

Trade, Structural Transformation and Development: Evidence from Argentina 1869-1914*

Pablo Fajgelbaum[†]
UCLA and NBER

Stephen J. Redding[‡]
Princeton and NBER

October 12, 2018

Abstract

We provide new theory and evidence on the relationship between economic development and international trade using Argentina's late-19th-century integration into the global economy. We show that structural transformation, from agriculture to non-agriculture, and across disaggregated goods within the agricultural sector, was central to Argentina's rapid export-led economic development. We provide evidence that the reductions in internal transport costs from the construction of the railroad network were important in enabling interior regions to participate in this process of structural transformation and economic development. We rationalize our empirical findings using a theoretical framework that emphasizes a spatial Balassa-Samuelson effect, in which regions with good access to world markets have higher population densities, urban population shares, relative prices of non-traded goods, and land prices relative to wages. In counterfactuals, we find that the construction of the railroad network increases the total population of Argentina by 49 percent under free international migration and raises the common real wage across all Argentinian districts by 8 percent under restricted international migration.

KEY WORDS: international trade, economic development, structural transformation

JEL: F11, F14, O13, O14

*We would like to thank Princeton University and UCLA for research support. Pablo Fajgelbaum thanks the UCLA Ziman Center for Real Estate for financial support. We are grateful to Andres Maggi, Ferdinand Rauch, Mariano Somale and particularly Iván Torre for excellent research assistance. We are grateful to Kei-Mu Yi, Dave Donaldson, Paula Bustos, Rick Hornbeck, Kiminori Matsuyama, Jeremy Adelman, and a number of other colleagues, conference and seminar participants for helpful comments. Responsibility for results, opinions and errors is the authors' alone.

[†]Department of Economics, 8283 Bunche Hall, Los Angeles, CA 90095. Tel: 310-794-7241. Email: pfajgelbaum@gmail.com.

[‡]Julis Romo Rabinowitz Building, Princeton University, Princeton, NJ 08544. Tel: 609-258-4016. Email: reddings@princeton.edu.

1 Introduction

The relationship between trade and development is widely-debated in the fields of both international economics and development economics. A key empirical challenge in this debate is determining the direction of causality in this relationship. A second unresolved issue is the mechanism underlying this relationship. Can rapid economic growth be achieved by scaling up all existing production activities? Or does it instead require a reorganization of production activity across different sectors, from rural to urban areas, and between peripheral and central locations? A third area of continuing dispute concerns the spatial incidence of both international trade and economic development. Does the process of economic growth necessarily increase inequality across regions within countries? When is economic development restricted to narrow coastal areas with good access to world markets and under what conditions can interior regions effectively participate in the global economy?

In this paper, we provide new theory and evidence on these questions using Argentina’s integration into the international economy in the late-19th century. We combine a newly-constructed, spatially-disaggregated dataset for the period 1869-1914 with a quantitative model of economic activity across regions and sectors. Our empirical setting has a number of advantages for addressing these empirical challenges. First, Argentina’s integration into the world economy was driven by late-19th-century reductions in maritime transport costs following the invention of the steamship, which was first developed for river transport in Europe and the United States, and was exogenous to this peripheral location. Second, we have disaggregated data on the distribution of economic activity across regions and sectors within Argentina over a long time horizon, which enables us to quantify the role of structural transformation in economic development. We find that Argentina’s 19th-century export boom was characterized by a high-level of commodity specialization, as observed in many developing countries today, with agriculture accounting for over 99 percent of the value of exports. We use our unusually detailed data on the organization of production activity within the agricultural sector (including crops, livestock and machinery) to establish the importance of structural transformation within agriculture and the emergence of new sources of comparative advantage in cereals and refrigerated and frozen meat. Third, the invention of steam railroads lowered inland transport costs, which enables us to examine the relationship between reductions in internal and external trade frictions. In analyzing this relationship, we use the historical context of Spanish colonial rule and Argentina’s late-19th-century integration into world markets to construct instruments for the railroad network to address the non-random placement of transport infrastructure.

We begin by showing that Argentina’s rapid export-led economic development in the late-19th century involved major changes in the distribution of economic activity across sectors and regions. In particular, we establish five stylized facts about patterns of economic development. First, population density is sharply decreasing in geographical distance from Argentina’s trade hub, as captured by its four leading ports of Buenos Aires, La Plata, Rosario and Bahía Blanca, which together account for more than 75 percent of its exports throughout our sample period, and form a semi-circle surrounding Buenos Aires and its hinterland. Second, this gradient is steeper for urban population density than for rural population density, with the result that the areas closest to world markets have higher urban population shares. Third, this gradient in population density steepens over our sample period, as economic activity expands in the immediate hinterland of Buenos Aires and its surrounding ports. Fourth, railroad access predicted by our instruments raises both urban and rural population density for a given geographical distance from Argentina’s trade hub, consistent with railroads reducing inland transportation costs. Fifth, proximity to Argentina’s trade hub and

railroad access predicted by our instruments are both accompanied by compositional changes within the agricultural sector, away from the traditional comparative advantage products of tanned hides and leather, and towards the new export goods of cereals and refrigerated and frozen meat.

To rationalize these empirical findings, we develop a theoretical model of the distribution of economic activity across both sectors and regions. The key new insight of the model is to establish an interaction between structural transformation across sectors and internal trade costs across regions. Consistent with the macroeconomic literature on structural transformation, we assume inelastic demand between traded goods (i.e. agriculture and manufacturing) and non-traded goods (i.e. services and manufacturing for the local market). In line with the extreme agricultural specialization observed in our export data, we assume that all regions within Argentina have a comparative advantage in agriculture, and we allow the extent of this comparative advantage to differ across disaggregated goods within the agricultural sector. We also make the natural assumption that this agricultural sector is land intensive relative to the non-traded sector. Under these assumptions, we show that our generic neoclassical production structure implies a *spatial Balassa-Samuelson effect*, such that regions with good access to world markets have higher population densities, urban population shares, relative prices of non-traded goods, and land prices relative to wages. The intuition for this result is straightforward. Locations with good access to world markets are attractive for the production and consumption of traded goods, which increases population density, and bids up the reward of the immobile factor (land) relative to the mobile factor (labor). Together the increase in population and the reduction in wages relative to land rents induce an expansion in the employment share of the labor-intensive non-traded sector, which requires a higher relative price for the non-traded good, given inelastic demand between sectors.

Therefore, the model highlights that internal geography not only affects the overall level of economic activity but also shapes structural transformation and the composition of economic activity. We show that this structural transformation occurs not only between agriculture and non-agriculture, but also across disaggregated commodities within the agricultural sector. In particular, we derive a composite measure of adjusted-agricultural productivity, which depends on prices, trade costs and productivity for each disaggregated good within the agricultural sector. This measure of adjusted-agricultural productivity for each location, together with the corresponding measure of productivity in the non-traded sector, is a sufficient statistic for population density and the urban population share in each location. As the construction of the railroad reduces transport costs for some goods more than for others (e.g. cereals and refrigerated and frozen meat versus tanned hides and leather), this induces a change in composition of economic activity within the agricultural sector, which acts like an increase in overall agricultural productivity. We show that the model can be inverted to recover unique values for the sufficient statistics of adjusted-agricultural productivity and non-traded productivity from the observed data on employment in rural and urban areas. We find that the construction of the railroad network predicted by our instruments has a statistically significant positive effect on adjusted-agricultural productivity but not on non-traded productivity, which is consistent with railroads reducing internal transport costs. Finally, we confirm that these changes in adjusted-agricultural productivity are strongly related to measures of specialization in the new export crops of cereals and refrigerated and frozen meat, in line with the compositional changes within the agricultural sector in the model.

Although our reduced-form empirical specifications reveal the relative impact of internal geography on locations with different levels of access to world markets, they do not capture general equilibrium effects or distinguish reallocation from the creation of economic activity. Therefore, we use the structure of the model to undertake counterfactuals,

in which we show that internal geography is important not only for the distribution of economic activity but also for aggregate economic outcomes. We find that these substantial aggregate effects for both the change in the spatial gradient of productivity relative to Argentina’s trade hub and the construction of the railroad network. Assuming that interior regions experienced the same increases in price-adjusted productivities as the coastal regions proximate to Argentina’s trade hub from 1869-1914, we find that total population would have been 271 percent larger under free international migration, and the common real wage would have been 60 percent higher. Focusing solely on the impact of the construction of the railroad network, we find an increase in total population of 49 percent under free international migration, and a rise in the common real wage by 8 percent under restricted international migration. Our estimate for the real income impact of the railroad network of 8 percent is larger than those of 2.7 and 3.2 percent for the late-19th century United States in [Fogel \(1964\)](#) and [Donaldson and Hornbeck \(2016\)](#) respectively, but smaller than the estimate of 16 percent of agricultural real income in [Donaldson \(2018\)](#) for late-19th century India. Whereas all of these studies focus on the agricultural sector, our estimates capture the impact of the railroad network on both urban and rural economic activity, which plausibly explains our estimates being somewhat larger than typically found in these other studies. Therefore, taking both our reduced-form and quantitative evidence together, we find that the reductions in internal transport costs from the construction of the railroad network were important in enabling interior regions to participate in 19th-century Argentina’s rapid export-led economic development.

Our paper is related to a number of different strands of research. First, our work contributes to the macroeconomic literature on structural transformation, including [Matsuyama \(1992, 2009\)](#), [Caselli and Coleman \(2001\)](#), [Ngai and Pissarides \(2007\)](#), [Herrendorf, Schmitz, and Teixeira \(2012\)](#), [Uy, Yi, and Zhang \(2012\)](#), [Michaels, Rauch, and Redding \(2012\)](#), [Lagakos and Waugh \(2013\)](#), [Gollin and Rogerson \(2014\)](#), [Gollin, Jedwab, and Vollrath \(2016\)](#), [Bustos, Garber, and Ponticelli \(2017\)](#), [McMillan, Rodrik, and Sepulveda \(2017\)](#), [Eckert and Peters \(2018\)](#), [Karádi and Koren \(2018\)](#), and [Sotelo \(2018\)](#). A related literature in development economics emphasizes structural transformation, as reviewed in [Syrquin \(1988\)](#) and [Foster and Rosenzweig \(2007\)](#). We make two main contributions relative to this line of work. First, whereas most existing macroeconomics research focuses on the aggregate economy, our analysis emphasizes the role of internal geography and transport costs in shaping structural transformation and the Balassa-Samuelson effect. Second, we use the natural experiment of Argentina’s late-19th-century integration into world markets and disaggregated data by sector and region over a long historical time period to provide quantitative evidence on the role of this structural transformation in the process of economic development.

Second, our paper is related to research on economic geography in the international trade literature, including [Hanson \(1996\)](#), [Fujita, Krugman, and Venables \(1999\)](#), [Davis and Weinstein \(2002\)](#), [Redding and Sturm \(2008\)](#), [Coşar and Fajgelbaum \(2012\)](#), [Davis and Dingel \(2012\)](#), [Allen and Arkolakis \(2014\)](#), [Ramondo, Rodriguez-Clare, and Saborio \(2016\)](#), [Redding \(2016\)](#), [Nagy \(2017\)](#), [Morten and Oliveira \(2017\)](#), and [Caliendo, Parro, Rossi-Hansberg, and Sarte \(2018\)](#). While most of this research focuses on the overall level of economic activity, our work highlights the role of internal geography in shaping the composition of economic activity between traded and non-traded sectors and across disaggregated goods within the traded sector.

Third, a growing empirical literature has examined the relationship between economic activity and transport infrastructure, including [Chandra and Thompson \(2000\)](#), [Baum-Snow \(2007\)](#), [Michaels \(2008\)](#), [Berlinski and Jaitman \(2011\)](#), [Banerjee, Duflo, and Qian \(2012\)](#), [Duranton and Turner \(2012\)](#), [Duranton, Morrow, and Turner \(2014\)](#), [Faber \(2014\)](#), [Coşar and Demir \(2016\)](#), [Donaldson and Hornbeck \(2016\)](#), [Storeygard \(2016\)](#), [Martincus and Cusolito \(2017\)](#),

Baum-Snow, Brandt, Henderson, Turner, and Zhang (2017), and Donaldson (2018), as reviewed in Redding and Turner (2015). Much of this empirical literature has followed a reduced-form approach and concentrated on the impact of reductions in internal transport costs. In contrast, we combine both reduced-form and structural approaches, and explore the role of internal trade frictions in shaping the impact of reductions in external trade barriers.¹

Fourth, we build on the historical literature on Argentine economic development, including Scobie (1971), Taylor (1992), Cortés Conde (1993), and Adelman (1994). Relative to this historical literature, we combine spatially-disaggregated data on economic activity by region and sector with a general equilibrium model to provide quantitative evidence on the relationship between trade, structural transformation and development.

The remainder of the paper is structured as follows. Section 2 provides some historical background. Section 3 introduces our data sources and definitions. Section 4 presents reduced-form evidence on the evolution of the spatial and sectoral distribution of economic activity over our sample period. Section 5 develops our theoretical model and uses its key prediction of the spatial Balassa-Samuelson effect to rationalize our reduced-form empirical findings. Section 6 undertakes a quantitative analysis of the model. We solve for unique values for the adjusted productivities in each sector that are sufficient statistics for the spatial and sectoral distribution of economic activity in the model. Section 7 reports counterfactuals, in which we examine the role of the construction of the railroad network in shaping the relationship between trade, structural transformation and development. Finally, Section 8 concludes. A separate web appendix collects together technical derivations and supplementary material.

2 Historical Background

The area that makes up present-day Argentina was first settled by Europeans in the early-sixteenth century. During this period of Spanish colonial rule, economic activity was centered around the silver mines in neighboring Bolivia.² Reflecting this orientation, official trade routes ran towards the Northwest through Panama, and trade was monopolized by Spanish merchants. In contrast, the Eastern coastal regions of Argentina, including Buenos Aires and the River Plate (*Río de la Plata*), were peripheral outposts for illegal trade with Brazil, Portugal and Britain.³

In response to the growth of this illegal trade and threats from encroaching Portuguese settlement, the Viceroyalty of the *Río de la Plata* was established in 1776 in Buenos Aires. With the decline in Spanish imperial power during the Napoleonic Wars, a local junta seized political power in 1810, which led to the first opening of direct trade with other foreign countries. After the failure of attempts to reassert Spanish colonial authority, full Argentinian independence was achieved in 1816. In the ensuing decades, there followed a gradual process of political consolidation, with the first national constitution agreed in 1853, the first constitutional government of all provinces meeting in 1862, and Buenos Aires absorbed into the federal structure of Argentina in 1880. Over these decades, successive military campaigns against native populations culminated in the “Conquest of the Desert” of 1879-80, which opened up the hinterland of Buenos Aires to economic development.⁴ Following the election of Julio Roca to the Presidency in 1880, liberal policies were pursued towards international flows of trade, capital and migrants, which were maintained until the

¹For empirical evidence on the role of domestic transport costs in shaping access to international markets, see also Atkin and Donaldson (2015) and Inter-American Development Bank (2013).

²For historical discussions of Argentine development, see for example Adelman (1994) and Scobie (1971).

³Early settlement patterns were heavily influenced by the availability of passive native Indian populations under the feudal *encomienda* system. Interior towns were established at Asunción (1537), Santiago del Estero (1553), Mendoza (1561), San Juan (1562) and San Miguel de Tucumán (1565). In contrast, the establishment of coastal towns lagged by several decades, including Santa Fe (1573), Buenos Aires (1580), Concepción del Bermejo (1585), and Corrientes (1588).

⁴Until 1880, there were periodic incursions from hostile native populations, as examined in Droller (2018).

outbreak of the First World War in 1914 drastically reduced these flows.⁵

During the late-19th century, a series of technological improvements centered on steam power dramatically reduced both international and domestic transport costs. The steam ship was first developed for river transportation in Europe and North America, with regular crossings of the North Atlantic by steam ship beginning in 1838. Following improvements in the speed, reliability and capacity of steam ships, international freight rates across the North Atlantic fell by around 1.5 percent per annum from around 1840 onwards, with a cumulative decline of around 70 percent points from 1840-1914, as documented in [North \(1958\)](#), [Harley \(1988\)](#) and [Pascali \(2017\)](#).⁶ In addition to this reduction in the overall level of transport costs, new technologies made possible trade in goods that were previously prohibitively costly to ship over long distances, as with the invention of the mechanical refrigeration of meat in 1861, which was first developed in Australia with a view to serving the U.K. market.

Steam technology also revolutionized domestic transportation through the construction of railroads. The first commercial use of mobile steam locomotives was to haul freight from mines at the Stockton and Darlington railway in the U.K. in 1825. The first railroad constructed in Argentina was the Buenos Aires Western Railway in 1857, with around 700 kilometers of track completed by 1869. From this point onwards, the railroad network expanded rapidly to grow to around 13,000 kilometers in 1895 and 30,000 kilometers in 1914.⁷ Whereas previously it had taken several months to transport goods by oxcart from Buenos Aires to an interior city such as Salta in the Northwest, the same journey could now be made in a matter of days (as discussed in [Scobie 1971](#)). Much of this railroad network was operated by private companies, which were predominantly British owned. However, these private companies operated alongside state-owned railroads, and the state influenced the development of the overall railroad network, through both land sales and the financing of railroads in rural areas.

With these reductions in international and domestic transport costs, Argentina experienced one of the largest recorded booms in international trade. Between 1869 and 1914, Argentina's real exports and imports increased by more than 500 and 200 percent respectively. In contrast to the Spanish colonial period, this international trade was now centered on the Eastern coastal regions. Following its emergence as the seat of political power, Buenos Aires rapidly developed into Argentina's main trade hub, even though its site was not particularly well suited for a port.⁸ Together, Buenos Aires and the three surrounding ports of La Plata (immediately adjacent to Buenos Aires), Rosario (directly upstream) and Bahía Blanca (developed as a satellite port to alleviate congestion in Buenos Aires port) account for more than 75 percent of the value of exports throughout our sample period.

As in many developing countries today, Argentina's exports were characterized by a high level of commodity specialization, with agriculture accounting for over 99 percent of export value throughout our sample period.⁹ Historically, agriculture in the hinterland of Buenos Aires had been based on cattle ranching on large estates (*estancias*), with sheep ranching becoming more important from the late-eighteenth century onwards. As the transport cost reduc-

⁵We end our sample period in 1914 to abstract from the effects of the First World War and subsequent more interventionist government policies, as discussed for example in [Taylor \(1992\)](#).

⁶These declines in freight rates were reflected in a convergence of commodity prices, with the gap between wheat prices in Liverpool and Chicago falling from 57.6 percent in 1870 to 17.8 percent in 1895 and 15.6 percent in 1913 ([Harley 1980](#)). See [O'Rourke and Williamson \(1999\)](#) for the seminal study of this increasing integration of the Atlantic economy.

⁷This rate of railroad expansion is comparable to that in the United States: between 1880 and 1913, railroad kilometers per 10,000 people rose from 9-42 in Argentina, compared to 29-44 in the United States.

⁸As noted in [Scobie \(1971\)](#), "Ironically, the sixteenth-century Spaniards, searching for an anchorage for their tiny ships, selected one of the poorest sites imaginable in terms of 19th-century sailing vessels and steamships" (p.95). As late as the 1880s, ships had to anchor several miles from shore in the open roads, until the construction of the Madero docks in 1897.

⁹As discussed in [Rocchi \(2008\)](#), the limited amount of domestic manufacturing activity involved either the processing of agricultural goods for export or was orientated towards the domestic consumer goods market.

tions in the late-19th-century occurred unevenly across goods, a major change in patterns of comparative advantage took place. Entirely new commodities now began to be exported, including in particular cereals and refrigerated and frozen beef and mutton. As a result, between 1869 and 1914, the export share of animal hides, bones and parts fell from 67 percent to 17 percent. In contrast, the export share of cereals rose from zero to more than 50 percent, and the export share of beef more than doubled from less than 5 percent to more than 10 percent.

This boom in agricultural production and exports was facilitated in part by large-scale international immigration, with Argentina's total population rising from 1.8 to 7.9 million between 1869 and 1914. Despite this substantial increase in labor supply, real wages and income per capita grew at average annual rates of 1.1 and 2.5 percent respectively over this period.¹⁰ This rapid economic growth was accompanied by structural transformation, as the share of agriculture in employment fell by around 7 percentage points between 1869 and 1914, and the share of the population living in towns and cities rose by about 20 percentage points. By 1914, Argentina was the eighth richest country in the world, with Buenos Aires accounting for around one fifth of its overall population.

3 Data

We construct a new spatially-disaggregated dataset for Argentina from 1869-1914. We combine six main sources of separate data.¹¹ First, we use the population censuses of 1869, 1895 and 1914 to measure the distribution of economic activity across spatially-disaggregated districts and provinces. We observe total population, rural population (which we associate with agricultural goods), urban population (which we associate with non-tradeables, including services and manufacturing for the local market), and geographical land area.¹² Across the three population censuses, there are changes in the boundaries of districts and provinces, both with the geographical expansion of Argentina's frontiers from 1869-1895 and the subdivision of districts from 1895-1914. Therefore, we construct time-invariant districts and provinces based on the boundaries in the 1895 census, using the maps and concordance in [Cacopardo \(1967\)](#). Our baseline sample consists of 380 districts and 23 provinces with constant boundaries.

Second, we use detailed data on the organization of economic activity within the agricultural sector from the 1895 and 1914 population censuses. We observe cultivated area for each district for twelve crops: Barley, Beans, Corn, Cotton, Flax, Peanuts, Potato, Sugar Cane, Vegetables, Tobacco, Wheat, and Wine. We also observe numbers of six types of livestock for each district: pure-breed cattle, mixed cattle, native cattle, pure-breed sheep, mixed sheep and native sheep. Finally, we have data on agricultural machinery use for 1895 and 1914, including up to twenty-one different categories of machines in 1914: Artesian Wells, Baling, Breakers, Cars, Carts and Wagons, Coaches, Combines, Dredges, Engines, Gleaners, Lawn Mowers, Other Mowers, Ploughs, Rakes, Rollers, Seeders, Shearers, Shellers, Threshers, Water machines and Wind machines.

Third, we have data on internal shipments by rail for 1895 and 1914 from the records of the Argentine railroads.¹³ We observe total quantities loaded at each railroad station for fifteen disaggregated products: Alfalfa, Cattle, Corn, Flax, Flour, Leather, Other Live Animals, Sand and Stone, Sheep, Sugar, Sugar Cane, Wheat, Wine, Wood and Wool.

¹⁰See [Taylor and Williamson \(1997\)](#). Argentina is the fastest-growing country in GDP per worker in their sample of 17 countries, which includes the richest countries of the period, such as the U.S., U.K., Australia and Canada.

¹¹See the data appendix for further discussion of the data definitions and sources.

¹²See [República Argentina \(1869, 1895, 1914\)](#). We use the definition of urban population from the population census, which corresponds to the population of all cities and towns. We find similar results with an alternative definition of urban population based on the population of cities with more than 2,000 inhabitants.

¹³See [Dirección General de Ferrocarriles \(1895, 1914\)](#). For further discussion of the historical development of the railroad network in Argentina, see for example [Lewis \(1983\)](#).

We first allocate railroad stations to districts using their latitude and longitude coordinates. We next compute the total quantity loaded by rail for each product for each of the districts in our sample.

Fourth, we combine this information on internal rail shipments with international trade data from customs records for 1870, 1895 and 1914.¹⁴ We observe the quantity and value of Argentine exports and imports by disaggregated product for each foreign country. Additionally, we observe exports by disaggregated product from each of the customs (ports) within Argentina. We use these data to document the high concentration of Argentine exports in agriculture and the large-scale changes in the volume and composition of these agricultural exports over time.

Fifth, we combine our production and trade data with a range of other geographical information. We constructed GIS shapefiles of the Argentinian railroad network in 1869, 1895 and 1914, the routes of navigable rivers, and Spanish colonial postal routes using the maps from [Randle \(1981\)](#). We use these GIS data to construct instruments for the railroad network, as discussed further below.

4 Reduced-Form Evidence

In this section, we provide reduced-form evidence on patterns of economic activity in late-19th-century Argentina, which guides the theoretical model that we develop below. In Section 4.1, we show the reorientation of economic activity away from the Spanish colonial cities of the North-West and towards the agricultural hinterland of Buenos Aires and its surrounding ports over our sample period.

In Section 4.2, we report regression evidence of a systematic gradient in overall population density and urbanization with distance from Argentina’s international trade hub in Buenos Aires and its surrounding ports. In Section 4.3, we present instrumental variables estimates of the impact of the railroad network on urban and rural population density, which are consistent with railroads reducing domestic transport costs and enabling interior regions to participate in international markets.

Finally, in Section 4.4, we show that the composition of economic activity within the agricultural sector is also systematically related to both distance from Argentina’s leading trade hub and railroad access.

4.1 Spatial Pattern of Economic Development

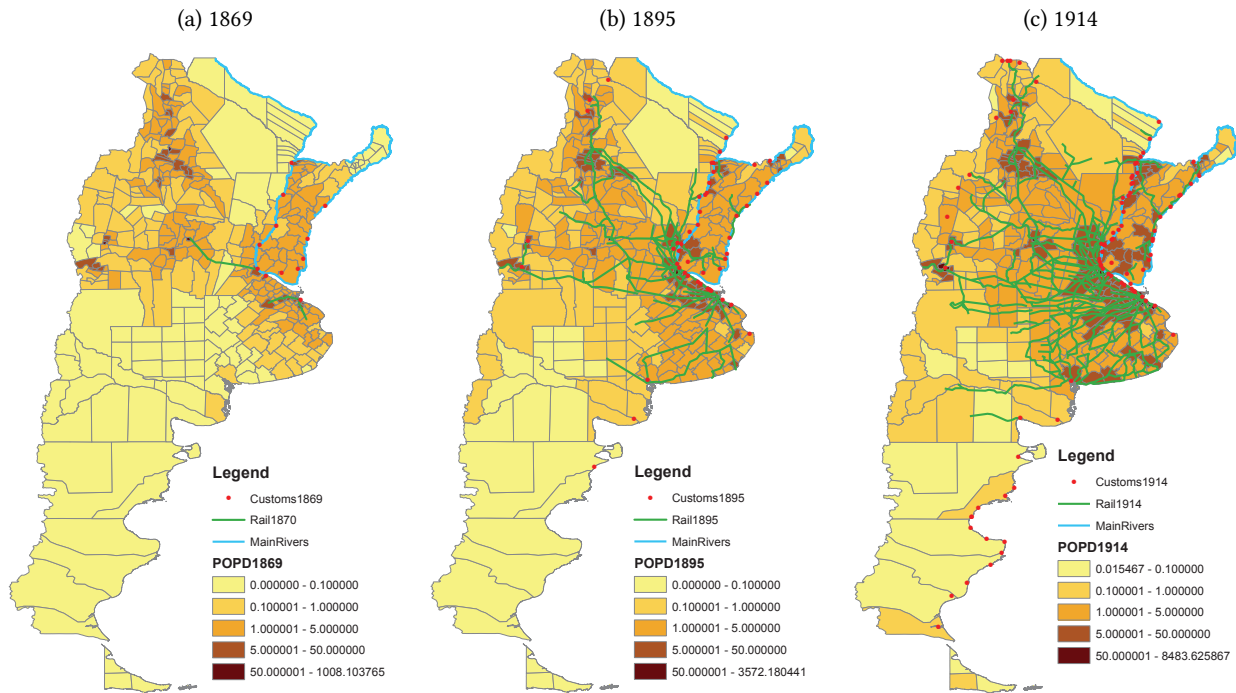
We begin by documenting the large-scale changes in the spatial distribution of economic activity within Argentina from 1869-1914. In Figure 1, we show the distribution of population density across our constant-boundary Argentinian districts in each of our census years. We divide the population density distribution in each year into the same five discrete cells, with darker shading indicating higher values. We show the railroad network in green, the main navigable rivers (the Paraná, Plate and Uruguay rivers) in blue, and the customs (ports) in red.

At the beginning of our sample period in 1869 (panel (a)), the main population concentrations were the Spanish colonial towns that served the mining region of Upper Peru (in the North-west) and the areas along the Paraná and Uruguay rivers and the River Plate estuary. Most of the remainder of Argentina was sparsely populated. The railroad network consisted of only 700 kilometers of track, including two lines radiating from Buenos Aires in the River Plate estuary and one line connecting the port of Rosario with the interior city of Córdoba.

¹⁴See [del Comercio Exterior \(1870\)](#) and [Compañía Sud-Americana de Billetes de Banco \(1895, 1914\)](#). We convert export and import values to U.S. dollars using the exchange rates from [Della Paolera \(1988\)](#) and [Bordo, Eichengreen, Klingebiel, and Peria \(2001\)](#). We convert nominal U.S. dollar values into 1869 prices using the GDP deflator from [Carter, Gartner, Haines, Olmstead, Sutch, and Wright \(2006\)](#).

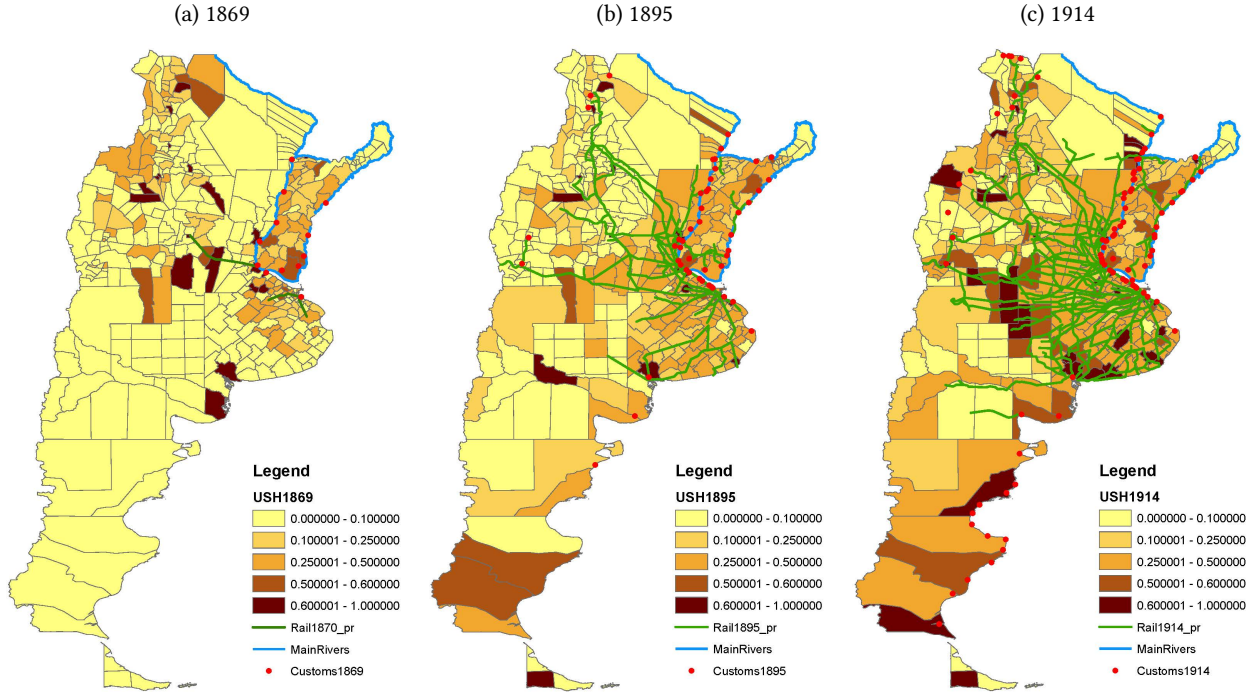
Between 1869 and 1895 (comparing panels (a) and (b)), we observe a substantial increase in overall population density, and a reorientation of the population density distribution towards the agricultural hinterland surrounding Buenos Aires and its neighboring ports. Over this period, there is a large-scale expansion in the railroad network to connect the agricultural hinterland with these ports and to link together the Spanish colonial towns. Between 1895 and 1914 (comparing panels (b) and (c)), we see a continuation of this pattern, with a further increase in population density, which now diffuses further inland from Buenos Aires and its surrounding ports. The railroad network now radiates further into the interior, with an increase in the density of lines serving the agricultural hinterland.

Figure 1: Spatial Distribution of Population Density from 1869-1914



In Figure 2, we find a similar pattern for urbanization, as measured by the share of the population living in towns and cities. In 1869 (shown in panel (a)), high urban population shares were concentrated around the Spanish colonial towns towards the North-West and along the main navigable rivers. Between each of the periods of 1869-95 and 1895-1914 (comparing panels (a) and (b)) and panels (b) and (c)), there is a general increase in the urban population share, which again radiates further inland from Buenos Aires and its neighboring ports. Therefore, we find that an increase in the overall level of economic activity (as reflected in population density) is accompanied by urbanization (a reallocation of economic activity from rural to urban areas). Additionally, with the expansion of economic activity into more peripheral locations, some remote areas with low population densities become dominated by few cities or towns, as reflected in high urban population shares.

Figure 2: Urban Population Share from 1869-1914



4.2 Gradients in Distance to Argentina's Trade Hub

We next use reduced-form regression specifications to establish three stylized facts about the distribution of economic activity within Argentina relative to its international trade hub: (i) We find a steep negative gradient in overall population density with distance from this trade hub; (ii) We show that this gradient is steeper for urban population density than for rural population density; (iii) We find that this gradient steepens over our sample period, as economic activity within Argentina reorients around this trade hub. We establish these stylized facts by estimating the following regression specification for each year separately:

$$\ln Y_t(\ell) = a_t + b_t \ln(\text{distport}(\ell)) + u_t(\ell), \quad (1)$$

where ℓ indexes districts and t corresponds to time; $\ln Y_t(\ell)$ is an economic outcome (e.g. log population density); $\ln(\text{distport}(\ell))$ is log geographical (Great Circle) distance to the nearest top-four port; $u_t(\ell)$ is a stochastic error.¹⁵ The key coefficient of interest is b_t , which captures the reduced-form elasticity of the economic outcome ($Y_t(\ell)$) with respect to distance from the nearest top-four port ($\text{distport}(\ell)$).

In Table 1, we report the estimation results. Panel A uses overall population density; Panel B examines urban population density; Panel C considers rural population density; and Panel D uses the urban population share. Columns (1), (2) and (3) report results for 1869, 1895 and 1914 respectively. Columns (4) and (5) report additional results for the last two years with controls. Already at the beginning of our sample period, we find that there is a substantial and statistically significant gradient in population density with respect to distance from Argentina's leading trade hub.

¹⁵ As discussed above, these top-four ports are Buenos Aires, La Plata, Rosario and Bahía Blanca, and account for more than 75 percent of export value throughout our sample period. In robustness checks, we find a similar pattern of results using Buenos Aires as the single leading port, or using Buenos Aires and the immediately adjacent port of La Plata, because the top-four ports are all clustered around Buenos Aires.

From Column (1) of Panel A, a doubling of distance to the closest top-four port (a 100 percent increase) is associated with around a 40 percent decline in population density. Comparing Column (1) of Panels B and C, we find that the urban elasticity of -0.65 is more than double the rural elasticity of -0.35, with this difference statistically significant at conventional levels. As a result, from Column (1) of Panel D, the urban population share decreases with remoteness from trade hubs with an elasticity of -0.05. Therefore, in the cross-section, we find that locations with better access to world markets are both more densely populated and more urbanized.¹⁶

All three of these gradients in population density steepen substantially over our sample period, with most of this change occurring in the first of our two sub-periods. From 1869-95, the gradient for overall population density more than doubles in absolute magnitude from -0.42 to -0.89 (Panel A, Columns (1) and (2)). We find that this steepening is smaller for urban population density (Panel B, Columns (1) and (2)) than for rural population density (Panel C, Columns (1) and (2)). Nevertheless, districts close to trade hubs have larger initial urban populations, which ensures that the gradient for the urban population share (Panel D, Columns (1) and (2)) more than doubles in absolute magnitude from -0.046 percent to -0.104 over this first sub-period. In the second of our two sub-periods, we find that all three of these gradients in population density are relatively flat, with none of the changes statistically significant at conventional levels. In Columns (4) and (5), we show that we find a similar pattern of results for 1895 and 1914 if we control for historical patterns of economic development using the 1869 values of population density or the urban population share. In this specification, the changes from 1895-1914 become larger and statistically significant, but remain smaller than those from 1869-1895 in Columns (1) and (2). Therefore, over time, we find that locations with better access to world markets become more densely populated and more urbanized.¹⁷

Overall, we find strong evidence that both the level and composition of economic activity within Argentina are systematically related to internal geography.

¹⁶ Although the natural experiment of Argentina's late-19th-century integration into world markets provides an attractive empirical setting, this property that locations with better access to world markets are both more densely populated and more urbanized is also found in other settings, such as in the Belt and Road Initiative (BRI) in Central Asia, as examined in the subsequent work by [Lall and Lebrand \(2018\)](#).

¹⁷ We find a similar pattern of results if we restrict the specifications in Table 1 to the sample of districts for which data exist for all three years. For example, comparing the 1869 and 1914 results from the specifications in Columns (1) and (3) for this sample of districts, we find increases in absolute magnitude of the gradients for population density (from -0.414 to -0.859), urban population density (from -0.582 to -1.115), rural population density (from -0.322 to -0.634), and the urban population share (from -0.046 to -0.105).

Table 1: Population Density and Distance to Top-Four Ports

	(1)	(2)	(3)	(4)	(5)
Panel A	Log Population Density 1869	Log Population Density 1895	Log Population Density 1914	Log Population Density 1895	Log Population Density 1914
Log Distance Top-Four Port	-0.415*** (0.089)	-0.889*** (0.091)	-0.957*** (0.088)	-0.364*** (0.040)	-0.526*** (0.054)
Log Population Density 1869	-	-	-	0.819*** (0.038)	0.806*** (0.046)
Observations	298	379	380	298	298
R-squared	0.09	0.20	0.26	0.83	0.77
Panel B	Log Urban Population Density 1869	Log Urban Population Density 1895	Log Urban Population Density 1914	Log Urban Population Density 1895	Log Urban Population Density 1914
Log Distance Top-Four Port	-0.650*** (0.126)	-1.174*** (0.104)	-1.133*** (0.106)	-0.494*** (0.072)	-0.527*** (0.077)
Log Population Density 1869	-	-	-	0.935*** (0.046)	0.965*** (0.048)
Observations	173	280	330	156	162
R-squared	0.17	0.35	0.30	0.86	0.83
Panel C	Log Rural Population Density 1869	Log Rural Population Density 1895	Log Rural Population Density 1914	Log Rural Population Density 1895	Log Rural Population Density 1914
Log Distance Top-Four Port	-0.345*** (0.071)	-0.724*** (0.083)	-0.758*** (0.070)	-0.269*** (0.036)	-0.418*** (0.043)
Log Population Density 1869	-	-	-	0.774*** (0.037)	0.692*** (0.039)
Observations	295	376	377	294	294
R-squared	0.07	0.15	0.22	0.77	0.70
Panel D	Urban Population Share 1869	Urban Population Share 1895	Urban Population Share 1914	Urban Population Share 1895	Urban Population Share 1914
Log Distance Top-Four Port	-0.046*** (0.014)	-0.104*** (0.010)	-0.100*** (0.011)	-0.094*** (0.010)	-0.087*** (0.010)
Urban Population Share 1869	-	-	-	0.464*** (0.061)	0.385*** (0.058)
Observations	298	379	380	298	298
R-squared	0.04	0.25	0.18	0.59	0.39

Notes: Observations for each cell are a cross-section of Argentinian districts in the respective year. Log population density is the log of the population per unit of land area. Urban population is the population living in cities and towns, as measured by the population census. Rural population is total population minus urban population. Urban Population share is the share of population living in cities and towns. Some urban and rural population variables are missing or zero in some years, as reflected in the smaller number of observations in the specifications using urban and rural population data. Distance Top-Four Port is the geographic (Great Circle) distance from the centroid of each district to the nearest top-four port (Buenos Aires, Rosario, La Plata and Bahia Blanca). Heteroskedasticity robust standard errors in parentheses. *** denotes significance at the 1 percent level; ** denotes significance at the 5 percent level; * denotes significance at the 10 percent level.

4.3 Impact of the Railroad Network

We next provide regression evidence on the role of the railroad network in enabling interior regions to experience increases in urban and rural population density. A key empirical challenge is that the placement of railroads could be non-random and targeted towards interior regions that would have experienced different trends in urban and rural population density, even in the absence of these railroads. On the one hand, much of the railroad network was operated by private-sector companies, whose search for profits could have led them to select regions that otherwise would have grown more rapidly. On the other hand, the Argentine state promoted the development of railways in rural areas that were unattractive to private-sector companies, which could have targeted locations that otherwise would have grown less rapidly. To address these concerns about non-random placement, we construct two instrumental variables that

exploit quite different sources of variation, one based on the historical location of Spanish colonial cities, and the other based on Argentina's late-19th-century integration into the world economy.

Our first instrument exploits the fact that the top-four ports are all clustered around Buenos Aires, which had already developed into Argentina's trade hub in the aftermath of the Napoleonic Wars, before the invention of the railroad in 1825. Once railroads were invented, we exploit the fact that interior regions were likely to be connected to this pre-existing trade hub, regardless of the economic characteristics of those interior regions. Therefore, our first instrument mechanically predicts the railroad network based on constructing least-cost paths between the centroid of each district and the top-four ports. In particular, we discretize Argentina into a raster of grid points. Starting from the grid point closest to the centroid of a given district, we construct least-cost paths across this raster of grid points to each of the top-four ports, assuming an equal cost of travel across each grid point.¹⁸ We then repeat this exercise for all districts. Finally, for each district, we compute the fraction of grid points within its boundaries that lie along one or more of these least-cost paths from a centroid to Argentina's trade hub.

Crucially, this instrument uses no information about the economic characteristics of districts, and hence cannot be influenced by some districts being economically more desirable destinations than others. Instead, this instrument predicts the structure of the railway network based purely on mechanically connecting all districts to the pre-existing trade hub. To address the concern that larger districts are other things equal more likely to be along these least-cost paths to the pre-existing trade hub, we control separately for log district land area. To address the concern that areas close to top-four ports could have different rates of economic growth for reasons unrelated to the railroad network, we also control separately for distance from the centroid of each district to the nearest top-four port. Therefore, our estimates exploit variation in the frequency with which a district lies along a least-cost path to Argentina's trade hub, conditional on a given distance from that trade hub. Finally, to control for potential heterogeneity in initial levels of economic development, we control separately for initial population in 1869. Conditional on these controls, our first instrument assumes that there is no direct effect on economic activity of frequently lying along a least-cost path to Argentina's trade hub, other than through the probability with which a district is connected to the railroad.

Our second instrument uses historical exploration and trade routes following [Duranton and Turner \(2012\)](#) and [Duranton, Morrow, and Turner \(2014\)](#). We use the fact that economic activity in the Spanish colonial period was orientated in a very different way from that in the late-19th-century export boom. In particular, official trade routes ran towards the North-West through Panama, instead of towards the Eastern coastal areas around Buenos Aires. Despite this very different orientation of economic activity, once existing population centers had formed, they were likely to be connected to the railroad after it had been invented. Hence, locations along the route between these historical centers were also likely to be connected. To implement this idea, we georeference a map of Spanish colonial postal routes from the eighteenth century from [Randle \(1981\)](#). For each district, we construct our instrument as the length of colonial postal routes within its boundaries as a share of the length of these routes for Argentina as a whole. We expect this instrument to have power in predicting the railroad network, because paths that are convenient for colonial postal routes using horses are also likely to be convenient for the construction of railroads. To address the concern that districts along colonial postal routes could differ in historical levels of economic activity, access to international markets or geographical size, we again control separately for the initial level of economic activity in

¹⁸As a robustness check, we constructed a version of this instrument based on constructing least-cost paths from all districts to Buenos Aires alone, and find a similar pattern of results, because the top-four ports are clustered around Buenos Aires, as discussed above.

1869, distance to the nearest top-four port, and land area. After conditioning on these controls, our second instrument assumes that there is no direct effect of lying along Spanish colonial postal routes on subsequent late-19th-century economic growth, other through the probability with which a district is connected to the railroad.

Importantly, our two instruments exploit quite different sources of variation. Our first instrument is based on connecting the interior to the late-19th century trade hub centered on the Buenos Aires coastal region. In contrast, our second instrument uses postal routes between the Spanish colonial cities that were orientated around trade routes through the North-Western interior regions towards Panama. Therefore, we can use these two different sources of variation to provide a check on our identifying assumptions, by reporting Hanson-Sargan overidentification tests and the results of specifications using only one of the two instruments. If we find a similar pattern of results using each of the two instruments separately, this implies either that both instruments are valid, or that both are invalid and there exists an implausible correlation structure, such that the error term has a similar correlation with these two quite different sources of variation.

In our instrumental variables estimation, we consider the following second-stage regression for long-differenced population growth (either overall, urban or rural) over our sample period:

$$\Delta \ln Y_{1914-1869}(\ell) = a + b \ln(\text{distport}(\ell)) + c(\text{sharerail}_{1914}(\ell)) + d_1 \ln(\text{area}(\ell)) + d_2 \ln Y_{1869}(\ell) + u(\ell), \quad (2)$$

where ℓ again indexes districts; $\Delta \ln Y_{1914-1869}(\ell)$ is log population growth from 1869-1914 (either overall, urban or rural); $\ln(\text{distport}(\ell))$ is log geographical (Great Circle) distance to the closest top-four port; $\text{sharerail}_{1914}(\ell)$ is the length of railroads in each district in 1914 as a percentage of this length for Argentina as a whole; this percentage railroad share in 1914 captures the expansion of the railroad network from 1869-1914, because this network was of negligible length in 1869;¹⁹ we measure access to railroads using a percentage share rather than a log length to permit the inclusion of zero values; $\ln(\text{area}(\ell))$ is the log geographical area of each district; $\ln Y_{1869}(\ell)$ is initial log population in 1869 (either overall, urban or rural); and $u(\ell)$ is a stochastic error.

This second-stage regression specification (2) controls for a fixed effect in the level of log population for each district, which is differenced out when we take long differences. We thus allow for time-invariant unobserved heterogeneity in location characteristics that affects population levels in each year. The constant a captures any common time effect that affects population growth across all Argentinian districts from 1869-14, such as common macro shocks. The corresponding first-stage regression for a district's share of the railroad network in 1914 is given by:

$$\begin{aligned} \text{sharerail}_{1914}(\ell) = & e + f_1(\text{routeport}(\ell)) + f_2(\text{sharepost}(\ell)) + g_1 \ln(\text{distport}(\ell)) \\ & + g_2 \ln(\text{area}(\ell)) + g_3 \ln Y_{1869}(\ell) + h(\ell), \end{aligned} \quad (3)$$

where $\text{routeport}(\ell)$ is the frequency with which a district lies along the least-cost path from all districts to a top-four port (our first instrument); $\text{sharepost}(\ell)$ is a district's share of Spanish colonial postal routes (our second instrument); $e(\ell)$ is a stochastic error; and the remaining variables are defined above.

Table 2 presents the results of estimating the second-stage regression (2) for population growth. In Columns (1)-(3), we report the OLS estimates. In Column (1), we include only distance from the nearest top-four port and log land area. Consistent with the steepening of the gradient in population density in Table 1, we find that districts

¹⁹Given the negligible railroad network in 1869, we find a similar results if we instead use the change in each district's percentage share of the railroad network between 1869 and 1914.

further from Argentina's trade hub experience statistically significantly slower population growth, with an elasticity of population growth with respect to distance from the nearest top-four port of around -0.45. In Column (2), we augment this specification with each district's percentage share of the total length of railroads in 1914, where recall that this 1914 value captures the expansion of the railroad network from 1869-1914, because this network was of negligible length in 1869. We find a strong positive correlation between population growth and the expansion of the railroad network, which is statistically significant at conventional levels. The estimated magnitude of the coefficient is also economically large. The estimates in Column (2) imply that a one standard deviation increase in the share of the railroad network is associated with 0.53 standard deviation increase in population growth, where the standard deviations of long-differenced population growth and our railroad variable are 0.98 and 0.35 respectively. In Column (3), we further augment this specification with initial log population density in 1869 and find a similar pattern of results, confirming that the correlation between population growth and railroad expansion is robust to controlling for historical patterns of settlement.

In Column (4), we report the two-stage least squares estimates of the specification from Column (3), using both our port and colonial post instruments. Consistent with a causal effect of the expansion of the railroad network on population growth, we find that the estimated railroad coefficient remains positive and statistically significant. The IV estimate is marginally larger than the OLS estimate, but this difference is not statistically significant, which could reflect the two offsetting effects discussed above. On the one hand, private-sector railroad companies have an incentive to target regions that otherwise would have grown more rapidly, which suggests that the OLS estimate should be greater than the IV estimate. On the other hand, the public-sector promotion of regions that otherwise would have grown more slowly implies that the OLS estimate should be smaller than the IV estimate. In principle, either one of these effects could dominate, and the fact that the IV and OLS estimates are close to one another is consistent with the idea that they approximately offset one another in our empirical setting. We find that the instruments have power in the first-stage regression, with the F-statistic for the significance of the instruments in the first stage equal to 34.96 (well above the conventional threshold of 10), as reported at the bottom of the column. In a Hansen-Sargan overidentification test, we are unable to reject the null hypothesis of the model's overidentifying restrictions (p -value = 0.72), as also reported at the bottom of the column. Therefore, assuming that one of the instruments is valid, we are unable to reject the null hypothesis that the other instrument only matters for population growth through railroad access.

Table 2: Population Growth and Railroad Access

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log Population Growth 1869-1914	Log Population Growth 1869-1914	Log Population Growth 1869-1914	Log Population Growth 1869-1914	Log Population Growth 1869-1914	Log Population Growth 1869-1914	Log Urban Population Growth 1869-1914	Log Rural Population Growth 1869-1914
log Distance Top-Four Port	-0.465*** (0.053)	-0.296*** (0.050)	-0.311*** (0.048)	-0.290*** (0.051)	-0.308*** (0.074)	-0.287*** (0.051)	-0.422*** (0.081)	-0.153*** (0.046)
Log Land Area	0.049 (0.048)	-0.160*** (0.037)	-0.147*** (0.037)	-0.173*** (0.045)	-0.151** (0.075)	-0.177*** (0.046)	-0.225*** (0.068)	-0.066 (0.052)
Share Rail Length 1914	-	1.466*** (0.141)	1.418*** (0.101)	1.596*** (0.184)	1.444*** (0.479)	1.623*** (0.194)	1.188*** (0.388)	1.599*** (0.236)
Log Population Density 1869	-	-	-0.412*** (0.053)	-0.409*** (0.052)	-0.411*** (0.052)	-0.409*** (0.052)	-	-
Log Urban Population Density 1869	-	-	-	-	-	-	-0.255*** (0.067)	-
Log Rural Population Density 1869	-	-	-	-	-	-	-	-0.492*** (0.049)
Instruments	-	-	-	Both	Port	Colonial Post	Both	Both
Observations	298	298	298	298	298	298	162	294
R-squared	0.24	0.45	0.59	-	-	-	-	-
First-stage F-Statistic	-	-	-	34.96	14.19	65.29	34.96	34.96
Overidentification test (p-value)	-	-	-	0.7164	-	-	0.6738	0.1673

Notes: Observations are a cross-section of Argentinian districts. Distance Top-Four Port is the geographic (Great Circle) distance from the centroid of each district to the nearest top-four port (Buenos Aires, Rosario, La Plata and Bahia Blanca). Log population density is the log of the population per unit of land area. Share rail length is the length of railroads in each district as a percentage of this length for Argentina as a whole. Port instrument is the percentage of grid points within each district that lie on the least-cost routes from the centroids of all Argentinian districts to the top-four ports. Colonial post is the length of Spanish colonial postal routes in each district as a percentage of this length for Argentina as a whole. First-stage F-statistic is a test of the statistical significance of the instruments in the first-stage regression. Overidentification test is a Hansen-Sargan test of the model's overidentifying restrictions. Heteroskedasticity robust standard errors in parentheses. *** denotes significance at the 1 percent level; ** denotes significance at the 5 percent level; * denotes significance at the 10 percent level.

As a further specification check, Columns (5) and (6) report exactly-identified specifications, in which we use each instrument separately. We find that each instrument has power, with a first-stage F-statistic in each case above the conventional threshold of 10. In both specifications, the IV estimates are marginally larger than the OLS estimate in Column (3), although the differences are again not statistically significant. This similarity of the estimates using instruments that exploit quite different sources of variation again provides support for our identifying assumptions. If only one of the instruments were invalid, we would expect to find a quite different pattern of results using that instrument. Hence, to explain the similarity of the results using all combinations of the instruments, we again need either both instruments to be valid or an improbable pattern of correlation to exist between the instruments and the error term in the second-stage regression. Finally, we interpret the fact that all three IV estimates are close to the OLS estimate as supporting the idea that, conditional on our controls, the expansion of the railroad network within Argentina was indeed mainly driven by connecting interior regions with the top-four ports, and connecting existing colonial centers, rather than targeting interior regions that would have grown more rapidly for other reasons, even in the absence of the railroad.

In Columns (7) and (8), we report analogous IV specifications for urban and rural population growth respectively. In the interests of brevity, we concentrate on our baseline specification using both instruments. We find a similar pattern of results as for overall population growth in Column (4). Expansions of the railroad network predicted by our instruments raise both urban and rural population growth. Although we find a larger estimated coefficient for rural than for urban population growth, the difference between these two coefficients is not statistically significant at conventional levels. Again the instruments have power in the first-stage regression and we pass the Hansen-Sargan test of the model's overidentifying restrictions.

Taken together, the results of this section are consistent with the idea that the railroad network played an important role in enabling interior regions to participate in Argentina's rapid 19th-century economic development. For a given distance from Argentina's trade hub, we find that an expansion in the railroad predicted by our instruments raises both urban and rural population density.

4.4 Structural Transformation Within the Agricultural Sector

We have so far established that both overall population density and the composition of the population between urban and rural areas are systematically related to internal geography. We now provide evidence of similar systematic differences in the composition of economic activities within the agricultural sector. Consistent with transport costs playing a role in shaping comparative advantage, as in [von Thünen \(1826\)](#), we find that the new export crops of cereals and refrigerated and frozen meat are concentrated close to Argentina's trade hub, and railroad access predicted by our instruments increases concentration in these new export crops.

We begin by establishing the dramatic expansion and transformation in the agricultural sector over our sample period. Total cultivated area increases from 40,000 to 129,000 kilometers squared between 1895 and 1914, with 74,000 kilometers squared of this increase made up of the new cereal crops of Barley, Corn and Wheat. As part of this large-scale expansion in cereals production, the total number of agricultural machines reported in the data rises from 15,000 to 56,000. Entirely new types of machines are recorded for the first time in 1914, such as combine harvesters, seeders and shellers, all of which are used for cereals production. Consistent with most cereals production being shipped outside the district where it was produced for the export market, we observe a large-scale expansion in rail shipments

of Corn (from 93,000 to 254,000 tons) and Wheat (from 108,000 to 192,000 tons) between 1895 and 1914. In contrast, the total number of reported cattle and sheep falls from 93 to 69 million over the same period. Additionally, there is a shift in livestock composition away from native breeds most suited for tanning and leather (from 35-26 percent), towards pure and mixed breeds better suited for refrigerated and frozen meat (from 65-74 percent).

In Table 3, we provide evidence on patterns of agricultural production across districts and over time. Each cell of the table corresponds to a separate regression, with the dependent variable reported in the rows of the table, and the independent variable given in the columns of the table. In Panel A, crop cultivated area is measured as a percentage of total land area for each district. In Panel B, agricultural machinery is measured as the number of each type of machine in each district as a percentage of the total for this type of machine for Argentina as a whole. In Panel C, livestock are measured as the number of each type for livestock in each district as a percentage of the total for this type of livestock for Argentina as a whole. In Panel D, railroad shipments are measured as the quantity of each good shipped from stations in each district as a percentage of the total for this same good for Argentina as a whole. In each case, we define the variable in terms of percentage shares to permit the inclusion of zero values.

We examine the spatial distribution of agricultural activities relative to Argentina's trade hub in the specifications reported in Columns (1)-(2) and (5)-(6). In all specifications, we include log land area as a control to capture the fact that larger districts are likely to have greater percentage shares of an agricultural activity, other things equal. As apparent from the table, we find substantial differences in the extent to which agricultural activities are concentrated close to Argentina's trade hub, which are typically statistically significant at conventional critical values. In Panel A, we find that the new export cereal crops (in particular Corn and Wheat) have particularly steep negative gradients in economic activity with distance from Argentina's trade hub. Both of these gradients steepen between 1895 and 1914, as the agricultural hinterland surrounding Buenos Aires is developed for export cereal production. Consistent with this concentration of cereal production in locations with good access to world markets, Panel B shows negative and statistically significant gradients for all types of agricultural machines, most of which are intensively used for cereals production (such as mowers, ploughs, seeders and combines). In contrast, as the agricultural hinterland surrounding Buenos Aires is developed for cereal production, we find a flattening of the gradient of all types of livestock with respect to distance from Argentina's trade hub in Panel C. This flattening is particularly marked for native cattle and sheep, which are disproportionately used for the traditional export goods of tanned and leather hides. Finally, consistent with most cereal production being shipped outside districts to the export market, we observe some of the steepest negative gradients in railway shipments with respect to distance from Argentina's trade hub for cereals and cereal products (such as Corn, Flour and Wheat) in Panel D.

Table 3: Agricultural Production, Distance to Top-Four Ports and Railroad Access

	(1)	(2)	(3)	(4)		(5)	(6)	(7)	(8)
	Log Distance Top-Four Port 1895	Log Distance Top-Four Port 1914	Rail Share 1895	Rail Share 1914		Log Distance Top-Four Port 1895	Log Distance Top-Four Port 1914	Rail Share 1895	Rail Share 1914
Panel A : Cultivated Area					Panel C : Livestock				
Corn	-1.3149***	-3.7057***	0.3499	5.1995**	Native Cattle	-0.0815***	0.0469***	0.3403*	0.5075***
Wheat	-1.6948***	-2.3140***	5.9179***	10.0983***	Mixed Cattle	-0.2115***	-0.1765***	-0.6117**	0.1969
Flax	-0.3661***	-1.3247***	0.7586*	2.8820**	Pure-breed Cattle	-0.2148***	-0.1968***	-0.3417	-0.0006
Vegetables	-0.1466***	-0.1158***	-0.2757	0.0628	Native Sheep	-0.1582***	0.0552**	0.1043	0.4838*
Barley	-0.0862***	-0.09291**	-0.0130	-0.1182	Mixed Sheep	-0.2219***	-0.1396***	-0.8116***	-0.6027***
Potato	-0.0598**	-0.0967	-0.1294*	-0.0952	Pure-breed Sheep	-0.2491***	-0.1588***	-0.4560*	-0.1629
Cotton	0.0007***	0.0015*	0.0005	0.0000	Panel D : Railroad Shipments				
Peanuts	-0.0017	0.0017	-0.0002	0.0322	Corn	-0.3018***	-0.3518***	0.5265*	1.5620**
Beans	-0.0719*	0.0023	-0.2486	0.0046	Flour	-0.2807***	-0.3151***	0.8920**	1.1587**
Tobacco	0.0247***	0.0306	-0.0499	-0.0036	Flax	-0.2986***	-0.2860***	1.6437***	1.6918***
Sugar Cane	0.1908**	0.2927**	0.6527	0.4425	Wool	-0.2582***	-0.2842**	-0.5757	-0.3982*
Wine	0.2456**	0.6378**	-0.2413	1.9266	Wheat	-0.2973***	-0.2657***	2.0781***	1.3773**
Panel B : Machinery					Leather	-0.1755***	-0.1669***	-0.0710	0.6567***
Wind Machines	-0.3378***	-0.2050***	-0.7465	0.8483**	Alfalfa	-0.3222***	-0.1640***	1.8546**	1.0319*
Water Machines	-0.2995***	-0.1736***	1.2557**	0.9814***	Sugar	0.2194*	-0.1640	0.5629	0.1299
Mowers	-0.2771***	-0.2018***	1.0514***	0.6082***	Sand and Stone	0.0333	-0.0657	1.2734	1.0617
Threshers	-0.2635***	-0.2249***	1.7420***	1.0365***	Wood	0.0189	-0.0303	0.9842*	0.7729***
Rakes	-0.2146***	-0.1940***	0.8873***	1.1325***	Wine	0.2555	0.0451	-0.8947	2.0400**
Ploughs	-0.0937***	-0.1184***	0.6054***	0.7724***	Sugar Cane	0.2679*	0.2656**	-1.2842	0.0590
Combines		-0.2921***		-0.3024**					
Seeders		-0.2453***		1.2172***					
Dredges		-0.2361***		0.8943***					
Gleaners		-0.2332***		1.6459***					
Cars		-0.2281***		0.6329***					
Lawn Mower		-0.2267***		1.0165***					
Rollers		-0.2117***		1.6520**					
Coaches		-0.2094***		0.8416***					
Breakers		-0.2014***		-0.1818					
Carts & Wagons		-0.1947***		0.7937***					
Engines		-0.1944***		0.8030***					
Artesian Wells		-0.1824***		0.5254**					
Shellers		-0.1699***		0.2583**					
Baling		-0.1109***		0.5083***					
Shearers		-0.0948***		-0.6150***					

Notes: Each cell of the table corresponds to a separate regression. Observations are a cross-section of Argentinian districts in the respective year. Rows correspond to the dependent variable. Cultivated area is crop cultivated area divided by total land area for each district. Agricultural machinery is the number of each type of agricultural machine for each district as a share of the total for that type for Argentina as a whole. Livestock is the number of each type of livestock for each district as a share of the total for that type for Argentina as a whole. Railroad shipments is the quantity shipped of each good from railroad stations in each district as a share of the total for that good for Argentina as a whole. Columns correspond to independent variables. Distance Top-Four Port is the log geographic (Great Circle) distance from the centroid of each district to the nearest top-four port (Buenos Aires, Rosario, La Plata and Bahia Blanca). Rail share is the length of railroads in each district as a percentage of this length for Argentina as a whole. We instrument the rail share using our port and colonial post instruments. The port instrument is the percentage of grid points within each district that lie on the least-cost routes from the centroids of all Argentinian districts to the top-four ports. The colonial post instrument is the length of Spanish colonial postal routes in each district as a percentage of this length for Argentina as a whole. In Columns (1), (2), (5) and (6), we control for the log land area of each district. In Columns (3), (4), (7) and (8), we control for the log distance to the nearest top-four port, log land area and the log population density in 1869. Statistical significance based on heteroskedasticity robust standard errors. *** denotes significance at the 1 percent level; ** denotes significance at the 5 percent level; * denotes significance at the 10 percent level.

We next examine the relationship between agricultural production and railroad access, conditional on a given distance from Argentina's trade hub. In Columns (3)-(4) and (7)-(8), we regress each district's percentage share of an agricultural activity on its percentage share of the length of the railroad network in 1895 and 1914. As in our earlier specifications for population density, we include controls for log land area, log distance to the nearest top-four port and log initial population density in 1869. We also instrument railroad access with our two instruments based on the frequency with which a district lies along a least-cost path to a top-four port and the frequency with which it lies along Spanish colonial postal routes. As shown in the table, we find substantial differences across agricultural activities

in the impact of railroad access predicted by our instruments. In Panel A, we find some of the largest positive and statistically significant effects for the new cereal crops of Corn and Wheat. In Panel B, we find positive and statistically significant effects for almost all categories of agricultural machines, consistent with these machines being intensively used for cereals production. In Panel C, we find positive and statistically significant effects for some categories of cattle, and negative and statistically significant effects for some categories of sheep, which is consistent with sheep farming being relatively more profitable in more remote locations. In Panel D, we find some of the largest positive and statistically significant effects for cereals and cereal products (such as Corn, Flour and Wheat), again consistent with access to world markets being relatively more important for these agricultural goods.

Taken together, these results for both distance from Argentina's trade hub and railroad access confirm the role of internal geography in shaping the composition of economic activity within the agricultural sector and the shift towards new sources of comparative advantage in cereals and refrigerated and frozen meat.

5 Theoretical Model

We now develop the theoretical framework that we use to rationalize these empirical findings.²⁰ The key new insight of the model is to establish an interaction between structural transformation across sectors and internal trade costs across regions. In particular, using general neoclassical assumptions on production, we demonstrate a *spatial Balassa-Samuelson effect*, such that regions with good access to world markets have higher population densities, urban population shares, relative prices of non-traded goods, and land prices relative to wages. Therefore, reductions in external transport costs induce structural transformation from agriculture to non-agriculture, and lead to changes in the composition of economic activities within the agricultural sector. Similarly, reductions in internal transport costs, through for example the expansion of the railroad network, increase the ability of interior regions to participate in international markets and undergo this process of structural transformation.

The world economy comprises three sectors: manufacturing (M), agriculture (A), and non-tradables (N). We consider a country that consists of a set of locations $\ell \in \mathcal{L}$. Some of these locations $\ell \in \mathcal{L}_C \subset \mathcal{L}$ are coastal and have direct access to world markets at prices $(\{P_g^*\}_{g=1}^G, P_M^*)$ that depend on external transport costs.²¹ Other locations $\ell \in \mathcal{L}_I \subset \mathcal{L}$ are interior regions that are connected to coastal locations through an internal transport network. We denote the trade cost between any pair of locations $(\ell, \ell') \in \mathcal{L}$ for good g by $\delta_g(\ell, \ell')$. Motivated by the overwhelming concentration of Argentinian exports in agriculture, we assume that all locations within Argentina have a comparative advantage in agriculture. To rationalize the observed differences in the composition of agricultural production across these locations, we assume that this agricultural sector consists of a discrete number of disaggregated goods indexed by $g = 1, \dots, G$. Each location ℓ has a land area $L(\ell)$ and a continuum of land plots $j \in [0, L(\ell)]$ that are heterogeneous in terms of their productivity for these disaggregated agricultural goods $g = 1, \dots, G$.²² Unless otherwise indicated, we suppress time subscripts from now onwards to simplify notation, but we take it as understood that all location-specific characteristics (such as productivities and bilateral trade costs) can change over time.

²⁰ A web-based technical appendix contains the derivation of the results and the proofs of the propositions in this section.

²¹ For most of our analysis, we are not required to take a stand on whether these prices at the port $(\{P_g^*\}_{g=1}^G, P_M^*)$ are exogenous or endogenous. When we undertake counterfactuals, we assume that Argentina is a small open economy that faces exogenous prices at the port, which is a reasonable approximation in our empirical setting. For example, for the major export product of wheat, [Bennett \(1933\)](#) estimates that world production in 1895 (1914) was 2,731 (3,618) bushels, which compares with Argentinian production of 46.4 (169.2) bushels.

²² While we make the conventional neoclassical assumption that units of any given good are homogeneous across locations, it is straightforward to introduce Armington differentiation by location of origin, and the spatial Balassa-Samuelson forces in the model continue to apply.

5.1 Preferences and Endowments

Preferences are defined over consumption of *traded* and *non-traded* goods and are assumed to take the constant elasticity of substitution (CES) form:

$$u(\ell) = \left[\beta_T c_T(\ell)^{\frac{\sigma-1}{\sigma}} + (1 - \beta_T) c_N(\ell)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \quad (4)$$

where $c_T(\ell)$ and $c_N(\ell)$ respectively denote consumption of the traded and non-traded goods. Following the literature on structural transformation in macroeconomics, we assume inelastic demand between these two sectors ($0 < \sigma < 1$). Tradables consumption is in turn defined over consumption of a composite manufacturing good and the set of agricultural goods $g = 1, \dots, G$ with the following homothetic price index:

$$E_T(\ell) = E_T(\{P_g(\ell)\}_{g=1}^G, P_M(\ell)), \quad (5)$$

where $P_g(\ell)$ is the price of agricultural good g in location ℓ and $P_M(\ell)$ is the corresponding price of the composite manufacturing good.

Each worker is endowed with one unit of labor that is supplied inelastically with zero disutility. We assume that workers are perfectly mobile across locations and hence arbitrage away real wage differences.²³ The labor market clearing condition for the economy as a whole can be written as:

$$\sum_{\ell \in \mathcal{L}} L(\ell) n(\ell) = N, \quad (6)$$

where $n(\ell) = N(\ell) / L(\ell)$ is population density at location ℓ ; and N is the economy's total population. For most of our quantitative analysis, we are not required to take a stand on the extent to which the economy's population is endogenous or exogenous, because we read this variable directly from the data. When we undertake counterfactuals, we report results under two alternative assumptions about international migration: (i) free international migration, in which case the real wage is exogenous and pinned down in the rest of the world, and total population adjusts; (ii) restricted international migration, in which case total population remains constant, and the real wage adjusts.

Land is owned by immobile landowners who consume where they live and do not own any labor.²⁴ Total income per unit of land equals the sum of payments to both labor and land and is denoted by $y(\ell)$.

5.2 Production Technology

Production in each sector is characterized by constant returns to scale. For simplicity, we assume a Cobb-Douglas technology, so that output per unit of land in the non-traded sector ($q_N(\ell)$), manufacturing ($q_M(\ell)$) and for an agricultural good g ($q_{g,j}(\ell)$) in land plot j in location ℓ are respectively:

$$\begin{aligned} q_N(\ell) &= z_N(\ell) n_N(\ell)^{1-\alpha_N}, \\ q_M(\ell) &= z_M(\ell) n_M(\ell)^{1-\alpha_M}, \\ q_{g,j}(\ell) &= z_{g,j}(\ell) n_{g,j}(\ell)^{1-\alpha_A}, \end{aligned} \quad (7)$$

²³Consistent with high labor mobility, we observe substantial changes in the distribution of population across locations within Argentina during our sample period. To allow for real wage differences across locations, it would be straightforward to generalize the analysis to allow for idiosyncratic worker preferences for locations, as in Redding (2016).

²⁴Under our assumption of identical and homothetic preferences, all equilibrium allocations are invariant to the number of these landowners.

where $z_N(\ell)$ is non-traded productivity; $z_M(\ell)$ is manufacturing productivity; $z_{g,j}(\ell)$ is productivity for a disaggregated agricultural good; $n_N(\ell)$ is non-traded employment per unit of land; $n_M(\ell)$ is manufacturing employment per unit of land; $n_{g,j}(\ell)$ is employment for a disaggregated agricultural good per unit of land; and $0 < \alpha_i < 1$ is the land intensity in sector $i = A, M, N$. We make the natural assumptions that agriculture is land intensive ($\alpha_A > \alpha_M$ and $\alpha_A > \alpha_N$) and that all sectors use at least some land ($\alpha_M, \alpha_N > 0$).

We allow productivity in all three sectors ($z_M(\ell)$, $z_N(\ell)$, $z_{g,j}(\ell)$) to differ across locations ℓ . In the manufacturing and non-traded sectors, productivity is the same across all land plots j within a given location ℓ . In the agricultural sector, we assume that land plots $j \in [0, L(\ell)]$ can differ in terms of their productivities for individual agricultural goods j ($z_{g,j}(\ell)$) within a given location ℓ . This variation in agricultural productivity enables us to rationalize the production of a range of agricultural goods within each location in the data and captures the impact of differences in soil conditions and topography. In particular, we assume that the realizations of productivity for each agricultural good and land plot $\{z_{g,j}(\ell)\}_{g=1}^G$ are drawn independently from the following Fréchet distribution:

$$\text{Prob}[z_{g,j}(\ell) < z] = e^{-T_g(\ell)z^{-\theta}}, \quad (8)$$

where $T_g(\ell)$ controls the average productivity of good g in location ℓ ; θ controls the dispersion of agricultural productivity across land plots, which we assume is the same for all goods. For most of our quantitative analysis, we are not required to specify whether productivity in each sector ($z_M(\ell)$, $z_N(\ell)$, $\{z_{g,j}(\ell)\}$) is exogenous or endogenous, because we use the equilibrium conditions of the model to solve for the value that productivity must take to rationalize the observed data on the endogenous variables of the model. When we undertake counterfactuals below, we examine the impact of exogenous changes in productivity in a given sector.

In the international trade literature following [Eaton and Kortum \(2002\)](#), the properties of Fréchet distribution are used across a continuum of *goods* to determine patterns of production for each country. In contrast, we use these properties across a continuum of *land plots* to characterize patterns of production for each good. This formulation enables us to consider a discrete number of goods, as observed in the data, and yet still obtain determinate predictions for production patterns for each good (by using the law of large numbers across the continuum of land plots). This specification also allows us to accommodate zero agricultural land shares for some goods in some locations, as observed in the data, because the Fréchet scale parameter that determines average productivity ($T_g(\ell)$) can vary by both good g and location ℓ . Therefore, we rationalize a zero agricultural land share for good g in location ℓ by taking the limit as this productivity parameter converges to zero ($\lim_{T_g(\ell) \rightarrow 0}$). Finally, our framework allows for zero populations in some locations in equilibrium, as observed for some locations and years in the data, which are rationalized in the model by zero productivities in both traded sectors ($\lim_{T_g(\ell) \rightarrow 0}$ for all g and $\lim_{z_M(\ell) \rightarrow 0}$).²⁵

5.3 Profit Maximization

Production in each sector is perfectly competitive. Firms choose employment density (employment per unit of land) to maximize profits, taking as given goods and factor prices and the location decisions of other firms and workers. In equilibrium, firms make zero profits in each sector and location with positive production. Therefore, if a plot of land in location ℓ is used for manufacturing or non-tradables $i = M, N$, land rents ($r_i(\ell)$) are equal to revenue per unit of

²⁵In the web appendix, we consider an extension of the model, in which landowners make an endogenous decision whether to leave land wild or convert it to productive use. In this extension, zero population in a location in equilibrium also can be rationalized by it not being profitable to convert land to productive use in that location.

land minus labor costs per unit of land at the equilibrium value of employment density:

$$r_i(\ell) = \max_{n_i(\ell)} \{P_i q_i(n_i(\ell)) - w(\ell) n_i(\ell)\} \quad \text{for } i = M, N, \quad (9)$$

where $w(\ell)$ is the wage. Alternatively, if a plot of land j in location ℓ is used in agriculture, it is allocated to the agricultural good that offers the highest land rent, and this land rent is again equal to revenue per unit of land minus labor costs per unit of land at the equilibrium value of employment density:

$$\begin{aligned} r_j(\ell) &= \max_{g=1, \dots, G} \{r_{g,j}(\ell)\}, \\ r_{g,j}(\ell) &= \max_{n_{g,j}(\ell)} \{P_g(\ell) q_{g,j}(n_{g,j}(\ell)) - w(\ell) n_{g,j}(\ell)\}. \end{aligned} \quad (10)$$

We assume that the decision whether to allocate a land plot to agriculture, manufacturing or non-tradables is made before observing the realizations for agricultural productivities $\{z_{g,j}(\ell)\}_{g=1}^G$, which captures the role of idiosyncratic shocks to agricultural productivity, such as weather shocks. Therefore, the land use decision across the three sectors depends on the comparison of expected land rents in agriculture ($r_A(\ell) = \mathbb{E}[r_j(\ell)]$) to land rents in the other two sectors ($r_M(\ell), r_N(\ell)$). Expected land rents in agriculture in turn depend on the probability distribution for agricultural productivities (8). After a landowner decides to allocate a land plot to agricultural use, she observes the realizations for productivity for each agricultural good, and decides which of these agricultural goods to produce.

5.4 Sectoral Employment and Wage-Rental Ratio

Using profit maximization and zero profits, equilibrium variables in each sector and location can be written in terms of the equilibrium wage-rental ratio $\omega_i(\ell) = w(\ell) / r_i(\ell)$, which in turn depends on wages ($w(\ell)$), productivity ($z_i(\ell)$) and prices ($P_i(\ell)$). For the manufacturing and non-traded sectors $i \in \{M, N\}$, employment per unit of land and the wage-rental ratio in each location with positive production must satisfy:

$$n_i(\ell) = \frac{1 - \alpha_i}{\alpha_i} \frac{1}{\omega_i(\ell)}, \quad (11)$$

$$\omega_i(\ell) = \left(\frac{w(\ell)}{P_i(\ell) z_i(\ell)} \right)^{\frac{1}{\alpha_i}}. \quad (12)$$

For the agricultural sector, once a plot of land j in location ℓ has been assigned to the production of a given agricultural good g , the equilibrium values of employment per unit of land ($n_{g,j}(\ell)$) and the wage-rental ratio ($\omega_{g,j}(\ell)$) take exactly the same form as above, except with price $P_g(\ell)$ and productivity $z_{g,j}(\ell)$.

We now establish a key aggregation property of the model. Under our assumption of a Fréchet distribution for agricultural productivity, there exists an aggregate measure of agricultural productivity ($z_A(\ell)$) that is a sufficient statistic for the impact of the prices and productivity of the disaggregated agricultural goods on aggregate employment and output in the agricultural sector:

$$z_A(\ell) = \Gamma \left(\frac{\alpha_A \theta - 1}{\alpha_A \theta} \right)^{\alpha_A} \left[\sum_{g=1}^G T_g(\ell) P_g(\ell)^\theta \right]^{1/\theta}, \quad (13)$$

where $\Gamma(\cdot)$ is the Gamma function; and we assume that θ is sufficiently large that $\alpha_A \theta > 1$.

This aggregation result (13) implies that we can treat the agricultural sector $i = A$ as if it consisted of a single good with the same productivity $z_A(\ell)$ across all land plots in location ℓ and a price equal to $P_A(\ell) = 1$. Importantly, this property that $P_A(\ell) = 1$ is not a price normalization, because the entire distribution of agricultural prices ($P_g(\ell)$)

is contained in the definition of agricultural productivity ($z_A(\ell)$). Using this aggregation result (13), employment density ($n_A(\ell)$) and the wage-rental ratio ($w_A(\ell)$) in the aggregate agricultural sector take the same form as for the manufacturing and non-traded sectors in equations (11) and (12), but using the expected land rent $r_A(\ell) = \mathbb{E}[r_j(\ell)]$, the ratio of wages to expected land rents $\omega_A(\ell) \equiv w(\ell)/r_A(\ell)$, productivity $z_A(\ell)$, and $P_A(\ell) = 1$.

5.5 Definition of Equilibrium

Under our neoclassical assumptions, the definition of general equilibrium takes a standard form, in which workers maximize utility and choose their location optimally, firms maximize profits and zero profits are made in each location with positive production, and markets clear.

Definition 1. A general equilibrium consists of a real wage u^* ; allocations of population density $n(\ell)$, land shares $\{L_i(\ell)\}_{i=N,M,A}$, and employment density $\{n_i(\ell)\}_{i=N,M,A}$; wages $w(\ell)$; land rents $r(\ell)$; and prices $\{P_g(\ell)\}_{g=1}^G$, $P_M(\ell)$, $P_N(\ell)$ for all $\ell \in \mathcal{L}$ such that:

- (i) workers maximize utility and choose their location optimally: $u(\ell) \leq u^*$ and $u(\ell) = u^*$ if $n(\ell) > 0$.
- (ii) land is allocated optimally across sectors: $r(\ell) = \max\{r_A(\ell), r_M(\ell), r_N(\ell)\}$.
- (iii) the land market clears in each location: $\sum_{i=N,M,A} L_i(\ell) = L(\ell)$.
- (iv) the labor market clears in each location: $\sum_{i=N,M,A} \frac{L_i(\ell)}{L(\ell)} n_i(\ell) = n(\ell)$.
- (v) the non-traded goods market clears in each location: $c_N(\ell) = \frac{L_N(\ell)}{L(\ell)} q_N(n_N(\ell))$.
- (vi) traded goods prices are determined by no arbitrage: If a location ℓ exports an agricultural good g to the rest of the world, its price equals the price at the nearest port less transport costs, $P_g(\ell) = P_g^*/\delta_g(\ell)$, where $\delta_g(\ell) = \min_{\ell' \in \mathcal{L}_C} \{\delta(\ell, \ell')\}$. If the location ℓ imports the manufacturing good M from the rest of the world its price equals $P_M(\ell) = \delta_M(\ell) P_M^*$, where $\delta_M(\ell) = \min_{\ell' \in \mathcal{L}_C} \{\delta_M(\ell, \ell')\}$.
- (vii) the common real wage u^* adjusts to clear the labor market for the economy as a whole, i.e. condition (6) holds.

Given our neoclassical assumptions, it is straightforward to establish the existence and uniqueness of the equilibrium, such that there exists a unique common real wage across locations (u^*) and a unique set of prices $\{w(\ell), r(\ell), \{P_g(\ell)\}_{g=1}^G, P_M(\ell), P_N(\ell)\}$ and allocations $\{n(\ell), \{L_i(\ell)\}_{i=N,M,A}, \{n_i(\ell)\}_{i=N,M,A}\}$ for each location $\ell \in \mathcal{L}$ that satisfies the above equilibrium conditions.

Proposition 1. There exists a unique general equilibrium.

Proof. See the web-based technical appendix □

In this characterization of equilibrium, a distinction can be drawn between (a) a “local equilibrium” in each location $\ell \in \mathcal{L}$ for given prices of traded goods $\{P_M(\ell), P_g(\ell)\}$ and the common real wage (u^*), which satisfies conditions (i)-(v); and (b) the full general equilibrium, in which these endogenous local prices of traded goods and the common real wage are endogenously determined through no-arbitrage in goods markets and population mobility, and conditions (i)-(vii) are satisfied. In characterizing the full general equilibrium, we either take the common real wage (u^*) as exogenous, and determined by its value in the rest of the world (under our assumption of free international migration), or we take the total population of Argentina (N) as exogenous (under our assumption of restricted international migration).

5.6 Specialization Across Sectors

We now use the equilibrium conditions of the model to characterize specialization across sectors. We show that our assumption of constant returns to scale implies complete specialization in the traded sector between agriculture and manufacturing. Therefore, assuming that all locations within Argentina have a comparative advantage in agriculture, these locations all produce and export agricultural goods, and import the manufacturing good. The model thus rationalizes the extreme concentration of Argentinian exports in agriculture observed in the data. Finally, our specification of CES preferences between traded and non-traded goods implies that the utility function satisfies the Inada conditions, which ensures that all populated locations produce and consume the non-traded good.

To establish these results, we begin by using population mobility, which implies that real wages are equalized across all populated locations,

$$u^* = \frac{w(\ell)}{E(\ell)} = \frac{w(\ell)}{\left[\beta_T E_T(\ell)^{1-\sigma} + (1-\beta_T) P_N(\ell)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}}. \quad (14)$$

A landowner's decision over how to use a land plot is determined by comparison of the wage-rental ratios across the three sectors ($\omega_M(\ell)$, $\omega_A(\ell)$, $\omega_N(\ell)$). As all populated locations produce the non-traded good and at least one traded good, factor mobility across sectors ensures that there is a common equilibrium wage-rental ratio between the non-traded sector and the traded sector(s) with positive production: $\omega_N(\ell) = \omega_i(\ell)$ for $i = A, M$ if $n_i(\ell) > 0$. Using population mobility from equation (14) and profit maximization and zero profits from equation (12), this common equilibrium wage-rental ratio must satisfy,²⁶

$$\left[\beta_T \left(\frac{P_i(\ell)}{E_T(\ell)} z_i(\ell) \omega_i(\ell)^{\alpha_i} \right)^{\sigma-1} + (1-\beta_T) (z_N(\ell) \omega_i(\ell)^{\alpha_N})^{\sigma-1} \right]^{\frac{1}{\sigma-1}} = u^*. \quad (15)$$

Under autarky, there is positive production in all three sectors, and hence a common wage-rental ratio across these three sectors $\{N, M, A\}$. Using equations (12) and (15), we can solve in closed-form for this autarkic wage-rental ratio ($\omega^a(\ell) = \omega_N^a(\ell) = \omega_A^a(\ell) = \omega_M^a(\ell)$) for each location ℓ ,

$$\omega^a(\ell) = \left(\frac{P_M(\ell) z_M(\ell)}{z_A(\ell)} \right)^{1/(\alpha_A - \alpha_M)}. \quad (16)$$

In contrast, when a location is open to trade, it produces the non-traded good and only one of the two traded goods. The reason is that the equilibrium wage-rental ratio in each traded sector in equation (12) depends solely on prices, productivity and the common wage across sectors, and does not depend on the scale of production in any sector. Therefore, depending on the values of prices and productivities, one of the two traded sectors in general will have a lower wage-rental ratio than the other in a given location, which implies that this location will produce only one of the two traded goods. We summarize this complete specialization result within the traded sector as follows.

Proposition 2. *If location ℓ trades, it is either fully specialized in agriculture, in which case $\omega_A(\ell) < \omega^a(\ell)$, or fully specialized in manufacturing, in which case $\omega_M(\ell) < \omega^a(\ell)$. Complete specialization in agriculture occurs for sufficiently high values of agricultural productivity ($z_A(\ell)$) relative to manufacturing productivity ($z_M(\ell)$).*

²⁶To obtain equation (15), first rewrite equation (14) as $u^* = \left[\beta_T \left(\frac{w(\ell)}{E_T(\ell)} \right)^{\sigma-1} + (1-\beta_T) \left(\frac{w(\ell)}{P_N(\ell)} \right)^{\sigma-1} \right]^{\frac{1}{\sigma-1}}$ and then eliminate $w(\ell)$ using the expressions $w(\ell) = P_i(\ell) z_i(\ell) \omega(\ell)^{\alpha_i}$ and $\frac{w(\ell)}{P_N(\ell)} = z_N(\ell) \omega(\ell)^{\alpha_N}$ implied by (ii) in Definition 1 and equation (12).

Proof. See the web-based technical appendix. □

As discussed above, based on the overwhelming concentration of exports in agriculture, we assume that all locations within Argentina have a comparative advantage relative to the rest of the world in agriculture (i.e., $z_A(\ell)$ is sufficiently large in each location that $\omega_A(\ell) < \omega^a(\ell)$). Whether any given location is closed or open to trade is determined by comparative advantage and transport costs. In particular, trade occurs if the relative price of the imported manufacturing good net of transport costs is less than the relative price of the manufacturing good under autarky. In contrast, for sufficiently large transport costs, the model features a “trade frontier” beyond which regions further inland are in autarky. As transport costs fall, this frontier expands further inland as additional regions are integrated into world markets. Finally, for a given value of transport costs, a location is open to trade for a sufficiently large comparative advantage in agricultural goods (a high enough value of $z_A(\ell)/z_M(\ell)$).

Under these assumptions on comparative advantage, each location that is open to trade specializes in agriculture and non-traded goods, which implies that the population mobility condition (15) can be re-written as:

$$\left[\beta_T (\tilde{z}_A(\ell) \omega(\ell)^{\alpha_A})^{\sigma-1} + (1 - \beta_T) (z_N(\ell) \omega(\ell)^{\alpha_N})^{\sigma-1} \right]^{\frac{1}{\sigma-1}} = u^*, \quad (17)$$

where we define

$$\tilde{z}_A(\ell) = \frac{z_A(\ell)}{E_T(\ell)} \quad (18)$$

as a measure of agricultural productivity adjusted by the tradables price index.

This measure of adjusted-agricultural productivity ($\tilde{z}_A(\ell)$) summaries the attractiveness of a location for the production and consumption of traded goods. We show below that this adjusted-agricultural productivity ($\tilde{z}_A(\ell)$) and non-traded productivity ($z_N(\ell)$) are sufficient statistics for determining the equilibrium population density and employment share in the traded and non-traded sectors for each location ℓ .

5.7 Specialization Within the Agricultural Sector

We now determine patterns of specialization across disaggregated goods within the agricultural sector. We show that the model implies systematic differences in the composition of agricultural production across locations, which are determined by relative productivity and trade costs for these disaggregated agricultural goods.

With a continuum of land plots within each location, the share of agricultural land allocated to good g equals the probability that an individual land plot is allocated to that good. Therefore, using the properties of the Fréchet distribution of agricultural productivities, the share of agricultural land allocated to each good depends on relative productivities $\{T_g(\ell)\}$, relative local prices $\{P_g(\ell)\}$, and the Fréchet shape parameter θ :

$$l_g(\ell) = \frac{T_g(\ell) P_g(\ell)^\theta}{\sum_{g'} T_{g'}(\ell) P_{g'}(\ell)^\theta}. \quad (19)$$

Combining this result for patterns of agricultural production with an assumption over the functional form for the tradables price index $E_T(\ell)$, we can solve for patterns of trade in the disaggregated agricultural goods. In particular, under the assumption that the tradeables price index is Cobb-Douglas, a constant share (γ_g) of overall spending on agriculture is allocated to each agricultural good:

$$E_T(\ell) = P_M(\ell)^{1-\gamma_A} \prod_{g=1}^G P_g(\ell)^{\gamma_g}, \quad \text{where} \quad \sum_{g=1}^G \gamma_g = \gamma_A. \quad (20)$$

Using this constant Cobb-Douglas expenditure share (γ_g) together with our expression for the share of agricultural land allocated to each good in equation (19), we obtain the following closed-form solution for exports of each disaggregated agricultural good ($x_g(\ell)$) as a share of overall agricultural exports ($x_A(\ell)$):

$$\frac{x_g(\ell)}{x_A(\ell)} = \frac{l_g(\ell) - \gamma_g}{1 - \gamma_A}. \quad (21)$$

While each location is a net exporter of agricultural goods, and a net importer of manufacturing goods, there is also internal bilateral trade in the disaggregated agricultural goods between locations within Argentina. This internal bilateral trade depends on the average realizations for idiosyncratic productivity across land plots (as determined by $T_g(\ell)$), local prices ($P_g(\ell)$) as determined by prices at the port and transport costs), and the Cobb-Douglas expenditure shares for each disaggregated agricultural good (γ_g). From equation (19), we have already shown that the share of agricultural land allocated to each good ($l_g(\ell)$) depends on relative values of agricultural productivity ($T_g(\ell)$) and local prices ($P_g(\ell)$). Using this relationship, we obtain the result in equation (21) that each location is a net exporter of an individual disaggregated agricultural good ($x_g(\ell) > 0$) if the share of agricultural land that it allocates to the production of that good is greater than its share of expenditure on that good ($x_g(\ell) > \gamma_g$).

We thus obtain a neoclassical prediction for chains of comparative advantage within the agricultural sector, such that if location ℓ exports good g , it necessarily exports all goods g' such that $l_{g'}(\ell) / l_g(\ell) > \gamma_{g'} / \gamma_g$:

$$\frac{x_{g'}(\ell)}{x_g(\ell)} = \frac{l_{g'}(\ell) - \gamma_{g'}}{l_g(\ell) - \gamma_g}. \quad (22)$$

5.8 Spatial Balassa-Samuelson Effect

A key prediction of the model is that internal geography determines not only the overall level of economic activity, as measured by population density, but also the composition of economic activity, both between the traded and non-traded sectors, and across goods within the traded sector. To establish this role for internal geography, we use the population mobility condition (14), zero-profits and profit maximization in each sector from equation (12), and the labor market clearing condition (6). Together these relationships imply the following closed-form expressions for equilibrium population density ($n(\ell)$) and the share of labor employed in agriculture ($\nu_A(\ell)$):

$$n(\ell) = \frac{N(\ell)}{L(\ell)} = \left(\frac{1}{\alpha_N + (\alpha_A - \alpha_N) \beta_T \left(\frac{E_T(\ell)}{E(\ell)} \right)^{1-\sigma}} - 1 \right) \frac{1}{\omega(\ell)}, \quad (23)$$

$$\nu_A(\ell) = \frac{N_A(\ell)}{N(\ell)} = \frac{(1 - \alpha_A) \beta_T \left(\frac{E_T(\ell)}{E(\ell)} \right)^{1-\sigma}}{1 - \left(\alpha_N + (\alpha_A - \alpha_N) \beta_T \left(\frac{E_T(\ell)}{E(\ell)} \right)^{1-\sigma} \right)}. \quad (24)$$

Equations (23) and (24) summarize the equilibrium relationship between population density ($n(\ell)$) and the agricultural employment share ($\nu_A(\ell)$) in the model and have an intuitive interpretation. A higher population density must be accommodated through some combination of both sectors using more labor-intensive production techniques (which requires a lower wage-rental ratio, $\omega(\ell)$) or a higher share of employment in the labor-intensive non-traded sector (which with $0 < \sigma < 1$ requires a lower relative price of traded goods, $E_T(\ell)/E(\ell)$).

Totally differentiating equations (23) and (24), and using profit maximization and zero profits from equation (12) and labor market clearing from equation (6), we obtain the following system of equations for changes in population

density ($\widehat{n}_t(\ell)$), the agricultural employment share ($\widehat{\nu}_{At}(\ell)$), the wage-rental ratio ($\widehat{\omega}(\ell)$) and the relative price of traded goods ($\widehat{\frac{E_T(\ell)}{E(\ell)}}$) as a function of changes in adjusted-agricultural productivity ($\widehat{z}_A(\ell)$), non-traded productivity ($\widehat{z}_N(\ell)$), the common level of utility across all locations (\widehat{u}^*) and total population (\widehat{N}):

$$\widehat{n}_t(\ell) = \frac{(\alpha_A - \alpha_N) \nu_{At}(\ell)}{\alpha_N (1 - \alpha_A) + (\alpha_A - \alpha_N) \nu_{At}(\ell)} \widehat{\nu}_{At}(\ell) - \widehat{\omega}_t(\ell), \quad (25)$$

$$\widehat{\nu}_{At}(\ell) = \left(1 + \frac{\alpha_A - \alpha_N}{1 - \alpha_A} \nu_{At}(\ell)\right) (1 - \sigma) \left(\widehat{\frac{E_{Tt}(\ell)}{E_t(\ell)}}\right), \quad (26)$$

$$\widehat{\omega}(\ell) = \frac{(1 - \alpha_N) \nu_A(\ell) (\widehat{u}^* - \widehat{z}_A(\ell)) + (1 - \alpha_A) (1 - \nu_A(\ell)) (\widehat{u}^* - \widehat{z}_N(\ell))}{\alpha_A (1 - \alpha_N) \nu_A(\ell) + \alpha_N (1 - \alpha_A) (1 - \nu_A(\ell))}, \quad (27)$$

$$\left(\widehat{\frac{E_T(\ell)}{E(\ell)}}\right) = \frac{(1 - \alpha_A) (1 - \nu_A(\ell)) [\alpha_A \widehat{z}_N(\ell) - \alpha_N \widehat{z}_A(\ell) - (\alpha_A - \alpha_N) \widehat{u}^*]}{\alpha_A (1 - \alpha_N) \nu_A(\ell) + \alpha_N (1 - \alpha_A) (1 - \nu_A(\ell))}, \quad (28)$$

$$\sum_{\ell} \nu(\ell) \widehat{n}(\ell) = \widehat{N}, \quad (29)$$

where $\nu(\ell) = N(\ell)/N$ is the share of location ℓ in the economy's total population; a hat above a variable denotes a proportional change, such that $\widehat{n}_t(\ell) \equiv dn_t(\ell)/n_t(\ell)$; and these proportional changes can either involve derivatives across locations at a given point in time or derivatives over time.

As discussed above, we solve this system of equations under two possible assumptions about the change in population (\widehat{N}) and the common level of utility across all locations (\widehat{u}^*). Either we assume free international migration, in which case the common level of utility across all locations is exogenously determined by its value in the rest of world ($\widehat{u}^* = 0$), and the change in population (\widehat{N}) is endogenously determined by this system of equations. Or we assume restricted international migration, in which case total population is constant ($\widehat{N} = 0$), and the change in the common level of utility across all locations (\widehat{u}^*) is endogenously determined by this system of equations.

We now connect the changes in adjusted-agricultural productivity ($\widehat{z}_A(\ell)$) in this system of equations to the underlying internal geography and trade costs. From equation (18), the change in adjusted-agricultural productivity ($\widehat{z}_A(\ell)$) equals the change in agricultural productivity ($\widehat{z}_A(\ell)$) minus that in the tradeables price index ($\widehat{E}_T(\ell)$):

$$\widehat{z}_A(\ell) = \widehat{z}_A(\ell) - \widehat{E}_T(\ell). \quad (30)$$

Using our aggregation result for agricultural productivity (13) and our solution for agricultural land shares from equation (19), we can express changes in agricultural productivity ($\widehat{z}_A(\ell)$) in terms of initial land shares ($l_g(\ell)$) and changes in the primitives of productivity for each agricultural good ($\widehat{T}_g(\ell)$), internal trade costs ($\widehat{\delta}_g(\ell)$) and the price for each agricultural good at the port (\widehat{P}_g^*):

$$\widehat{z}_A(\ell) = \sum_{g=1}^G l_g(\ell) \left(\frac{\widehat{T}_g(\ell)}{\theta(\ell)} + \widehat{P}_g^* - \widehat{\delta}_g(\ell) \right). \quad (31)$$

Finally, under the assumption of a Cobb-Douglas tradables consumption index (20), and using the fact that all locations are net importers of manufacturing goods, we can also express changes in the tradables consumption price index ($\widehat{E}_T(\ell)$) in terms of changes in the primitives of internal trade costs ($\widehat{\delta}_g(\ell)$) and prices at the port (\widehat{P}_g^*):

$$\widehat{E}_T(\ell) = (1 - \gamma_A) [\widehat{P}_M^* - \widehat{\delta}_M(\ell)] + \sum_{g=1}^G \gamma_g [\widehat{P}_g^* - \widehat{\delta}_g(\ell)]. \quad (32)$$

Together the system of equations (25)-(32) determines the response of the endogenous spatial distribution of economic activity to exogenous changes in productivity, internal trade costs and prices at the port for the disaggregated agricultural goods $\{\hat{T}_g(\ell), \hat{\delta}_g(\ell), \hat{P}_g^*\}$, non-traded productivity $(\hat{z}_N(\ell))$, and either total population (\hat{N}) or the common level of utility (\hat{u}^*) . We now in a position to establish our key Spatial Balassa-Samuelson result:

Proposition 3. (*Spatial-Balassa Samuelson Effect*) Assume that traded and non-traded goods are complements ($\sigma < 1$) and agriculture is land-intensive relative to non-tradables ($\alpha_A > \alpha_N$). Under these assumptions, low trade-cost locations (locations ℓ with lower transport costs $\delta(\ell, \ell')$ to coastal locations $\ell' \in \mathcal{L}_C$) have (i) higher adjusted-agricultural productivity $(\hat{z}_A(\ell))$, (ii) lower wage-rental ratios $(\omega(\ell))$, (iii) higher relative prices of non-traded goods $(E_T(\ell)/E(\ell))$, (iv) higher population densities $(n(\ell))$, and (v) lower agricultural employment shares $(\nu_A(\ell))$.

Proof. The proposition follows from equations (25)-(32), as shown in the web-based technical appendix. □

This proposition is related to the conventional Balassa-Samuelson effect in macroeconomics, in which higher productivity in tradeables at the level of the economy as a whole causes a rise in the relative price of the non-traded good. In this conventional specification, with inelastic demand between sectors, higher productivity in the traded sector can either raise or reduce employment in that sector, depending on whether the economy is open or closed to international trade, as in Matsuyama (1992) and Uy, Yi, and Zhang (2012). In contrast to this conventional specification, our spatial Balassa-Samuelson effect operates across locations within an open economy that are linked through goods trade and factor mobility, and arises because internal trade costs induce endogenous differences across these locations in price-adjusted productivity in the traded sector. The intuition for our spatial Balassa-Samuelson effect is as follows. Locations with good access to world markets are attractive for the production and consumption of traded goods, which increases population density, and bids up the reward of the immobile factor (land) relative to that of the mobile factor (labor). Together the increase in population and the reduction in wages relative to land rents induce an expansion in the employment share of the labor-intensive non-traded sector, which requires a higher relative price for the non-traded good, given inelastic demand between sectors ($0 < \sigma < 1$).

Through this spatial Balassa-Samuelson effect, our model provides a microeconomic rationale for our earlier reduced-form evidence on patterns of development across sectors and regions in Argentina. As locations close to Buenos Aires and its surrounding ports face lower trade costs in accessing world markets, the model rationalizes the high population densities and urban population shares in these locations. With the reorientation of trade routes away from the North-West and towards Buenos Aires and its surrounding ports, and the fall in external trade costs to markets in Europe and North America, the model also predicts an increase in population density and urban population shares in these locations. Finally, as the expansion of the railroad network reduces the trade costs faced by interior regions in accessing world markets, and changes the relative trade costs for different agricultural goods, the model predicts an increase in population density, structural transformation from agriculture to non-agriculture, and a change in the composition of agricultural production in these interior regions.

As we derive this spatial Balassa-Samuelson effect from a general neoclassical production structure, it captures a generic feature of patterns of economic development that applies more broadly. As long as there is specialization according to comparative advantage in the traded sector, non-traded goods are labor intensive, and demand is inelastic between sectors, this force will operate. Therefore, our model provides a natural rationale for a common pattern of

economic development, in which locations close to world markets have higher population densities, urban population shares, relative prices of non-traded goods, and land prices relative to wages.

6 Quantitative Analysis

We now show that our model not only rationalizes the qualitative properties of our reduced-form empirical findings but can also account for the data quantitatively. First, we show that the model can be inverted to recover unique values for our two sufficient statistics of adjusted-agricultural productivity ($\tilde{z}_{At}(\ell)$) and non-traded productivity ($z_{Nt}(\ell)$) from the observed values of population density ($n_t(\ell)$) and the agricultural employment share ($\nu_{At}(\ell)$), where throughout the remainder of this section we make explicit the time subscripts. Second, we examine the role of internal geography in determining structural transformation through these two sufficient statistics, in the form of both proximity to Argentina's trade hub and railroad access.

6.1 Sufficient Statistics

We begin by using the observed data on population density ($n_t(\ell)$) and the agricultural employment share ($\nu_{At}(\ell)$) to solve for unique values for the sufficient statistics of adjusted-agricultural productivity ($\tilde{z}_{At}(\ell)$) and non-traded productivity ($z_{Nt}(\ell)$). We assume central values for the model's parameters from the existing empirical literature. In particular, we set the shares of land in production costs in agriculture and non-tradeables as $\alpha_A = 0.2$ and $\alpha_N = 0.10$, which are line with the values in [Caselli and Coleman \(2001\)](#). We assume an elasticity of substitution between traded and non-traded goods of $\sigma = 0.5$, which is consistent with a long line of research in macroeconomics that assumes inelastic demand between sectors, including for example [Ngai and Pissarides \(2007\)](#). We set the weight of tradeables in consumer expenditure as equal to $\beta_T = 0.3$, which ensures that the model is consistent with historical shares of tradables in consumer expenditure. Finally, we assume a value for the Fréchet shape parameter for the dispersion of agricultural productivities across land plots of $\theta = 5$, which is line with the values for the dispersion of productivities across goods in the Ricardian trade literature following [Eaton and Kortum \(2002\)](#) and [Donaldson \(2018\)](#).

Using these assumed parameters and the recursive structure of the model, we solve for our two sufficient statistics. First, we determine the unique relative price of traded goods ($E_{Tt}(\ell)/E_t(\ell)$) in each location ℓ from the observed agricultural employment share ($\nu_{At}(\ell)$) using equation (24):

$$\frac{E_{Tt}(\ell)}{E_t(\ell)} = \left(\frac{1}{\beta_T} \frac{(1 - \alpha_N) \nu_{At}(\ell)}{(1 - \alpha_A) + (\alpha_A - \alpha_N) \nu_{At}(\ell)} \right)^{\frac{1}{1-\sigma}}. \quad (33)$$

Second, we recover the unique wage-rental ratio ($\omega_t(\ell)$) for each location ℓ from the observed agricultural employment share ($\nu_{At}(\ell)$) and population density ($n_t(\ell)$) using equations (23) and (24):

$$\omega_t(\ell) = \frac{(1 - \alpha_A)(1 - \alpha_N)}{\alpha_N(1 - \alpha_A) + (\alpha_A - \alpha_N) \nu_{At}(\ell)} \frac{1}{n_t(\ell)}. \quad (34)$$

Third, using these solutions for the relative price of traded goods ($E_{Tt}(\ell)/E_t(\ell)$) and the wage-rental ratio ($\omega_t(\ell)$), together with profit maximization and zero-profits (12) and population mobility (14), we obtain the following closed-form solutions for adjusted-agricultural productivity ($\tilde{z}_{At}(\ell)$) and non-traded productivity ($z_{Nt}(\ell)$):

$$\tilde{z}_{At}(\ell) = \frac{u_t^*}{\omega_t(\ell)^{\alpha_A}} \frac{1}{(E_{Tt}(\ell)/E_t(\ell))}, \quad (35)$$

$$z_{Nt}(\ell) = \frac{u_t^*}{\omega_t(\ell)^{\alpha_N}} \left(\frac{1 - \beta_T}{1 - \beta_T \left(\frac{E_{Tt}(\ell)}{E_t(\ell)} \right)^{1-\sigma}} \right)^{\frac{1}{1-\sigma}}. \quad (36)$$

From equations (35) and (36), we recover adjusted-agricultural productivity ($\tilde{z}_{At}(\ell)$) and non-traded productivity ($z_{Nt}(\ell)$) up to the common level of utility across all locations (u_t^*). We choose units in which to measure this common level of utility such that it takes the value one in 1914, and we calibrate its values for 1869 and 1895 to match the estimates of real wage growth in Argentina between these years from [Taylor and Williamson \(1997\)](#).²⁷ These choices for the common level of utility $\{u_{1914}^*, u_{1895}^*, u_{1869}^*\}$ leave the distributions of employment and population across locations unchanged, because these distributions depend solely on the relative value of these productivities across locations. From equations (33)-(36), adjusted-agricultural productivity ($\tilde{z}_{At}(\ell)$) and non-traded productivity ($z_{Nt}(\ell)$) are defined for districts with positive values for both agricultural and non-traded employment. To avoid extreme values for productivity for districts with agricultural employment shares close to zero or one, we therefore focus in our quantitative analysis on the sample of districts for which agriculture accounts for more than 5 and less than 95 percent of employment, of which there are 318 districts in 1914.

6.2 Spatial Gradients and Railroad Access

Using these model solutions $\{E_{Tt}(\ell)/E_t(\ell), \omega_t(\ell), \tilde{z}_{At}(\ell), z_{Nt}(\ell)\}$, we now examine the quantitative magnitude of the spatial Balassa-Samuelson effect in late-19th century Argentina. We begin by considering the impact of distance from Argentina's international trade hub. In Table 4, we present the results of estimating the same specification as reported in Table 1 above, but using our model solutions instead of the observed data. As all of our model solutions are ultimately derived from the observed population density ($n_t(\ell)$) and agricultural employment share ($\nu_{At}(\ell)$), we begin by reporting results for these two observed variables in Panels A and B for the sample of districts used in our quantitative analysis. We measure the agricultural employment share ($\nu_{At}(\ell)$) in the model using the rural population share in the data. Consistent with the results for the full sample in Table 1, we find a sharp negative and statistically significant gradient in distance from Argentina's trade hub for population density, which rises from -0.45 to -0.80 from 1869-1914. We also find a marked positive and statistically significant gradient for the rural population share, which rises from 0.07 to 0.12 between these same years.

These patterns for the two observed variables imply substantial and statistically significant differences in the relative price of non-traded goods and the wage-rental ratio. As shown in Panels C and D, we find that a doubling of distance to the closest top-four port (a 100 percent increase) is associated with a 12 percent increase in the relative price of traded goods and a 43 percent increase in the wage-rental ratio in 1869, with these elasticities approximately doubling to 23 percent and 75 percent in 1914. As displayed in Panels E and F, we find that the higher relative price of traded goods and higher wage-rental ratio in more remote locations are explained by a combination of lower adjusted-agricultural productivity and higher non-traded productivity. Both gradients in distance from Argentina's trade hub steepen over time, with the negative and statistically significant elasticity for adjusted-agricultural productivity increasing in absolute magnitude from -0.21 to -0.38 from 1869-1914, and the positive and statistically significant elasticity for non-traded productivity rising from 0.27 to 0.40 between these same years.²⁸ Therefore, we find substantial

²⁷The implied normalizations for utility are as follows: $u_{1914}^* = 1$, $u_{1895}^* = 0.79$ and $u_{1869}^* = 0.65$.

²⁸As for the full sample in Table 1, we find a similar pattern of results in Table 4 if we restrict attention to the balanced panel of districts that are present in all three years. For example, we find a steepening of the gradient with respect to distance from the nearest top-four port between 1869

effects of internal geography on the relative values of adjusted-agricultural productivity and non-traded productivity. Locations close to world markets have substantially higher values of adjusted-agricultural productivity at the beginning of our sample period in 1869, and the magnitude of these differences increases sharply over our sample period with Argentina's growing integration into the world economy during the late-19th century.

Table 4: Spatial Balassa-Samuelson Effect and Distance to Top-Four Ports

	(1)	(2)	(3)
Panel A	Log Population Density 1869	Log Population Density 1895	Log Population Density 1914
Log Distance Top-Four Port	-0.459*** (0.089)	-0.823*** (0.089)	-0.796*** (0.085)
Observations	164	255	318
R-squared	0.16	0.29	0.26
Panel B	Log Rural Population Share 1869	Log Rural Population Share 1895	Log Rural Population Share 1914
Log Distance Top-Four Port	0.065** (0.030)	0.098*** (0.023)	0.123*** (0.031)
Observations	164	255	318
R-squared	0.03	0.10	0.08
Panel C	Log Relative Price Tradeables 1869	Log Relative Price Tradeables 1895	Log Relative Price Tradeables 1914
Log Distance Top-Four Port	0.120** (0.057)	0.181*** (0.043)	0.231*** (0.060)
Observations	164	255	318
R-squared	0.03	0.10	0.08
Panel D	Log Wage-Rental Ratio 1869	Log Wage-Rental Ratio 1895	Log Wage-Rental Ratio 1914
Log Distance Top-Four Port	0.430*** (0.087)	0.779*** (0.089)	0.748*** (0.082)
Observations	164	255	318
R-squared	0.15	0.27	0.24
Panel E	Log Adjusted Agricultural Productivity 1869	Log Adjusted Agricultural Productivity 1895	Log Adjusted Agricultural Productivity 1914
Log Distance Top-Four Port	-0.206*** (0.062)	-0.337*** (0.047)	-0.381*** (0.067)
Observations	164	255	318
R-squared	0.07	0.22	0.16
Panel F	Log Non-Traded Productivity 1869	Log Non-Traded Productivity 1895	Log Non-Traded Productivity 1914
Log Distance Top-Four Port	0.272*** (0.084)	0.488*** (0.070)	0.395*** (0.061)
Observations	164	255	318
R-squared	0.05	0.16	0.10

Notes: Observations are a cross-section of Argentinian districts in the respective year. Log population density is the log of the population per unit of land area and is observed in the data. Log rural population share is the log of the rural population as a share of the total population and is observed in the data. Log relative prices of tradeables, log wage-rental ratio, log adjusted agricultural productivity, and log non-traded productivity are model solutions, as discussed in the main text. Distance to Top-Four Port is the geographic (Great Circle) distance from the centroid of each district to the nearest top-four port (Buenos Aires, Rosario, La Plata and Bahia Blanca). Heteroskedasticity robust standard errors in parentheses. *** denotes significance at the 1 percent level; ** denotes significance at the 5 percent level; * denotes significance at the 10 percent level.

We next examine the role of the expansion in the railroad network in reducing the remoteness of interior regions. In Table 5, we present the results of estimating the same instrumental variables specification for railroad access as reported in Table 2, but using our model solutions instead of the observed data. In the interests of brevity, we focus on our two sufficient statistics of adjusted-agricultural productivity ($\tilde{z}_{At}(\ell)$) and non-traded productivity ($z_{Nt}(\ell)$). We begin by regressing adjusted-agricultural productivity at the end of our sample period in 1914 on railroad ac-

and 1914 from 0.08 to 0.32 for the relative price of traded goods in Panel C, from 0.37 to 0.70 for the relative wage rental ratio in Panel D, from -0.15 to -0.46 for adjusted agricultural productivity in Panel E, and from 0.20 to 0.45 for non-traded productivity in Panel F.

cess, log distance to the nearest top-four port and log land area, including initial log rural population in 1869 as a control for historical patterns of rural development. In Column (1), we report the OLS estimates. In Column (2), we present the two-stage least squares estimates, using both our port and colonial post instruments. As predicted by the model, railroad access has a positive and statistically significant impact on adjusted-agricultural productivity, with the instrumental variables estimates marginally larger than but not statistically significantly different from the OLS estimates. We find that the instruments have power in the first-stage regression, with a first-stage F-statistic equal to 34.5 (above the conventional threshold of 10). In a Hansen-Sargan overidentification test, we are again unable to reject the model's overidentifying restrictions (p -value=0.24).

Table 5: Spatial Balassa-Samuelson Effect and Railroad Access

	(1)	(2)	(3)	(4)
	Log Adjusted Agricultural Productivity 1914	Log Adjusted Agricultural Productivity 1914	Log Non- Traded Productivity 1914	Log Non- Traded Productivity 1914
Log Distance Top-Four Port	-0.216*** (0.075)	-0.213** (0.085)	0.240*** (0.075)	0.333*** (0.096)
Log Land Area	-0.470*** (0.078)	-0.473*** (0.087)	-0.221* (0.132)	-0.360*** (0.127)
Share Rail Length 1914	0.634*** (0.142)	0.657*** (0.255)	-0.852*** (0.277)	0.146 (0.670)
Log Rural Population Density 1869	-0.005 (0.067)	-0.004 (0.067)		
Log Urban Population Density 1869			-0.471*** (0.084)	-0.492*** (0.082)
Frequency Along Least Cost Path to Top-Four Port				
Frequency Along Spanish Colonial Postal Routes				
Instruments	-	Both	-	Both
Observations	259	259	155	155
R-squared	0.40	-	0.35	-
First-stage F-Statistic	-	34.5	-	16.82
Overidentification test (p-value)	-	0.2444	-	0.2659

Notes: Observations are a cross-section of Argentinian districts. Distance Top-Four Port is the geographic (Great Circle) distance from the centroid of each district to the nearest top-four port (Buenos Aires, Rosario, La Plata and Bahia Blanca). Log adjusted-agricultural productivity and log non-traded productivity are model solutions. Share rail length is the length of railroads in each district as a percentage of this length for Argentina as a whole. Port instrument is the percentage of grid points within each district that lie on the least-cost routes from the centroids of all Argentinian districts to the top-four ports. Colonial post is the length of Spanish colonial postal routes in each district as a percentage of this length for Argentina as a whole. First-stage F-statistic is a test of the statistical significance of the instruments in the first-stage regression. Overidentification test is a Hansen-Sargan test of the model's overidentifying restrictions. Heteroskedasticity robust standard errors in parentheses. *** denotes significance at the 1 percent level; ** denotes significance at the 5 percent level; * denotes significance at the 10 percent level.

Finally, we regress non-traded productivity at the end of our sample period on railroad access, log distance to the nearest top-four port and log land area, including initial log urban population in 1869 as a control for historical patterns of urban development. In Column (3), we report the OLS estimates. In Column (4), we present the two-stage least squares estimates, again using both of our instruments. In the OLS specification, we find a negative and statistically significant coefficient on railroad access for non-traded productivity, although this coefficient changes sign and becomes statistically insignificant once we instrument for railroad access. This pattern of results, where we

only find the positive impact of the railroad network for adjusted agricultural productivity, but not for non-traded productivity, provides support for the mechanism in the model. As the expansion of the railroad network improves the access of interior regions to world markets, the model implies increases in export prices and reductions in import prices, which raise adjusted-agricultural productivity. Again, we find that the instruments have power in the first-stage, with a first-stage F-statistic equal to 16.82 (above the conventional threshold of 10), and we pass the Hansen-Sargan test of the model's overidentifying restrictions (p -value=0.27).

Therefore, taking the results of this section together, we find that the spatial Balassa-Samuelson effect is quantitatively relevant for late-19th century Argentina. We find substantial and statistically significant effects of internal geography on adjusted-agricultural productivity, both for distance from Argentina's trade hub and railroad access.

6.3 Further Evidence

Our solutions for adjusted-agricultural productivity ($\tilde{z}_{At}(\ell)$) and non-traded productivity ($z_{Nt}(\ell)$) in the model are derived from the observed data on population density ($n_t(\ell)$) and the agricultural employment share ($\nu_{At}(\ell)$). We now report the results of an external validation exercise, in which we examine whether the model's predictions for adjusted-agricultural productivity are correlated with other observed variables that were not used in our quantitative analysis but are expected to be closely related to adjusted-agricultural productivity.

In Table 6, we report conditional correlations between adjusted-agricultural productivity and a range of observed variables. Each cell of the table corresponds to a separate regression, with the dependent variable reported in the columns of the table, and the independent variable given in the rows of the table. Each of these regressions corresponds to a correlation between endogenous variables, where we control in all specifications for log land area as a determinant of the scale of agricultural production. In Panel A, we consider land values, which are reported in the 1895 population census. Under our assumption that all locations within Argentina have a comparative advantage in agriculture, the model implies that land values are closely related to agricultural productivity through the zero-profit condition for production in agriculture: $r_t(\ell) = z_{At}(\ell) \omega_t(\ell)^{\alpha_A - 1}$. Although we only recover adjusted-agricultural productivity from our solution of the model ($\tilde{z}_{At}(\ell) = z_{At}(\ell) / E_{Tt}(\ell)$), we expect this to be highly correlated with agricultural productivity ($z_{At}(\ell)$). Therefore, as a first validation exercise, we regress the log of adjusted-agricultural productivity ($\tilde{z}_{At}(\ell)$) on the log of observed land values ($r_t(\ell)$). As reported in Panel A, we find a strong positive and statistically significant relationship, with an elasticity somewhat above one, confirming that land values are indeed closely related to the model's sufficient statistic for employment and output in the agricultural sector.²⁹

In the remaining panels of Table 6, we consider the range of observed measures of economic activity within the agricultural sector from Table 3 above. Panel B examines crop cultivated area, measured as a percentage of total land area for each district. Panel C investigates agricultural machinery, measured as the number of each type of machine for a district as a percentage of the total for this type of machine for Argentina as a whole. Panel D explores livestock, measured as the number of each type of livestock for each district as a percentage of the total for this type of livestock for Argentina as a whole. Finally, Panel E considers railroad shipments, measured as the quantity of each good shipped from stations in each district as a percentage of the total for this same good for Argentina as a whole. In each case, we define the variable in terms of percentage shares to permit the inclusion of zero values.

²⁹In 1914, the population census reports the distribution of estates (*estancias*) across a number of discrete land value intervals. Constructing land values for each district in 1914 using these distributions and the mid-point for each interval, we find a similar relationship between adjusted-agricultural productivity and land values.

Table 6: Adjusted-Agricultural Productivity and Measures of Agricultural Production

	(1)	(2)	(3)	(4)
	Log Adjusted Agricultural Productivity 1895	Log Adjusted Agricultural Productivity 1914	Log Adjusted Agricultural Productivity 1895	Log Adjusted Agricultural Productivity 1914
Panel A : Land Values			Panel D : Livestock	
Land Value	1.3798***	-	Native Cattle	0.2471***
Panel B : Cultivated Area			Mixed Cattle	-0.1233
Corn	0.0479**	0.0116	Pure-breed Cattle	0.0959**
Wheat	0.0128***	0.0203***	Native Sheep	0.2065***
Flax	0.0714**	0.0137	Mixed Sheep	0.1144*
Vegetables	0.0920**	0.5490**	Pure-breed Sheep	0.1089**
Barley	0.2957	0.0367		0.3138***
Potato	0.7746**	0.4657**	Panel E : Railroad Shipments	
Cotton	-9.7003	0.5421	Corn	0.0942**
Peanuts	0.1785	-0.2640	Flour	0.2733***
Beans	0.0601	-0.6139	Wheat	0.0979
Tobacco	-0.7064***	-0.3348***	Flax	0.1045***
Sugar Cane	0.0792***	0.0400	Wool	0.3847***
Wine	0.0149*	0.0192***	Leather	0.0422*
Panel C : Machinery			Alfalfa	0.1702***
Wind Machines	0.0518	0.3615***	Sugar	0.1072***
Water Machines	0.0807**	0.2182***	Sand and Stone	0.3324***
Mowers	0.1770***	0.2789***	Wood	0.0366***
Threshers	0.1123***	0.2725***	Wine	0.1018
Rakes	0.1802***	0.3634***	Sugar Cane	0.7111***
Ploughs	0.3676***	0.4762**		0.0874***
Combines		0.0888*		0.0154
Seeders		0.2553***		0.0759
Dredges		0.3663***		-0.0154***
Gleaners		0.1749***		0.0218
Cars		0.8438***		
Lawn Mower		0.3854***		
Rollers		0.0881**		
Coaches		0.6689***		
Breakers		0.1216		
Carts & Wagons		0.9670***		
Engines		0.5760***		
Artesian Wells		0.4821***		
Shellers		0.4037***		
Baling		0.6772***		
Shearers		0.1793***		

Notes: Each cell of the table corresponds to a separate regression. Observations are a cross-section of Argentinian districts in the respective year. Columns correspond to the dependent variable. Rows correspond to the independent variable. Log adjusted-agricultural productivity is from the solution of the model, as discussed in the main text. Cultivated area is crop cultivated area divided by total land area for each district. Agricultural machinery is the number of each type of agricultural machine for each district as a share of the total for that type for Argentina as a whole. Livestock is the number of each type of livestock for each district as a share of the total for that type for Argentina as a whole. Railroad shipments is the quantity shipped of each good from railroad stations in each district as a share of the total for that good for Argentina as a whole. All specifications control for the log land area of each district. Statistical significance based on heteroskedasticity robust standard errors. *** denotes significance at the 1 percent level; ** denotes significance at the 5 percent level; * denotes significance at the 10 percent level.

As shown in the table, we find that higher adjusted-agricultural productivity is typically positively correlated with these different measures of production activity within the agricultural sector, especially for the new export goods of cereals and refrigerated and frozen meat. As shown in Panel B, we find positive correlations for cereal cultivated area, particularly in 1895 and especially for wheat. As reported in Panel C, we find strong positive relationships for agricultural machinery, much of which is intensively used for cereals production. As indicated in Panel D, we find especially strong correlations for mixed and pure-breed livestock that are used disproportionately for refrigerated and frozen meat, with these correlations increasing in magnitude over time. Finally, as displayed in Panel E, we find strong positive relationships for railroad shipments including cereals, consistent with most agricultural production

being shipped outside districts for the export market.

Therefore, across this wide range of different observed variables, we find that our model solutions for adjusted-agricultural productivity ($\tilde{z}_{At}(\ell)$) have the expected relationship with separate data closely related to production activity in this traded sector.

7 Counterfactuals

Having quantified the model, we are now in a position to undertake counterfactuals to examine the role of internal geography in shaping structural transformation and economic development in late-19th century Argentina. Internal geography affects levels of economic activity in the model through the two sufficient statistics of adjusted-agricultural productivity ($\tilde{z}_{At}(\ell)$) and non-traded productivity ($z_{Nt}(\ell)$). Based on our earlier reduced-form evidence, we distinguish two key dimensions of internal geography: (i) proximity to Argentina's trade hub and (ii) the expansion of the railroad network. We undertake counterfactuals to examine the impact of these two dimensions of internal geography on the distribution of economic activity through the model's two sufficient statistics.

In our quantitative analysis in the previous section, we inverted the model to solve for the unique values of adjusted-agricultural productivity ($\tilde{z}_{At}(\ell)$) and non-traded productivity ($z_{Nt}(\ell)$) that exactly rationalize the observed data on population density ($n_t(\ell)$) and the agricultural employment share ($\nu_{At}(\ell)$), where we chose the common real wage in each year (u_t^*) to match the estimates of real wage growth for Argentina as a whole from [Taylor and Williamson \(1997\)](#). In contrast, we now assume a counterfactual change in adjusted-agricultural productivity ($\tilde{z}_{At}(\ell)$) and non-traded productivity ($z_{Nt}(\ell)$), and solve for the unique equilibrium values of the model's endogenous variables, including population density ($n_t(\ell)$) and the agricultural employment share ($\nu_{At}(\ell)$). Therefore, the model's counterfactual predictions for these endogenous variables need not equal their observed values in the data.

We undertake these counterfactuals starting from the actual equilibrium observed in the data at the end of our sample period in 1914, which is the year with the largest number of populated districts in our sample. In our first set of counterfactuals, we assume free international migration, which is consistent with the large flows of international migrants observed during our sample period. In this case, the real wage in Argentina is pinned down by its exogenous value in the rest of the world. Therefore, we solve for the counterfactual equilibrium holding the real wage constant at its actual equilibrium value in 1914, and allowing total population in Argentina to adjust to achieve real wage equalization. In our second set of counterfactuals, we assume restricted international migration. In this case, we hold the total population of Argentina constant at its actual equilibrium value in 1914, and solve for the common equilibrium real wage across all Argentina districts consistent with labor market clearing.

7.1 Free International Migration

We start with our first set of counterfactuals under free international migration. We use the recursive structure of the model to obtain a tractable characterization of the counterfactual equilibrium. First, we assume counterfactual sectoral productivities ($\tilde{z}'_{At}(\ell)$, $z'_{Nt}(\ell)$), hold the common real wage ($u_t^{*'} = u_t^*$) constant at its 1914 value, and use the population mobility condition (17) to solve for the counterfactual wage-rental ratio ($\omega'_t(\ell)$):

$$\left[\beta_T (\tilde{z}'_{At}(\ell) \omega'_t(\ell)^{\alpha_A})^{\sigma-1} + (1 - \beta_T) (z'_{Nt}(\ell) \omega'_t(\ell)^{\alpha_N})^{\sigma-1} \right]^{\frac{1}{\sigma-1}} = u_t^{*'}, \quad (37)$$

where we use a prime (\prime) to denote a counterfactual value.

Second, using this solution for the counterfactual wage-rental ratio ($\omega'_t(\ell)$) and equation (35), we immediately recover the counterfactual equilibrium relative price for tradeables ($E'_{Tt}(\ell)/E'_t(\ell)$):

$$\frac{E'_{Tt}(\ell)}{E'_t(\ell)} = \frac{u_t^{*'}}{\tilde{z}'_{At}(\ell)\omega'_t(\ell)^{\alpha_A}}. \quad (38)$$

Third, using these solutions for the counterfactual wage-rental ratio ($\omega'_t(\ell)$) and relative price for tradeables ($E'_{Tt}(\ell)/E'_t(\ell)$), together with equations (23) and (24), we obtain the counterfactual equilibrium values for population density ($n'_t(\ell)$) and the agricultural employment share ($\nu'_{At}(\ell)$):

$$n'_t(\ell) = \left(\frac{1}{\alpha_N + (\alpha_A - \alpha_N) \beta_T \left(\frac{E'_{Tt}(\ell)}{E'_t(\ell)} \right)^{1-\sigma}} - 1 \right) \frac{1}{\omega'_t(\ell)}, \quad (39)$$

$$\nu'_{At}(\ell) = \frac{(1 - \alpha_A) \beta_T \left(\frac{E'_{Tt}(\ell)}{E'_t(\ell)} \right)^{1-\sigma}}{1 - \left(\alpha_N + (\alpha_A - \alpha_N) \beta_T \left(\frac{E'_{Tt}(\ell)}{E'_t(\ell)} \right)^{1-\sigma} \right)}, \quad (40)$$

where we recall that we associate agricultural employment with rural population in the data and non-traded employment with urban population in the data. Therefore, the urban population share is $\nu'_{Nt}(\ell) = 1 - \nu'_{At}(\ell)$.

Finally, from equation (39), summing the counterfactual population density ($n'_t(\ell)$) in each location multiplied by its supply of land ($L(\ell)$), we obtain the counterfactual equilibrium total population for Argentina as a whole:

$$N'_t = \sum_{\ell \in \mathcal{L}} n'_t(\ell) L(\ell) = \sum_{\ell \in \mathcal{L}} \left(\frac{1}{\alpha_N + (\alpha_A - \alpha_N) \beta_T \left(\frac{E'_{Tt}(\ell)}{E'_t(\ell)} \right)^{1-\sigma}} - 1 \right) \frac{L(\ell)}{\omega'_t(\ell)}. \quad (41)$$

7.2 Restricted International Migration

We next consider our second set of counterfactuals with restricted international migration. Again we use the recursive structure of the model to obtain a tractable characterization of the counterfactual equilibrium. We assume counterfactual productivities ($\tilde{z}'_{At}(\ell)$, $z'_{Nt}(\ell)$), and instead of holding the common real wage constant at its 1914 value, we now hold the total population of Argentina as a whole constant at its 1914 value:

$$N'_t = \sum_{\ell \in \mathcal{L}} n'_t(\ell) L(\ell) = N_t(\ell). \quad (42)$$

Using the system of five equations (37), (38), (39), (40) and (42), we solve for the five unknowns of the counterfactual common real wage ($u_t^{*'}), wage-rental ratio ($\omega'_t(\ell)$), relative price for tradeables ($E'_{Tt}(\ell)/E'_t(\ell)$), population density ($n'_t(\ell)$) and agricultural employment share ($\nu'_{At}(\ell)$) for each location, such that the counterfactual equilibrium total population equals the actual 1914 population ($N'_t = N_t$).$

7.3 Counterfactual Changes in Productivities

For both the free and restricted international migration specifications, we explore the role of internal geography in shaping on structural transformation and economic development using three counterfactuals. In our first counterfactual, we examine the overall impact of the changes in the levels and spatial gradients of adjusted-agricultural productivity ($\tilde{z}_{At}(\ell)$) and non-traded productivity ($z_{Nt}(\ell)$), holding all else constant. Using the specification in equation (1)

and the results reported in Table 4, we adjust the 1914 productivities for each sector and location by the estimated change in the intercept and gradient from the nearest top-four port between 1869 and 1914:

$$\begin{aligned}\tilde{z}'_{A,1869}(\ell) &= \left(\frac{a_{A,1869}}{a_{A,1914}}\right) (\text{distport}(\ell)^{b_{A,1869}-b_{A,1914}}) \tilde{z}_{A,1914}(\ell), \\ z'_{N,1869}(\ell) &= \left(\frac{a_{N,1869}}{a_{N,1914}}\right) (\text{distport}(\ell)^{b_{N,1869}-b_{N,1914}}) z_{N,1914}(\ell),\end{aligned}\quad (43)$$

where a_{it} is the estimated intercept and b_{it} is the estimated slope coefficient on distance to the nearest top-four port for sector $i \in A, N$ and year t . These counterfactual productivities differ from the actual productivities in 1869, because equation (43) omits the change in the regression error from equation (1). We thus abstract from idiosyncratic productivity shocks in order to focus on the average pattern of productivity growth across sectors and locations.

In our second counterfactual, we consider the pure impact of the change in the spatial gradients of productivities in each sector, holding all else constant. Again we use the specification in equation (1) and the results reported in Table 4 to adjust the 1914 productivities for each sector and location by the estimated change in the gradient from the nearest top-four port between 1869 and 1914:

$$\begin{aligned}\tilde{z}''_{A,1869}(\ell) &= (\text{distport}(\ell)^{b_{A,1869}-b_{A,1914}}) \tilde{z}_{A,1914}(\ell), \\ z''_{N,1869}(\ell) &= (\text{distport}(\ell)^{b_{N,1869}-b_{N,1914}}) z_{N,1914}(\ell),\end{aligned}\quad (44)$$

where b_{it} is the estimated slope coefficient on distance to the nearest top-four port for sector $i \in A, N$ and year t ; and we use a double prime to denote our second counterfactual. These counterfactual productivities differ from the actual productivities in 1869, because equation (44) omits the change in both the intercept and the regression error from equation (1). The key difference between our first and second counterfactuals is that this second counterfactual focuses solely on the change in the spatial gradient of productivity with respect to distance from the nearest top-four port, and abstracts from changes in the overall level of productivity.

In our third counterfactual, we evaluate the impact of the construction of the railroad network on the spatial distribution of economic activity through our two sufficient statistics of adjusted-agricultural productivity ($\tilde{z}_{At}(\ell)$) and non-traded productivity ($z_{Nt}(\ell)$). In particular, we use our baseline instrumental variables estimates for the impact of the railway network on these sufficient statistics from Table 5, which are based on both our port and colonial post instruments. We evaluate the impact of the construction of the railway network by adjusting the 1914 productivities for each sector and location to remove the predicted impact of railroad access:

$$\begin{aligned}\tilde{z}'''_{A,1869}(\ell) &= (\text{sharerrail}_{1914}(\ell)^{-c_{A,1914}}) \tilde{z}_{A,1914}(\ell), \\ z'''_{N,1869}(\ell) &= (\text{sharerrail}_{1914}(\ell)^{-c_{N,1914}}) z_{N,1914}(\ell),\end{aligned}\quad (45)$$

where $\text{sharerrail}_{1914}(\ell)$ is the length of railroads in each district in 1914 as a percentage of this length for Argentina as a whole; $c_{i,1914}$ is the estimated coefficient on this variable; and we use a triple prime to indicate our third counterfactual. Although this third counterfactual removes the entire railroad network, almost all of this network was constructed after 1869, and hence most of the variation comes from eliminating railroad construction from 1869-1914.

For each counterfactual, we substitute the assumed changes in productivity for each sector and location into the system of equations (37)-(41), using the assumption of either free international migration ($u^{*'} = u^{*''} = u^{*'''} = u^*$) or restricted international migration ($N' = N'' = N''' = N$). For ease of interpretation, we report the results from all these specifications as increases from the counterfactual equilibrium in 1869 to the actual equilibrium in 1914.

7.4 Counterfactual Predictions

In Table 7, we report the results of these counterfactuals. Panel A presents the results under free international migration, while Panel B contains those under restricted international migration. Within each panel, row 1 corresponds to our first counterfactual that changes both the level and gradient of productivity in each sector; row 2 captures our second counterfactual that changes only the gradient of productivity in each sector; and row 3 summarizes our third counterfactual for the removal of the railroad network.

As reported in Columns (1)-(2) of row 1 (both panels), our first counterfactual for the level and gradient of productivity involves an average increase in adjusted-agricultural productivity across all locations of 68 percent, and a corresponding average increase in non-traded productivity of 18 percent. Under our assumption of free international migration (Panel A), the equilibrium common real wage is pinned down by its exogenous value in the rest of the world (Column (4)). Therefore, as these counterfactual increases in productivity create upward pressure on the common real wage, they attract a population inflow from the rest of the world, until the increase in the economy's total population and a diminishing marginal productivity of labor with a fixed supply of land restore the common real wage to this exogenous value in the rest of the world. We find that these counterfactual increases in productivity result in a substantial increase in the total population of Argentina, which rises from 0.59 to 5.8 million between the counterfactual equilibrium in 1869 and the actual equilibrium in 1914 (a log increase of 229 percent, as reported in Column (3)).³⁰ As implied by the spatial Balassa-Samuelson effect in Proposition 3, the larger increase in adjusted-agricultural productivity than in non-traded productivity leads to structural transformation away from agriculture under our assumption of inelastic demand between sectors. We find a counterfactual increase in the mean urban population share of 6 percent in Column (5), which is somewhat smaller than the actual increase in the mean urban population share of 11 percent between 1869 and 1915.

Under our assumption of restricted international migration (Panel B), the total population of Argentina is held constant at its value in the actual equilibrium in 1914 (Column (3)). Given this inelastic supply of labor, the increases in productivity in each sector in our first counterfactual lead to a substantial increase in the common real wage of 35 percent, as reported in Column (4). Again the spatial Balassa-Samuelson effect from Proposition 3 implies that the larger increase in adjusted-agricultural productivity than in non-traded productivity leads to a reallocation of employment away from agriculture with inelastic demand between sectors. We find that the resulting increase in the mean urban population share of 12 percent in Column (5) is larger than both the counterfactual increase with an endogenous population (6 percent) and the actual increase in the data from 1869-1914 (11 percent).

As reported in Columns (1)-(2) of row 2 (both panels), our second counterfactual involves an average increase across all locations of 101 percent for adjusted-agricultural productivity, and a corresponding average decrease in non-traded productivity of -71 percent. This pattern reflects the steepening of the gradients in distance to the nearest top-four port for adjusted-agricultural productivity and non-traded productivity reported in Table 4 above. As a result, when we replace the more negative gradient in adjusted-agricultural productivity in 1914 with the less negative gradient in 1869, we raise the adjusted-agricultural productivity of interior regions, and hence increase average adjusted-productivity in this sector. In contrast, when we replace the more positive gradient in non-traded produc-

³⁰This counterfactual increase is somewhat larger than the actual increase in the total population of Argentina between 1869 and 1914 from 1.02 to 5.8 million (a log growth of 175 percent). As discussed above, this difference reflects our abstraction in these counterfactuals from idiosyncratic shocks to productivity in each sector for individual locations and the assumption of free international migration.

tivity in 1914 with the less positive gradient in 1869, we reduce the non-traded productivity of interior regions, and hence decrease average productivity in this sector.

Table 7: Counterfactuals

	(1)	(2)	(3)	(4)	(5)
	Mean Log Growth Adjusted- Agricultural Productivity	Mean Log Growth Non- Traded Productivity	Log Aggregate Population Growth	Log Common Real Wage Growth	Mean Actual Divided by Mean Counterfactual Urban Population Share
Panel A : Free International Migration (Endogenous Population)					
1. Both Constant and Gradient	0.68	0.18	2.29	0	1.06
2. Only Gradient	1.01	-0.71	2.71	0	1.42
3. Railroad Access	0.15	0.03	0.49	0	1.02
Panel B : Restricted International Migration (Exogenous Population)					
1. Both Constant and Gradient	0.68	0.18	0	1.35	1.12
2. Only Gradient	1.01	-0.71	0	1.6	1.5
3. Railroad Access	0.15	0.03	0	1.08	1.03

Notes: Columns (1) and (2) report means of log productivity growth across Argentinian districts; Columns (3) and (4) report log changes in Argentina's total population and the common real wage; Column (5) reports the mean across Argentinian districts of the urban population share in the counterfactual equilibrium divided by that in the actual equilibrium. Free international migration in Panel A changes adjusted-agricultural productivity and non-traded productivity, holding the common real wage constant at its 1914 value, and allowing total population in Argentina to adjust. Restricted international migration in Panel B changes adjusted-agricultural productivity and non-traded productivity, holding total population in Argentina constant at its 1914 value, and allowing the common real wage to adjust. Row 1 undertakes a counterfactual in which the intercept and gradient with respect to distance from the nearest top-four port are set equal to their 1869 value rather than their 1914 value, and reports the increase from the counterfactual 1869 equilibrium to the actual 1914 equilibrium. Row 2 undertakes a counterfactual in which the gradient with respect to distance from the nearest top-four port is set equal to its 1869 value rather than its 1914 value, and reports the increase from the counterfactual 1869 equilibrium to the actual 1914 equilibrium. Row 3 undertakes a counterfactual for the removal of the railroad network, and reports the increase from the counterfactual equilibrium without the railroad network to the actual 1914 equilibrium. In row 2, the estimates for the intercept and gradient for adjusted-agricultural productivity and non-traded productivity are taken from Table 4. In row 3, the instrumental variables estimates for the impact of the railroad network using both the port and colonial post instruments are taken from Table 5. Columns (1)-(2) report the mean log growth in productivity across Argentine districts in each counterfactual. Column (5) reports the mean urban share in the actual 1914 equilibrium divided by the mean urban share in the counterfactual equilibrium.

Consistent with the change in gradients in Table 4 being larger for adjusted-agricultural productivity than for non-traded productivity, we find net positive effects of these counterfactual changes in productivity on either the total population of Argentina (with free international migration in Panel A) or the common real wage (with restricted international migration in Panel B). Furthermore, the impact of the change in the spatial gradients of productivity in this second counterfactual (row 2) is large relative to the effect of the change in both the level and spatial gradient of productivity in our first counterfactual (row 1). With free international migration (Panel A), flattening the spatial gradients in productivity raises the total population of Argentina by 271 percent (Column (3)) and the mean urban population share by 42 percent (Column (5)). In comparison, with restricted international migration (Panel B), diminishing these spatial disparities in productivity raises the common real wage by 60 percent (Column (4)) and the mean urban population share by 50 percent (Column (5)). Therefore, holding the spatial gradients in productivity unchanged at their 1869 values, and thereby allowing the interior regions of Argentina to experience the same changes in productivity from 1869-1914 as those regions proximate to Argentina's trade hub, has quantitatively large effects on aggregate economic outcomes in the form of total population or the common real wage.

Although this second counterfactual highlights the quantitative relevance of the change in the spatial gradients of productivities within Argentina for aggregate economic outcomes, these spatial gradients capture a range of factors (including compositional differences within the agricultural sector), and some of these factors could be difficult to

change through feasible interventions. Therefore, in our third counterfactual, we consider one source of these spatial gradients that is amenable to change through realistic interventions, namely the construction of the railroad network. Our instrumental variables estimates for the impact of the railway network in Columns (2) and (4) of Table 5 imply smaller counterfactual changes in adjusted-agricultural productivity and non-traded productivity than the changes in levels and spatial gradients of productivity considered in our first two counterfactuals. As reported in Columns (1)-(2) of row 3 (both panels), these instrumental variables estimates imply an average increase in adjusted-agricultural and non-traded productivity of 15 and 3 percent respectively, where the larger impact for the agricultural sector is consistent with the railroad network reducing transport costs.

Despite these smaller magnitudes, we find substantial effects of the construction of the railroad network. Under our assumption of free international migration (Panel A), we find that the total population of Argentina increases 49 percent (Column (3)) and the mean urban population share rises by 2 percent (Column (5)). Alternatively, under our assumption of restricted international migration (Panel B), the common real wage increases by 8 percent (Column (4)), and the mean urban population share rises by 3 percent (Column (5)). While our earlier reduced-form estimates implied substantial impacts of the railroad network on urban and rural population growth, these earlier “difference-in-differences” specifications capture relative comparisons between locations receiving more versus less transportation infrastructure. Therefore, they cannot capture general equilibrium effects or distinguish reallocation from the creation of economic activity, as discussed in [Redding and Turner \(2015\)](#). In contrast, these counterfactual predictions in Table 7 reveal substantial general equilibrium effects of the railroad network on aggregate economic outcomes for Argentina as a whole, as measured by either aggregate population or the common real wage across all locations.

To provide a point of comparison, [Fogel \(1964\)](#) estimates that the social saving from railroads in the agricultural sector was no more than 2.7 percent of gross national product (GNP) in the United States in 1890; [Donaldson and Hornbeck \(2016\)](#) estimate that the reduction in market access from removing the 1890 railroad network in the United States would have decreased agricultural land values by 3.22 percent of GNP; and [Donaldson \(2018\)](#) estimates an impact of railroad access on real agricultural income of around 16 percent. While all of these studies focus on the agricultural sector, our estimates capture the economy-wide impact of the railroad network on urban and rural economic activity, which provides a natural explanation for our real income estimate of 8 percent being typically larger than these other estimates. Furthermore, another implication of our results is that the railroad network can have indirect effects on aggregate population and economic activity through international migration. In our specification with free international migration, even the modest changes in productivity in each sector from the construction of the railroad network result in a more than three times larger increase in total population.

Taken together, these findings are consistent with the view that the construction of the railroad network in 19th-century Argentina played an important role in improving the access of interior regions to world markets, thereby increasing their price-adjusted productivity in the traded sector. Our instrumental variables estimates for the impact of the railroad network on price-adjusted productivity imply a substantial increase in the common real wage or aggregate population, depending on assumptions about international migration. Finally, the larger impact of the railroad network on adjusted agricultural productivity than on non-traded productivity contributed to structural transformation away from agriculture and the observed rise in the urban population share.

8 Conclusions

We provide new theory and evidence on the relationship between economic development and international trade using Argentina’s late-19th-century integration into the global economy. We combine the natural experiment from reductions in transportation costs from the invention of steam ships and railroads, disaggregated data on economic activity across sectors and regions within Argentina, and a quantitative general equilibrium model.

We begin by showing that Argentina’s rapid export-led economic development in the late-19th century involved major changes in the distribution of economic activity across sectors and regions. First, we show that population density is sharply decreasing in distance from Argentina’s trade hub in Buenos Aires and its surrounding ports. Second, we establish that this gradient is steeper for urban population density than for rural population density, so that the areas closest to world markets are more urbanized. Third, we find that this gradient in population density steepens over our sample period, as economic activity expands in the immediate hinterland of Buenos Aires. Fourth, we show that exogenous variation in access to railroads predicted by our instruments increases both urban and rural population density for a given geographical distance from Argentina’s trade hub. Fifth, we demonstrate that both proximity to Argentina’s trade hub and access to railroads induce compositional changes within the agricultural sector, away from Argentina’s traditional comparative advantage products of tanned hides and leather, and towards its new export goods of cereals and refrigerated and frozen meat.

We rationalize these empirical findings by developing a theoretical model that emphasizes the interaction between structural transformation across sectors and internal trade costs across regions. We make the natural assumptions that all Argentinian locations have a comparative advantage in agriculture, demand is inelastic between sectors, and agriculture is land intensive. Under these assumptions, we show that our general neoclassical production structure implies a *spatial Balassa-Samuelson effect*, such that regions with good access to world markets have higher population densities, urban population shares, relative prices of non-traded goods, and land prices relative to wages. The intuition is that locations with good access to world markets are attractive for the production and consumption of traded goods, which increases population density, and bids up the reward of the immobile factor (land) relative to the mobile factor (labor). Together the increase in population and the reduction in wages relative to land rents induce an expansion in the employment share of the labor-intensive non-traded sector, which requires a higher relative price for the non-traded good, given inelastic demand between sectors.

We show that our theoretical framework can account for our empirical findings, not only qualitatively, but also quantitatively. We invert the model to determine unique values of adjusted-agricultural and non-traded productivity for each location, which are sufficient statistics for the distribution of economic activity. We establish a steepening of the gradient in these productivities with respect to distance from Argentina’s trade hub over our sample period, which is particularly large for adjusted-agricultural productivity. We show that the construction of the railroad network predicted by our instruments has a statistically significant positive effect on adjusted-agricultural productivity, but not on non-traded productivity, which is consistent with railroads reducing internal transport costs. In counterfactuals, we demonstrate that these changes in the spatial gradients of productivity are consequential for aggregate economic outcomes. We estimate that the reduction in internal transport costs from the construction of the railroad network either raises the total population of Argentina by 49 percent under free international migration or raises the common real wage across all Argentinian districts by 8 percent under restricted migration.

References

- ADELMAN, J. (1994): *Frontier Development: Land, Labor and Capital on the Wheatlands of Argentina and Canada 1890-1914*. Oxford University Press, Oxford.
- ALLEN, T., AND C. ARKOLAKIS (2014): “Trade and the Topography of the Spatial Economy,” *Quarterly Journal of Economics*, 129(3), 1085–1140.
- ATKIN, D., AND D. DONALDSON (2015): “Who’s Getting Globalized? The Size and Nature of Intranational Trade Costs,” *NBER Working Paper*, 21439.
- BANERJEE, A., E. DUFLO, AND N. QIAN (2012): “On the Road: Transportation Infrastructure and Economic Development,” *NBER Working Paper*, 17897.
- BAUM-SNOW, N. (2007): “Did Highways Cause Suburbanization?,” *Quarterly Journal of Economics*, 122(2), 775–780.
- BAUM-SNOW, N., L. BRANDT, J. HENDERSON, M. A. TURNER, AND Q. ZHANG (2017): “Roads, Railroads and Decentralization of Chinese Cities,” *Review of Economics and Statistics*, 99(2), 435–448.
- BENNETT, M. K. (1933): “World Wheat Crops, 1885-1932: New Series, with Areas and Yields, by Countries,” *Wheat Studies*, 9(7), Stanford: Food Research Institute.
- BERLINSKI, SAMUEL, S. G., AND L. JAITMAN (2011): “Impact of Railways: Expansion on Spatial Population Patterns,” *Lincoln Institute of Land Policy Working Paper*, WP11SB1.
- BORDO, M., B. EICHENGREEN, D. KLINGEBIEL, AND S. M. M. PERIA (2001): “Is the Crisis Problem Growing More Severe?,” *Economic Policy*, 24, 51–82.
- BUSTOS, P., G. GARBER, AND J. PONTICELLI (2017): “Capital Accumulation and Structural Transformation,” CEMFI, mimeograph.
- CACOPARDO, M. C. (1967): *Republica Argentina: Cambios en Los Limites Nacionales, Provinciales y Departmentales, A Traves de Los Censos Nacionales de Poblacion*. Instituto Torcuato Di Tella, Buenos Aires.
- CALIENDO, L., F. PARRO, E. ROSSI-HANSBERG, AND P.-D. SARTE (2018): “The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy,” *Review of Economic Studies*, forthcoming.
- CARTER, S. B., S. S. GARTNER, M. R. HAINES, A. L. OLNSTEAD, R. SUTCH, AND G. WRIGHT (2006): *Historical Statistics of the United States Millennial Edition*. Cambridge University Press, Cambridge.
- CASELLI, F., AND W. J. I. COLEMAN (2001): “The U.S. Structural Transformation and Regional Convergence: A Reinterpretation,” *Journal of Political Economy*, 109(3), 584–616.
- CHANDRA, A., AND E. THOMPSON (2000): “Does Public Infrastructure affect Economic Activity? Evidence from the Rural Interstate Highway System,” *Regional Science and Urban Economics*, 30(4), 457–490.
- COMPAÑÍA SUD-AMERICANA DE BILLETES DE BANCO (1895): *Anuario de la Dirección General de Estadística Correspondiente al Año 1895*. Compañía Sud-Americana de Billetes de Banco, Buenos Aires.

- (1914): *Anuario de la Dirección General de Estadística Correspondiente al Año 1895*. Compañía Sud-Americana de Billetes de Banco, Buenos Aires.
- CORTÉS CONDE, R. (1993): “The Growth of the Argentine Economy, 1870-1914,” in *Argentina Since Independence*, ed. by L. Bethell, vol. 1, chap. 2. Cambridge University Press, Cambridge.
- COŞAR, K., AND B. DEMİR (2016): “Domestic Road Infrastructure and International Trade: Evidence from Turkey,” *Journal of Development Economics*, 118, 232–244.
- COŞAR, K., AND P. FAJGELBAUM (2012): “Internal Geography, International Trade, and Regional Specialization,” *American Economic Journal: Microeconomics*, 8(1), 24–56.
- DAVIS, D. R., AND J. DINGEL (2012): “A Spatial Knowledge Economy,” NBER Working Paper, 18188.
- DAVIS, D. R., AND D. E. WEINSTEIN (2002): “Bones, Bombs, and Break Points: The Geography of Economic Activity,” *American Economic Review*, 92(5), 1269–1289.
- DEL COMERCIO EXTERIOR, C. G. (1870): *Estadística Comercial de la República Argentina*. Cuadro General del Comercio Exterior, Buenos Aires.
- DELLA PAOLERA, G. (1988): “How the Argentine Economy Peformed during the International Gold Standard: A Reexamination,” University of Chicago, PhD Thesis.
- DIRECCION GENERAL DE FERROCARRILES (1895): *Estadísticas de los Ferrocarriles en Explotacion*. Compania Sudamericana de Billetes de Banco, Buenos Aires.
- (1914): *Estadísticas de los Ferrocarriles en Explotacion*. Compania Sudamericana de Billetes de Banco, Buenos Aires.
- DONALDSON, D. (2018): “Railroads of the Raj: Estimating the Impact of Transportation Infrastructure,” *American Economic Review*, 108(4-5), 899–934.
- DONALDSON, D., AND R. HORNBECK (2016): “Railroads and American Economic Growth: A Market Access Approach,” *Quarterly Journal of Economics*, 131(2), 799–858.
- DROLLER, F. (2018): “Migration and Long-run Development: Evidence from Settlements in the Pampas,” *Economic Journal*, 128(614), 2321–2352.
- DURANTON, G., P. MORROW, AND M. A. TURNER (2014): “Roads and Trade: Evidence from the US,” *Review of Economic Studies*, 81(2), 681–724.
- DURANTON, G., AND M. A. TURNER (2012): “Urban Growth and Transportation,” *Review of Economic Studies*, 79(4), 1407–1440.
- EATON, J., AND S. KORTUM (2002): “Technology, Geography, and Trade,” *Econometrica*, 70(5), 1741–1779.
- ECKERT, F., AND M. PETERS (2018): “Spatial Structural Change,” Yale University, mimeograph.

- FABER, B. (2014): "Integration and the Periphery: The Unintended Effects of New Highways in a Developing Country," *Review of Economic Studies*, 81(3), 1046–1070.
- FOGEL, R. W. (1964): *Railroads and American Economic Growth: Essays in Econometric History*. Johns Hopkins University Press, Baltimore, MD.
- FOSTER, A. D., AND M. R. ROSENZWEIG (2007): "Economic Development and the Decline of Agricultural Employment," in *Handbook of Development Economics*, ed. by T. P. Schultz, and J. A. Strauss, vol. 4, pp. 3051–3083. Elsevier, North Holland, Amsterdam.
- FUJITA, M., P. KRUGMAN, AND A. J. VENABLES (1999): *The Spatial Economy: Cities, Regions, and International Trade*. MIT Press, Cambridge MA.
- GOLLIN, D., R. JEDWAB, AND D. VOLLRATH (2016): "Urbanization With and Without Industrialization," *Journal of Economic Growth*, 21(1), 35–70.
- GOLLIN, D., AND R. ROGERSON (2014): "Productivity, Transport Costs and Subsistence Agriculture," *Journal of Development Economics*, 107, 38–48.
- HANSON, G. H. (1996): "Localization Economies, Vertical Organization, and Trade," *American Economic Review*, 86(5), 1266–1278.
- HARLEY, C. K. (1980): "Transportation, World Wheat Trade, and the Kuznets Cycle, 1850-1913," *Explorations in Economic History*, 17, 218–250.
- (1988): "Ocean Freight Rates and Productivity, 1740-1913: The Primacy of Mechanical Invention Reaffirmed," *The Journal of Economic History*, 48(4), 851–876.
- HERRENDORF, B., J. A. SCHMITZ, AND A. TEIXEIRA (2012): "Transportation and Development: Insights from the U.S., 1840-1860," *International Economic Review*, 53(3), 693–716.
- INTER-AMERICAN DEVELOPMENT BANK (2013): *Too Far To Export: Domestic Transport Costs and Regional Export Disparities in Latin America and the Caribbean*. IADB, Washington DC.
- KARÁDI, P., AND M. KOREN (2018): "Cattle, Steaks and Restaurants: Development Accounting when Space Matters," Central European University, mimeograph.
- LAGAKOS, D., AND M. E. WAUGH (2013): "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economic Review*, 103(2), 948–980.
- LALL, S. V., AND M. LEBRAND (2018): "Who Wins, Who Loses? Understanding the Spatially Differentiated Effects of Belt and Road Within Central Asia," World Bank, mimeograph.
- LEWIS, C. M. (1983): *British Railways in Argentina 1857-1914: A Case Study of Foreign Investment*. Institute for Latin American Studies, London.
- MARTINCUS, CHRISTIAN VOLPE, J. C., AND A. CUSOLITO (2017): "Routes, Exports, and Employment in Developing Countries: Following the Trace of the Inca Roads," *Journal of Development Economics*, 125, 21–39.

- MATSUYAMA, K. (1992): "Agricultural Productivity, Comparative Advantage and Economic Growth," *Journal of Economic Theory*, 58, 317–334.
- (2009): "Structural Change in an Interdependent World: A Global View of Manufacturing Decline," *Journal of the European Economic Association*, 7, 478–486.
- MCMILLAN, M., D. RODRIK, AND C. SEPULVEDA (2017): "Structural Change, Fundamentals and Growth: A Framework and Case Studies," *NBER Working Paper*, 23378.
- MICHAELS, G. (2008): "The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System," *Review of Economics and Statistics*, 90(4), 683–701.
- MICHAELS, G., F. RAUCH, AND S. J. REDDING (2012): "Urbanization and Structural Transformation," *Quarterly Journal of Economics*, 127, 535–586.
- MORTEN, M., AND J. OLIVEIRA (2017): "The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City," *NBER Working Paper*, 22158.
- NAGY, D. (2017): "City Location and Economic Development," CREI, mimeograph.
- NGAI, R., AND C. PISSARIDES (2007): "Structural Change in a Multisector Model of Growth," *American Economic Review*, 97(1), 429–443.
- NORTH, D. (1958): "Freight Rates and Economic Development 1750-1913," *Journal of Economic History*, 18(4), 537–555.
- O'ROURKE, K. H., AND J. G. WILLIAMSON (1999): *Globalization and History: The Evolution of a Nineteenth-Century Atlantic Economy*. MIT Press, Cambridge MA.
- PASCALI, L. (2017): "The Wind of Change: Maritime Technology, Trade and Economic Development," *American Economic Review*, 107(9), 2821–54.
- RAMONDO, N., A. RODRIGUEZ-CLARE, AND M. SABORIO (2016): "Scale Effects and Productivity: Does Country Size Matter?," *American Economic Review*, 106(10), 3159–84.
- RANDLE, P. H. (1981): *Atlas del Desarrollo Territorial de la Argentina*. Oikos, Buenos Aires.
- REDDING, S. J. (2016): "Goods Trade, Factor Mobility and Welfare," *Journal of International Economics*, 101, 148–167.
- REDDING, S. J., AND D. M. STURM (2008): "The Costs of Remoteness: Evidence from German Division and Reunification," *American Economic Review*, 98(5), 1766–1797.
- REDDING, S. J., AND M. A. TURNER (2015): "Transportation Costs and the Spatial Organization of Economic Activity," in *Handbook of Regional and Urban Economics*, ed. by G. Duranton, J. V. Henderson, and W. Strange, chap. 20, pp. 1339–1398. Elsevier, Amsterdam.
- REPÚBLICA ARGENTINA (1869): *Primer Censo de la República Argentina*. Superintendente del Censo, Buenos Aires.
- (1895): *Segundo Censo de la República Argentina*. Comisión Directiva, Buenos Aires.

- (1914): *Tercer Censo Nacional*. Comisión Nacional, Buenos Aires.
- ROCCHI, F. (2008): *Chimneys in the Desert: Industrialization in Argentina During the Export Boom Years 1870-1930*. Stanford University Press, Stanford.
- SCOBIE, J. R. (1971): *Argentina: A City and a Nation*. Oxford University Press, New York.
- SOTELO, S. (2018): “Domestic Trade Frictions and Agriculture,” University of Michigan, mimeograph.
- STOREYGARD, A. (2016): “Farther on Down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa,” *Review of Economic Studies*, 83, 1263–1295.
- SYRQUIN, M. (1988): “Patterns of Structural Change,” in *Handbook of Development Economics*, ed. by H. Chenery, and T. N. Srinivasan, vol. 1, chap. 7, pp. 203–273. Elsevier, North Holland, Amsterdam.
- TAYLOR, A. (1992): “External Dependence, Demographic Burdens, and Argentine Economic Decline After the Belle Époque,” *Journal of Economic History*, 52(4), 907–936.
- TAYLOR, A., AND J. G. WILLIAMSON (1997): “Convergence in the Age of Mass Migration,” *European Review of Economic History*, 1, 27–63.
- UY, T., K.-M. YI, AND J. ZHANG (2012): “Structural Change in an Open Economy,” *Journal of Monetary Economics*, 60(6), 667–682.
- VON THÜNEN, J. H. (1826): *Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie*. Perthes, Hamburg.