50% case is very uniform across links. The grid-cell level change in investment has a correlation coefficient of 98%; the main difference in the two counterfactuals is the total level of investment allocated to each link which is reduced under the smaller expansion.

### 4.5 Labor Mobility

In this section, we relax the assumption of immobile labor and allow for perfect labor mobility across space within countries. In this version, we calibrate the model to match not only income per

---

²⁴ Table 1 shows the grid cell size and the total number of grid cells for each country in our sample.
capita in each location, as in the case of immobile labor, but also to match the distribution of the
correlation between welfare gains in each country under the assumption of fixed labor and
under the assumption of perfect labor mobility, for both the reallocation and expansion
counterfactuals. In the latter case, workers can reallocate across space given the change in
infrastructure and income in each location implied by each counterfactual. In turn, the distribution of
investment outcomes factors in that workers are mobile. The correlation between the welfare gains
under these two different assumptions is very high, around 99%.

Figure 14. Welfare Gains Under Labor Mobility Assumptions

We next study the determinants of population change in different grid cells under each
counterfactual scenario. Following Fajgelbaum and Schaal (2020), we estimate:

\[
\text{population growth}_{ic} = \alpha + \beta_1 \log(\text{population}_{ic}) + \beta_2 \log(\text{tradable income per capita}_{ic}) + \\
\beta_3 \log(\text{consumption per capita}_{ic}) + \beta_4 \log(\text{initial infrastructure}_{ic}) + \\
\beta_5 \text{infrastructure growth}_{ic} + \beta_6 \text{Differentiated Producer} + \delta_c + \epsilon_{ic}
\]  

We next study the determinants of population change in different grid cells under each
counterfactual scenario. Following Fajgelbaum and Schaal (2020), we estimate:

Results from estimating equation (9), where the outcome variable is infrastructure growth, are shown
in columns (1) and (2) of Table 5 for the expansion and reallocation scenarios, respectively. The results
here are very similar to the case of immobile labor, shown in Table 2 of Section 3.1. Infrastructure
grows more in areas with lower initial levels of infrastructure and higher initial population. Results
from estimating equation (9), where the outcome variable is population growth, are shown in columns
(3) and (4). Population grows in areas with higher levels of tradable income per capita and lower levels
of initial consumption per capita. As noted by Fajgelbaum and Schaal (2020), who observe the same
pattern, this results from a reduction in the variance of the marginal utility of consumption across
locations under the optimal network configuration. Population also grows more in the locations with
differentiated producers. This is because differentiated producers receive relatively higher levels of
infrastructure growth: if the indicator for differentiated producer is omitted from the regression, the
coefficient on infrastructure growth becomes significant.
Table 5. Infrastructure and Population Growth with Mobile Labor

<table>
<thead>
<tr>
<th></th>
<th>Infrastructure Growth</th>
<th>Population Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Expansion</td>
<td>Reallocation</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>-0.215***</td>
<td>-0.421***</td>
</tr>
<tr>
<td></td>
<td>(0.0103)</td>
<td>(0.0299)</td>
</tr>
<tr>
<td>Population</td>
<td>0.167***</td>
<td>0.478***</td>
</tr>
<tr>
<td></td>
<td>(0.0117)</td>
<td>(0.0266)</td>
</tr>
<tr>
<td>Tradable Income per</td>
<td>0.0682</td>
<td>0.153</td>
</tr>
<tr>
<td>capita</td>
<td>(0.0361)</td>
<td>(0.0941)</td>
</tr>
<tr>
<td>Differentiated Producer</td>
<td>0.0991*</td>
<td>0.157</td>
</tr>
<tr>
<td></td>
<td>(0.0411)</td>
<td>(0.0758)</td>
</tr>
<tr>
<td>Consumption per capita</td>
<td>-0.0368***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00833)</td>
<td></td>
</tr>
<tr>
<td>Infrastructure Growth</td>
<td>0.00186</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00132)</td>
<td></td>
</tr>
<tr>
<td>Country FE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1249</td>
<td>1249</td>
</tr>
<tr>
<td>R2</td>
<td>0.645</td>
<td>0.598</td>
</tr>
</tbody>
</table>

*Note: Standard errors in parentheses, * p < 0.05, ** p < 0.01, *** p < 0.001.

5. Transnational Trade Networks

In this section, we evaluate optimal expansions and reallocation in two transnational road networks. First, we consider road connectivity within the group of countries that are signatories to the MERCOSUR free trade agreement (FTA) – Argentina, Brazil, Paraguay, and Uruguay. Second, we explore the road networks connecting countries in the Andean Community, a free trade area including Bolivia, Ecuador, Peru, and Colombia. With this analysis we would like to assess the extent to which road infrastructure plays a role in inflating trade costs and thus limiting the gains from these two regional free trade agreements.25

5.1 Discretization

To study optimal expansions and reallocations among groups of countries, we deviate slightly from our original methodology and follow a discretization procedure similar to the one used in the case of Brazil. Although in theory we could implement exactly the same process on groups of countries as we did in earlier sections, in practice the countries in our sample are too large when grouped together for this to be computationally feasible.26 Thus, for our transnational analyses, we do not rely on square, uniformly-sized grid cells but instead use the Level 1 administrative borders (province or state-level) to construct “grid cells” in each country. One challenge with this approach is that it can result in a wide range of cell sizes. For example, Paraguay’s Level 1 regions are much smaller both in surface area and population compared with those in Argentina and Brazil. Thus, in the cases of Paraguay and Uruguay, we combine small provinces together. For example, in our MERCOSUR analysis, Paraguay consists of three regions and Uruguay consists of four regions.

---

25 We do not include any border frictions or tariffs that would affect trade of goods across countries in this analysis.
26 Fajgelbaum and Schaal (2020) adopt this approach for the union of 24 European countries, but the countries in our setting are far too large for this approach to be feasible.
As in our single-country analysis, we use SEDAC’s GPW to identify the most populous place in each cell; these points become the nodes corresponding to the state or province in which they lie. Thus, the total number of nodes in the connected network will be the total number of states across the group of countries. Edges between nodes are determined based on which grid cells are neighbors, meaning that they share a border. This set of nodes and edges forms the basis for our discretized graph. We compute the total population in each state using SEDAC’s GPW and total value added in each state using G-ECON.\(^{27}\) Since we are no longer working with rectangular cells and we observe population at a very fine granularity, we estimate value added in each grid cell as the population-weighted value added. Figure 15 shows the grid that we work with, as well as the relative levels of population in each cell. Purple lines indicate country borders. To measure distances between links and infrastructure quality across links, we use OpenStreetMap to compute travel distances and travel speeds between each node in the network.\(^{28}\) As in our OSM robustness check, we define infrastructure quality along a link as the average travel speed along that link. We also use OSM to identify whether links exist in the real road network: if there is no road network path between two points, we exclude the edge connecting those two points from our discretized road network.

Figure 15, panel (c) shows the discretized road network for the MERCOSUR group of countries; in this case, all links are colored in green because there is no differentiation between primary versus secondary roads. Thicker, brighter links are those with higher quality infrastructure as measured by high travel speeds, while thinner, dimmer lines are those with lower infrastructure quality. Figure 16 shows the same set of figures for the Andean Community countries. Maintaining the same assumptions on preferences and technology described in Section 2.3, we then calibrate the fundamentals of the model as in our single-country analysis. Following Fajgelbaum and Schaal’s application to transnational road networks within Europe, we assume that each country produces a country-specific differentiated product, in addition to a homogeneous good and use the same parameters as in the benchmark case. We assume that the largest locations in terms of observed population within each country produce the differentiated product of that country, while the remaining locations produce the homogeneous product. In the case of MERCOSUR, we set the number of differentiated producers to 7 in Brazil (the largest and most populous country in this setting), 5 in Argentina, and one each in Paraguay and Uruguay. In the case of the Andean Community, we set the number of differentiated producers to 3 in each country except Bolivia, where we assume only one location produces a differentiated product.\(^{29}\)

\(^{27}\) Another option is to use official statistics for population and/or income as we did in the case of Brazil. However, in this cross-country case, we prefer to use a uniform source because we would like to avoid combining slightly different data on countries within each connected group.

\(^{28}\) This is the same methodology used in one of our robustness checks. The road networks of each country are extremely complex, and this approach eases the computational burden associated with combining them together.

\(^{29}\) Choices governing the number of differentiated producers in each country are motivated by the number of cells in each country’s grid; for example, Bolivia has only 9 cells compared with Colombia where there are 33 cells.
Figure 15. MERCOSUR Discretization

(a) MERCOSUR Population

(b) MERCOSUR Nodes and Edges

(c) MERCOSUR Discretized Roads

Note: Brighter cells in panel (a) show cells with larger population. Magenta lines show country borders.
Figure 16. Andean Community Discretization

(a) Andean Community Population
(b) Andean Community Nodes and Edges
(c) Andean Community Discretized Roads

Note: Brighter cells in panel (a) show cells with larger population. Thicker brighter links show higher quality infrastructure in panel (c). Magenta lines show country borders.
5.2 Results

We study the same 50% expansion and reallocation counterfactuals examined on a country-by-country basis in Section 4. First, we consider these counterfactuals in the case of the MERCOSUR countries. Panels (a) and (b) of Figure 17 show the results for the optimal 50% expansion and the optimal reallocation for this group of countries, respectively. In the expansion case, the model tells us that it is best to improve the roads connecting the largest cities of each member country (which in our model are also the locations producing the differentiated good). Given the location of these cities, these are improvements mostly along the coastal highways. Most of the investments, as measured by the percentage of total infrastructure growth, are in Brazil (71%) and Argentina (22%), while the remainder is split between Uruguay and Paraguay. The optimal reallocation produces similar results, with resources allocated away from the less populous areas of Brazil and Argentina to finance the expansion. We find that the expansion yields an annual welfare increase of 2%, while the reallocation would yield a welfare gain of 1.8%. The results indicate that deficiencies in the road network connecting major cities in MERCOSUR member countries increase trade costs and limit the gains from regional trade.

Second, we consider these counterfactual scenarios for countries in the Andean Community. The results are shown in Figure 18. We find that the optimal expansion yields a welfare gain of 1.5% while the optimal reallocation yields a welfare gain of 1.6%. Both the optimal expansion and reallocation improve connections between La Paz in Bolivia, along the coast of Peru to Lima, and through Quito to Medellin. In the case of the reallocation scenario, resources for this investment are drawn from the interior of each country. In the optimal expansion, 50% of the infrastructure growth is in Colombia, 25% is in Peru, 23% in Ecuador, and the remainder in Bolivia.

Welfare gains under both counterfactual scenarios vary across countries. In Figures 19 and 20, we show the welfare gains obtained by each country within each grouping under both counterfactual scenarios. Figure 19 shows country-level welfare gains for MERCOSUR and Figure 20 shows the same for the countries in the Andean Community. Among the MERCOSUR countries, we find that Paraguay (+3%) and Brazil (+2.3%) experience the largest gains; Argentina experiences the smallest gains and a small decline in welfare in the reallocation scenario (-0.22%). Among the Andean Community, gains are largest in Bolivia (+5%) and smallest in Peru (+1%).

Figure 17. MERCOSUR Counterfactual Results
Figure 18. Andean Counterfactual Results

Figure 19. MERCOSUR Welfare Gains by Country
6. Conclusion

We explore the extent to which road connectivity issues affect the efficient spatial distribution of economic activity within and across countries in Latin America either because the existing road infrastructure is spatially misallocated or because it is insufficient. Using the general equilibrium spatial framework of Fajgelbaum and Schaal (2020) and data from multiple sources, we construct optimal transport networks and optimal expansions to existing networks in most Latin American countries, as well as within MERCOSUR and the Andean Community. We assess the average annual welfare losses due to inefficient domestic road networks in Latin America at 1.7%, if weighted by population, and 1.1% in simple average terms. Spatial inefficiencies are highest in Brazil, Argentina, and Bolivia, where they average 2.5%, 2.4% and 1.9%, respectively, and lowest in El Salvador where losses are 0.2%. These results are robust to changes in data sources and model assumptions and suggest that domestic trade costs associated with inefficient road networks are sizable in the most populous economies in Latin America.

We identify optimal expansions to existing networks that can correct these inefficiencies and we show that these investments tend to reduce spatial inequality. The average regional welfare gains of these investments are on par with those assessed by Fajgelbaum and Schaal (2020) for Europe. Model-implied optimal investments in improving and expanding existing networks correlate relatively well with World Bank road investments because both the model and the World Bank prioritize projects in high population areas. Within countries, we show that the optimal road infrastructure investments tend to boost consumption in areas with lower levels of per capita income before the expansion. This is consistent with the model’s objective to equalize the marginal utility of consumption across locations.

Identifying the optimal transnational road networks for the countries that are signatories to MERCOSUR and the Andean Community allows us to determine the extent to which trade costs deter regional trade. We find that spatial misallocation of transnational road networks is associated with average annual welfare losses of 1.8% in MERCOSUR and 1.6% in the Andean Community.
These losses can be remedied with road investments that improve and expand the existing road networks. In the case of MERCOSUR, expansion yields an annual welfare increase of 2%, while in the case of the Andean Community, the gain is 1.5%. In both cases, the transnational expansions benefit the most the poorest country in the trade bloc. The model improves connectivity between the largest cities within MERCOSUR and between the largest cities in each member country. Given the location of these cities, these are improvements mostly along the coastal highways. Most of the investments occur in Brazil (71%) and in Argentina (22%), with the remainder split equally between Uruguay and Paraguay. Within the Andean community, half of the infrastructure growth occurs in Colombia, a quarter each in Peru and Ecuador, and only 2% in Bolivia. Optimal investments improve connectivity between La Paz in Bolivia, along the coast of Peru to Lima, and through Quito to Medellín.

It is important to keep in mind the following caveat. The 50% expansion of the road network depicts a scenario equivalent to a major road infrastructure push. However, the paper does not factor in the financing costs of increasing the size of the infrastructure budget. If resources are raised by increasing taxes or pulling resources from other public investments, the welfare gains of road infrastructure investments would be smaller. Regardless, the findings of the paper are useful and timely because they are indicative of the optimal spatial distribution of road infrastructure projects. By optimally locating their road projects, governments can lower trade costs and achieve a bigger growth boost per dollar spent.

References


Appendix

Figure A.1. Expansions and Reallocations

A1. Chile Expansion

A2. Chile Reallocation
Figure A.1 (continued). Expansions and Reallocations

B1. Venezuela Expansion

B2. Venezuela Reallocation

C1. Nicaragua Expansion

C2. Nicaragua Reallocation

D1. Panama Expansion

D2. Panama Reallocation
Figure A.1 (continued). Expansions and Reallocations

E1. Costa Rica Expansion

E2. Costa Rica Reallocation

F1. El Salvador Expansion

F2. El Salvador Reallocation

G1. Ecuador Expansion

G2. Ecuador Reallocation
Figure A.1 (continued). Expansions and Reallocations

H1. Guatemala Expansion

H2. Guatemala Reallocation

I1. Uruguay Expansion

I2. Uruguay Reallocation
**Figure A.2. Correlations Between Population and World Bank Investments**

Note: These are correlations across grid cells within a country between population and the level of World Bank infrastructure projects, as measured by the number of projects (purple bars) or the total amount of spending (green bars). Only countries with more than one observed World Bank Project are included.

**Figure A.3. Population, Model-Implied Infrastructure, and World Bank Investments**

Note: Each point on the scatterplot is a country. The x-axis shows the correlation between World Bank projects and population within a country while the y-axis shows the correlation between World Bank projects and model-implied infrastructure growth. Blue points use the number of World Bank projects as the measure of WB investment, while red points use the level of spending. Only countries with more than one observed World Bank Project are included.
<table>
<thead>
<tr>
<th>COUNTRY</th>
<th>Expansion corr(Baseline, WorldPop)</th>
<th>Expansion corr(Baseline, OSM)</th>
<th>Expansion corr(Baseline, WP+OSM)</th>
<th>Reallocation corr(Baseline, WorldPop)</th>
<th>Reallocation corr(Baseline, OSM)</th>
<th>Reallocation corr(Baseline, WP+OSM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRAZIL</td>
<td>N/A</td>
<td>99.6</td>
<td>N/A</td>
<td>99.1</td>
<td>N/A</td>
<td>99.3</td>
</tr>
<tr>
<td>ARGENTINA</td>
<td>100</td>
<td>99.6</td>
<td>99.6</td>
<td>100</td>
<td>99.3</td>
<td>99.3</td>
</tr>
<tr>
<td>BOLIVIA</td>
<td>96.3</td>
<td>98</td>
<td>95.7</td>
<td>96.5</td>
<td>96.9</td>
<td>94.5</td>
</tr>
<tr>
<td>PERU</td>
<td>85</td>
<td>98.7</td>
<td>86</td>
<td>85.7</td>
<td>98.2</td>
<td>86.7</td>
</tr>
<tr>
<td>COLOMBIA</td>
<td>96.8</td>
<td>99.3</td>
<td>96.5</td>
<td>97.3</td>
<td>98.7</td>
<td>96.2</td>
</tr>
<tr>
<td>VENEZUELA, RB</td>
<td>98.5</td>
<td>99.5</td>
<td>98.2</td>
<td>97.7</td>
<td>98.9</td>
<td>96.8</td>
</tr>
<tr>
<td>CHILE</td>
<td>99</td>
<td>99.7</td>
<td>99.3</td>
<td>98.6</td>
<td>99.5</td>
<td>98.7</td>
</tr>
<tr>
<td>PARAGUAY</td>
<td>99.6</td>
<td>99.4</td>
<td>99</td>
<td>99.6</td>
<td>98.4</td>
<td>98</td>
</tr>
<tr>
<td>ECUADOR</td>
<td>62.3</td>
<td>99.7</td>
<td>62.8</td>
<td>65.7</td>
<td>99</td>
<td>65.3</td>
</tr>
<tr>
<td>NICARAGUA</td>
<td>98.3</td>
<td>99.7</td>
<td>97.6</td>
<td>98.1</td>
<td>98.9</td>
<td>96.5</td>
</tr>
<tr>
<td>GUATEMALA</td>
<td>99.8</td>
<td>99.8</td>
<td>99.5</td>
<td>99.7</td>
<td>99.4</td>
<td>99.1</td>
</tr>
<tr>
<td>URUGUAY</td>
<td>99.5</td>
<td>99.9</td>
<td>99.4</td>
<td>99.3</td>
<td>99.8</td>
<td>99.2</td>
</tr>
<tr>
<td>PANAMA</td>
<td>56.1</td>
<td>99.8</td>
<td>56</td>
<td>56.1</td>
<td>99.6</td>
<td>56.1</td>
</tr>
<tr>
<td>COSTA RICA</td>
<td>99.4</td>
<td>99.7</td>
<td>99.2</td>
<td>98.7</td>
<td>99.3</td>
<td>98.4</td>
</tr>
<tr>
<td>EL SALVADOR</td>
<td>95.7</td>
<td>99.5</td>
<td>95.6</td>
<td>95.7</td>
<td>98.9</td>
<td>94.7</td>
</tr>
</tbody>
</table>

Note: Correlations represent the correlation across grid cells within a country between the change in infrastructure under our baseline model estimate and each of the corresponding robustness checks. The “Base” column lists our baseline welfare estimate, using GRIP road network quality data and SEDAC-GPW population data. The “WorldPop” column uses WorldPop data on populations of grid cells in lieu of SEDAC-GPW data, and the “OSM” uses travel speeds as computed with OpenStreetMap to measure infrastructure quality in lieu of GRIP measures of road segment quality. Finally, “WorldPop + OSM” uses WorldPop data on the populations of grid cells and OSM data on travel speeds to measure infrastructure quality.