NBER WORKING PAPER SERIES

EXTERNAL INTEGRATION, STRUCTURAL TRANSFORMATION AND ECONOMIC DEVELOPMENT: EVIDENCE FROM ARGENTINA 1870-1914

Pablo Fajgelbaum
Stephen J. Redding

Working Paper 20217
http://www.nber.org/papers/w20217

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2014

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 Jeremy Adelman, Arnaud Costinot, Dave Donaldson, Jonathan Eaton, Pablo Gerchunoff, Gordon Hanson, Sam Kortum, Richard Rogerson, Ferdinand Rauch, Esteban Rossi-Hansberg, Alan Taylor, Mike Waugh and seminar participants at AEA, Brown, Columbia, EIEF, LACEA, NYU, Paris, Penn State, SED, UC Berkeley, UCLA, UTDT, Wisconsin and Yale for helpful comments. Thanks to Chris Meissner and Alan Taylor for sharing data. Responsibility for results, opinions and errors is the authors' alone. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

External Integration, Structural Transformation and Economic Development: Evidence from Argentina 1870-1914
Pablo Fajgelbaum and Stephen J. Redding
NBER Working Paper No. 20217
June 2014
JEL No. F11,F14,O13,O14

ABSTRACT

This paper uses the natural experiment of Argentina's integration into world markets in the late-nineteenth century to provide evidence on the role of internal geography in shaping the effects of external integration. We develop a quantitative model of the distribution of economic activity across regions and sectors. The model predicts a spatial Balassa-Samuelson effect, in which locations with better access to world markets have higher population densities, higher shares of employment in the non-traded sector, higher relative prices of non-traded goods, and higher land prices relative to wages. We use the model and data on population density and sectoral employment shares to recover sufficient statistics that isolate the economic mechanisms through which external and internal integration affect economic development. Our analysis highlights the role of complementary investments in internal infrastructure and technology adoption in mediating the economy's response to external integration.

Pablo Fajgelbaum
Department of Economics
University of California, Los Angeles
8283 Bunche Hall
Los Angeles, CA 90095
and NBER
pfajgelbaum@gmail.com

Stephen J. Redding
Department of Economics
and Woodrow Wilson School
Princeton University
Fisher Hall
Princeton, NJ 08544
and NBER
reddings@princeton.edu

1 Introduction

External economic integration is often argued to be an important driver of economic development, as it raises income through specialization in comparative-advantage sectors, provides low-cost access to imported goods, and shapes the pattern of structural transformation from agricultural into non-agricultural activities. These relationships are typically examined at the aggregate level, implicitly treating each country as a point in space. In reality, however, countries differ substantially in terms of their internal geography, and internal trade costs hamper the ability of interior regions to participate in world markets. How important is geographic heterogeneity within countries for the aggregate impact of external integration?

In policy circles, the role of domestic infrastructure in enabling countries to participate in world markets has received renewed attention, and a growing body of evidence suggests that internal trade costs can be large.[1] Yet there is little quantitative evidence on the role of internal trade costs in shaping the effects of external integration on the pattern of economic development and welfare within countries. This scarcity of evidence reflects in part both empirical and theoretical challenges. Empirically, it is difficult to find large-scale changes in both external and internal integration. Theoretically, to explore the relationship between them, we require a general-equilibrium model that can incorporate trade within and between countries, structural change across sectors as a key part of the development process, and factor mobility across space. At the same time, we require this model to remain sufficiently tractable as to be amenable to quantitative analysis.

In this paper, we address both of these challenges. Empirically, we make use of the natural experiment provided by Argentina's integration into world markets in the late-nineteenth century. This large-scale increase in external integration was driven by a cluster of related technological innovations that reduced international transport costs. Increases in the size of ships and the spread of steam navigation made it profitable to ship wheat, corn, and other cereals from Argentina to European markets. New technologies such as meat refrigeration, first invented in Australia in the 1860s to serve British colonial markets, made it possible for the first time to trade frozen and chilled meat from Argentina to Europe. These reductions in external transport costs propelled an export boom and an "agricultural revolution on the pampas," as Argentina's traditional production of animal hides, salted meat and wool was progressively replaced by specialization in the new comparative-advantage products of cereals and frozen and chilled beef. This increase in external trading opportunities stimulated an expansion of the railroad network to connect the agricultural hinterland with ports such as Buenos Aires, mass immigration that enlarged the labor force, and increased imports of manufacturing goods.[2]

---

[1] See, for example, Inter-American Development Bank (2013), United Nations Economic and Social Council and United Nations Economic Commission for Africa (2009), World Bank (2009, 2011), and World Trade Organization (2004). Limao and Venables (2001) find that improving a country's infrastructure from the median to top 25 percent of countries would reduce its transport costs by the equivalent of 481 kilometers of overland travel and 3,989 kilometres of travel by sea. Atkin and Donaldson (2012) provide further evidence on the importance of domestic trade costs within several developing countries.

[2] For further discussion of the history of Argentine economic development, see for example Adelman (1994), Amaral (2002), Cortes Conde (1993), Scobie (1971) and Taylor (1992, 1997).

Theoretically, we develop a tractable quantitative general equilibrium model that determines the distribution of economic activity across both regions and sectors. Both of these distributions are central to understanding the process of economic development, as this process is typically characterized by growing urbanization and spatial inequalities as well as structural transformation from agriculture to non-agriculture. We model an economy with many locations (districts in our data), some of which are interior, and others of which are coastal. Some coastal or riverine locations are directly connected to world markets (through ports) while interior locations are connected to world markets through an internal transportation network that can change over time, as for example railroads are constructed. The model includes two tradable sectors (Agriculture and Manufacturing) and a single non-tradable sector (Services or Manufacturing that is only supplied to the local market). Each of these sectors uses land and labor as factors of production with different intensities. To make contact with disaggregated data on trade and land use across goods within sectors, each sector may in turn comprise several goods. Labor and land in each location is endogenously allocated across Agriculture (produced in rural areas) and Manufacturing and Services (produced in urban areas). Workers are mobile across regions and choose their location to arbitrage away real wage differences.

We first characterize the key analytic properties of the framework. We identify a simple general-equilibrium channel for the impact of internal geography on the pattern of development that we term the spatial Balassa-Samuelson effect. At the country level, the Balassa-Samuelson effect refers to the fact that nontraded factors of production are more expensive in more productive countries. In our model of internal geography within countries, the *spatial Balassa-Samuelson effect* implies that locations with low trade costs to international markets, such as regions close to ports or railway lines, feature a high relative price in the nontraded sector and high land rents relative to wages. These differences in relative prices govern the pattern of economic development. As long as traded and nontraded goods are complements in final consumption, the high relative price of nontraded activities in well-connected locations drives large shares of employment in the non-traded sector. In turn, because labor is cheap relative to land and sectorial specialization is biased toward the labor-intensive nontraded activities, output in well-connected locations is produced with labor-intensive techniques, leading to high labor density. Thus, the model offers a unifying rationale for patterns of spatial development within countries observed in the historical Argentinean data and today in developing countries: Proximity to trade hubs is associated with high employment density, high land rents relative to wages, and structural transformation away from agriculture.

We next provide empirical evidence on these predictions of the model. An advantage of our empirical setting is the availability of rich spatially-disaggregated data for a long historical time period characterized by large-scale changes in external integration. We combine historical censuses, official trade statistics, and railway records, among other sources, to assemble a new dataset on rural and urban employment, specialization patterns across agricultural goods, and railway shipments of these goods for 386 Argentinian districts from 1869-1914. Over this period, real exports and imports increased by more than 500 and 200 percent respectively, with agricultural and livestock products accounting for more than 95 percent of the total value of exports in each year. Total population increased from 1.8 to 7.9 million, and real income per capita grew so rapidly that Argentina became the eighth richest country in the world by 1914.

We find strong empirical confirmation of the spatial Balassa-Samuelson effect in both the cross-section and time-series. At the beginning of our sample period, population density and the urban population share are both sharply decreasing in measures of geographical remoteness from world markets. Over time, despite the fall in internal trade costs from the construction of the railroad network, there is a *steepening* of the gradients of both population density and the urban population share with respect to remoteness. We show how the structure of the model can be used together with observed data on population density and the urban population share to recover two sufficient statistics for the distribution of economic activity for each location: (a) productivity in the export sector (agriculture) adjusted by the tradables consumption price index (including imported manufactures) and (b) productivity in non-tradables. Together with the population of the economy as a whole and geographical land area for each location, these sufficient statistics determine all of the model's endogenous variables, including the relative price of tradables and the relative wage-rental ratio for each location, as well as population density and sectoral specialization for each location.

We use the structure of the model to undertake counterfactuals for changes in each of these sufficient statistics and quantify their role in explaining changes in aggregate welfare and the distribution of economic activity. Reducing Argentina's frontiers from 1914 to 1869 boundaries, while holding the economy's total population and productivity in all other locations constant, reduces real wages to 93 percent of 1914 values, as less of the immobile factor land is available per person. In contrast, reducing both Argentina's frontiers and its total population from 1914 to 1869 values, while holding productivity in all other locations constant, increases real wages to 103 percent of 1914 values, as the reduction in population dominates the reduction in land area.

In comparison, changing adjusted agricultural productivity and non-agricultural productivity for each location has effects that are large relative to those for total land area and population. Adjusting frontiers, total population and all productivities to 1869 values reduces real wages to 62 percent of 1914 values. Adjusting frontiers, total population and only adjusted productivities in agriculture to 1869 values reduces real wages to 92 percent of 1914 values, an effect that is around the same magnitude as for total population.

We use these counterfactuals to show the role of adjusted agricultural and non-agricultural productivities in understanding changes in the internal distribution of economic activity across regions and sectors. The combination of an agricultural export boom and increased urbanization is explained in the model by faster growth in adjusted agricultural productivity than in non-agricultural productivity, which with inelastic demand between sectors reallocates employment away from agriculture (produced in rural areas) towards non-tradables (produced in urban areas). The steepening of the gradients of population density and the urban population share with respect to remoteness is rationalized in the model by a steepening of the productivity gradients, which is particularly marked for adjusted agricultural productivities.

Finally, we provide evidence on the economic mechanisms underlying the changes in adjusted agricultural productivities. In the model, the gradient of adjusted agricultural productivity depends on the rate at which the prices of exported agricultural goods decline with remoteness, the rate at which the prices of imported agricultural goods rise with remoteness, and the spatial distribution of the technologies for producing agricultural goods. Other things equal, the expansion of the railroad network might be expected to flatten the gradient of adjusted agricultural productivity, both by reducing internal trade costs and facilitating the diffusion of technology. However, both the expansion of the railroad network and the adoption of agricultural machinery are themselves geographically uneven. We find that a substantial component of the steepening of the gradient of adjusted agricultural productivity can be explained statistically by these internal investments.

To address the non-random assignment of railroads and agricultural machinery, we use an instrument based on the idea that locations can be treated with transport infrastructure, not because of their own unobserved characteristics, but because they happen to lie along the route between other locations (see Chandra and Thompson 2000 and Michaels 2008). After controlling for initial levels of development and geographical remoteness, we show that locations along the shortest route from the centroids of districts to 16th-century cities are more likely to obtain railroad connections, which in turn stimulates the adoption of agricultural machinery. We find that this source of quasi-experimental variation leads to large changes in both these internal investments and adjusted agricultural productivity.

Our paper is related to a number of literatures. Our use of the natural experiment of Argentina's integration into world markets in the late-nineteenth century relates to a small number of other studies that have used natural experiments in trade. Bernhofen and Brown (2004, 2005) examine Japan's opening in the nineteenth century but are not primarily concerned with the internal distribution of economic activity within Japan; Davis and Weinstein (2002, 2005) exploit the large-scale bombing of Japanese cities during the Second World War; Hanson (1996a,b) considers the implications of Mexican trade liberalization for the spatial distribution of employment and wages; Redding and Sturm (2008) investigate the impact of Germany's division in the aftermath of the Second World War on the distribution of population across West German cities. None of these papers examines the relationship between external integration, structural transformation and economic development.

Our paper also connects with the theoretical literature on new economic geography, as synthesized in Fujita, Krugman and Venables (1999). The complexity of these models typically restricts attention to stylized examples assuming symmetry and/or a handful of regions. Nevertheless, a small number of papers have recently begun to develop quantitative models of trade with endogenous internal distributions of economic activity, including Allen and Arkolakis (2013), Caliendo et al. (2013), Coşar and Fajgelbaum (2012), Ramondo et al. (2012) and Redding (2012). Theoretically, a key distinguishing feature of our model is the relationship between internal trade costs and structural transformation across sectors through the spatial Balassa-Samuelson effect. Empirically, in contrast to these papers, our analysis is geared toward the quantitative analysis of a natural experiment involving large-scale changes in external and internal trade costs and geographic reallocation of labor.

Our research is also related to the literature on transport infrastructure investments, including Banerjee et al. (2012), Baum Snow (2007), Berlinski et al. (2011), Chandra and Thompson (2000), Coşar and Demir (2014), Donaldson (2013), Donaldson and Hornbeck (2013), Duranton and Turner (2011, 2012), Duranton et al. (2013), Faber (2013), Martincus et al. (2012), Michaels (2008), and Sotelo (2014). The main concern of this literature has been finding exogenous sources of variation in transport infrastructure to estimate its causal impact on relative outcomes in treated versus untreated locations. In contrast, we develop a quantitative general equilibrium model of the internal distribution of economic activity across regions and sectors. Distinctive features of our quantitative analysis are the emphasis on the role of transport infrastructure in enabling interior regions to participate in world markets and in driving structural transformation.[3]

Our analysis also connects with the macroeconomics and development literatures on structural transformation from agriculture into non-agriculture, including Bustos et al. (2012), Caselli and Coleman (2001), Foster and Rosenzweig (2007), Gollin et al. (2012, 2013), Karádi and Koren (2013), Lagakos and Waugh (2013), Matsuyama (1992), Michaels et al. (2012), Ngai and Pissarides (2007), Herrendorf et al. (2013), Swiecki (2013), and Uy, Yi, and Zhang (2012).[4] In the macroeconomics literature the Balassa-Samuelson effect is driven by differences in productivity between the traded and non-traded sector for the aggregate economy as a whole.[5] In contrast, in our model, these effects emerge endogenously from geographical location alone: more remote locations have higher relative prices of tradables, higher wage-rental ratios, higher agricultural employment shares and lower population densities. This role for internal geography in turn influences the aggregate magnitude of structural transformation from agriculture towards non-agriculture.

The remainder of the paper is structured as follows. Section 2 discusses the historical background. Section 3 develops the theoretical model that we use to guide our empirical analysis. Section 4 discusses the data sources and definitions. Section 5 undertakes a quantitative analysis of the model and reports counterfactuals. Section 6 provides further evidence on the economic mechanisms in the model. Section 7 concludes. Technical derivations and supplementary material are collected together in the web appendix.

---

[3] Reduced-form studies of the impact of trade liberalization on local labor markets include Kovak (2013), Topalova (2010), and McCaig and Pavcnik (2012).

[4] Much of this literature excludes geographic variation within countries. A small number of studies do consider transport costs as a potential determinant of agricultural employment shares, including Adamopoulos (2012), Gollin and Rogerson (2012) and Herrendorf et al. (2012). But these studies consider stylized settings of for example two regions that do not trade internationally and are less well suited to taking to our spatially disaggregated data.

[5] See Balassa (1964) and Samuelson (1964) for the conventional explanation of this relationship.

2 Historical Background and Aggregate Trends

2.1 Pre-Export Boom Era

The earliest Spanish explorations of present-day Argentina date back to the first half of the Sixteenth Century.[6] Initially, economic activity was orientated towards the silver mines at Potosí in Bolivia rather than the Atlantic coast.[7] Official trade routes with Spain ran towards the Northwest, through Potosí and Lima, to Panama. In contrast, international trade was officially forbidden from Buenos Aires, so that the River Plate region (*Río de la Plata*) lied on the periphery of the Spanish Empire as an outpost for illegal trade with Brazil, Portugal and Britain.

Eventually, the growth of this illegal trade and threats from Portuguese settlement along the Río de la Plata spurred the opening of Buenos Aires to official trade and the establishment in 1776 of the Viceroyalty of the Río de la Plata. Throughout the colonial era, Spanish merchants retained a monopoly of all official trade. However, population growth and economic development led to growing demands for political autonomy from Spain. When the Napoleonic Wars undermined Spanish imperial power, these growing pressures brought about a transfer of political power to a local junta in 1810 and the opening of official direct trade with the merchants of other countries (in particular Britain and Portugal).

Despite initial attempts to restore Spanish imperial power, Argentinian independence was ultimately achieved in 1816. The decades immediately following independence were taken up with internal power struggles between Buenos Aires and the interior regions of Argentina. However, there was a move towards political stability from 1850 onwards. The first national constitution was agreed in 1853, the first constitutional government of all provinces met in 1862, and Buenos Aires was absorbed into the federal structure of Argentina in 1880. Further consolidation came with a series of campaigns against native populations in the hinterland of Buenos Aires that culminated with the "Conquest of the Desert" in 1879-80.[8] The election of Julio Roca to the Presidency in 1880 ushered in a sequence of liberal regimes open to foreign trade, capital and migration.

2.2 External and Internal Trade Costs Reductions in the Late-19th Century

From the mid-nineteenth century onwards, a series of technological innovations led to substantial reductions in maritime transport costs. According to the freight indices of North (1958) and Harley (1988), 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.[9]

---

[6] This section draws in particular on the discussions in Adelman (1994) and Scobie (1971).

[7] Early settlement patterns were heavily influenced by the availability of passive native Indian populations that were used as a source of forced labor 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).

[8] Until 1880, the development of large areas of the land subsequently used for agricultural production was limited by incursions from hostile native populations (see for example Droller 2013).

[9] These declines in freight rates were associated with a convergence in commodity prices: the gap between wheat prices in Liverpool and Chicago fell from 57.6 percent in 1870 to 17.8 percent in 1895 and 15.6 percent in 1913 (Harley

Alongside this general decline in international transport costs, technological innovations dramatically reduced transport costs for particular goods. The technology for freezing meat was invented in Australia in 1861, with an initial trial shipment from Australia to Britain occurring in 1876, and an initial trial shipment from Buenos Aires to France occurring in 1877. Shortly thereafter, the first freezing plant (*frigorifico*) in Buenos Aires was built by British investors in 1882. After 1908 further improvements in refrigeration technology permitted the shipment of chilled as well as frozen meat. This new technology made possible for the first time long-distance trade in chilled and frozen meat, opening up the markets of Europe and the United States to Argentinian producers.

The late nineteenth century also saw substantial reductions in internal transport costs as a result of railroad construction. When water transport was unavailable, the previous state of the art mode of land transport was by oxcart across dirt tracks. Using this mode of transport, it took around 3-4 months to travel from Buenos Aires to Salta, the largest city in the country's Northwest, and it cost around thirteen times as much to move a ton of goods from Salta to Buenos Aires as to move it from Liverpool to Buenos Aires (Scobie 1971, p. 94). The Buenos Aires Western Railway was the first to be constructed in 1857 and by 1869 around 700 kilometers of track had been completed. From this point onwards, the railroad network expanded rapidly to grow to around 13,000 kilometers in 1895 and 30,000 kilometers in 1914.[10]

2.3 Export Boom and Economic Growth

Prior to the nineteenth century, the economy of Buenos Aires was initially based on cattle that were the descendants of the escaped animals of early Spanish settlers. Extensive cattle ranching was undertaken on large estates (*estancias*). The main export goods were cattle hides and tallow to Europe and salted meat to the slave plantations of Brazil and Cuba. Processing of these export goods occurred in salting-plants (*saladeros*) that were concentrated in Buenos Aires, which rapidly developed into the main port and export processing center serving the surrounding Pampas agricultural region. As the nineteenth century progressed, extensive sheep ranching became increasingly important, and wool began to account for an increasing share of the value of exports.

The late-nineteenth century reductions in transport costs discussed above precipitated an export boom. As shown in Panel A of Figure 1, real exports more than quadrupled between 1869 and 1910.[11] This export boom was almost entirely driven by natural resource-based specialization in agriculture. As shown in Panel A of Table 1, Livestock and Agriculture accounted for more than 95 percent of the value of Argentina's exports throughout this period. Almost no exports of manufacturing goods are observed, although some production of manufacturing for the local market occurred.[12] In contrast, imports were much more diversified across sectors, with agricultural exports

---

1980). For the classic analysis of this integration of Atlantic commodity markets, see O'Rourke and Williamson (1999).

[10] This 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.

[11] We end our sample in 1914 to avoid the effects of the First World War and the political instability and more interventionist government policies from the 1920s and 1930s onwards. For an analysis of the role of changes in external integration in influencing Argentine economic development after 1914, see Taylor (1992).

[12] As discussed in Rocchi (2006), the limited amount of domestic manufacturing activity involved either the pro-

largely being exchanged for manufacturing imports, as shown in Panel A of Table 2 for 1895.[13]

Consistent with the trade boom being driven by Argentina's comparative advantage in agricultural products relative to the rest of world, we find that external exports from Argentine customs to foreign countries are large relative to internal shipments between Argentine customs (Panel B of Figure 1). Exports are also highly concentrated across customs, with Buenos Aires accounting for over 88 percent of the value of exports in 1869, and together with its neighbor La Plata continuing to account for 50 percent of the value of exports in 1914. Across the period 1869-1914, the top four customs of Buenos Aires, Rosario, La Plata and Bahia Bahia account for 75 percent of the value of exports (Panel C of Figure 1).[14]

Underlying this export boom was an "agricultural revolution on the pampas." Total cultivated area more than trebled from around 40,000 to 143,000 square kilometers between 1895 and 1914. Reductions in both external and internal trade costs precipitated large-scale changes in the composition of agricultural exports and the allocation of cultivated area across agricultural goods. As shown in Panel B of Table 1, hides, bones and other animal parts accounted for the majority of the value of Argentinian exports in 1869, with wool the other largest item included in the other product category. In contrast, Cereals rose from a negligible export share in 1869 to more than 50 percent of exports by 1914. In the first half of our sample period, beef's export share declined, because of the emergence of new export goods and the replacement of salted beef exports with live cattle exports. In the second half of our sample period, this initial decline is reversed, as exports of frozen and chilled beef expanded. As shown in Panel B of Table 2, imports of manufacturing goods included substantial imports of agricultural machinery (e.g. mowers, plows, rakes, threshers and metal wire), as reflected in the increased adoption of agricultural machinery over time.[15]

This export boom also involved large-scale immigration and rapid economic development. As shown in Panel D of Figure 1, Argentina's population increased from around 1.8 to 7.9 million between 1869 and 1914, with around half this increase achieved through net immigration. Despite this increase in labor supply, the real wage and income per capita grew at annual average rates of 1.1% and 2.5% respectively over the same period, so that Argentina became the eighth richest country in the world by 1914.[16] This rapid economic development involved both urbanization and structural transformation: The share of the population living in cities and towns grew by around 20 percentage points from 1869-1914 (see Figure 2), while the share of the population employed in

---

cessing of agricultural goods for export or was orientated towards the domestic market, including consumer goods industries such as Food, Beverages and Tobacco.

[13] Aggregate import categories are less stable over time than aggregate export categories and hence we focus here on import composition for 1895. We observe a similar pattern for other years in our sample.

[14] The establishment of the Viceroyalty of the Río de la Plata in Buenos Aires stimulated its development as the main port, even though the shallow shores of the estuary were not well suited for a port: "Ironically, the sixteenth century Spaniards, searching for an anchorage for their tiny ships, elected one of the poorest sites imaginable in terms of nineteenth-century sailing vessels and steamships." (Scobie 1971, p. 95) As late as the 1880s, ships had to anchor several miles from shore in the open roads, and construction of the Madero docks was not completed until 1897.

[15] As discussed in Adelman (1994), almost all of Argentina's agricultural implements were imported.

[16] See 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.

agriculture fell by around 7 percentage points from 1895-1914.[17]

2.4 Spatial Pattern of Economic Development

Rapid economic development also involved major changes in the spatial distribution of economic activity within Argentina. Figures 3-5 display the population density distribution in 1869, 1895 and 1914 respectively. In each figure, we use constant district boundaries (based on 1895 districts) and we divide the population density distribution into the same five discrete cells, with darker shading indicating higher population densities. We show the railroad network in green, the main navigable rivers (the Paraná, Plate and Uruguay) in blue, and the location of customs in red.

As shown in Figure 3, most of the country was sparsely populated in 1869. 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. Buenos Aires was by far the largest city and most important custom among the network of customs orientated around the Paraná-Plate-Uruguay river system. The railroad network of consisted of only 700 kilometers of track, including a couple of lines radiating from Buenos Aires and a line connecting the port of Rosario with the interior city of Córdoba.

Comparing Figures 3 and 4, population density from 1869-95 radiated inland from Buenos Aires towards its surrounding agricultural hinterland. The railroad network also expanded substantially to connect the Spanish colonial towns and to integrate the interior agricultural regions with Buenos Aires and the other customs on the coast or Paraná-Plate-Uruguay river system. The increase in population, economic activity and foreign trade is reflected in the growth in both the number and geographical spread of customs.

From Figures 4-5, the period 1895-1914 saw a continuation of this radiation of population density inland from Buenos Aires. The railroad network extended further into the interior and the density of railroad lines in the agricultural hinterland around Buenos Aires increased. This period also saw a further expansion in the number and geographical spread of customs, which began to encroach into the previously remote and undeveloped areas towards the South.

These changes in the spatial distribution of population density are accompanied by changes in the pattern of urbanization and the composition of employment across sectors. As shown in Figure A1 in the web appendix, high urban population shares in 1869 were concentrated around Buenos Aires and the Spanish colonial towns that served the mining region of Upper Peru. As shown in Figures A2 and A3 in the web appendix, these high urban population shares radiate outwards from Buenos Aires towards its agricultural hinterland over time. As shown in Figure A4 in the web appendix, high urban population shares are reflected in high shares of employment in non-agricultural activities.

---

[17] In Figure 2, we use the population census definition of rural and urban areas (based on the share of the population living in cities and towns). We also find a rise in urbanization of around the same magnitude using a more conservative definition based on the share on the population in cities with more than 2,000 inhabitants.

3 Theoretical Model

In this section, we develop our quantitative model of the distribution of economic activity across regions and sectors within countries.[18] The features of the model map approximately one-to-one to the objects observed in our data. We first characterize the model's properties analytically, before showing in a later section how it can be taken directly to the data.

The distribution of economic activity across regions and sectors is determined by productivity and relative prices, where these relative prices depend on both external and internal trade costs. The economy consists of three sectors: manufacturing (M), agriculture (A), and non-tradables (N). Since our data contain information on individual goods within the agricultural sector, we model agriculture as a composite sector that includes a discrete number of agricultural goods indexed by $g = 1, \cdots, G$. We assume that the economy is small relative to world markets and hence faces exogenous prices for traded goods that can change over time.[19]

The economy as a whole consists of a set of locations $\ell \in \mathcal{L}$ that differ in terms of their geographical position and natural endowments. Some of these locations $\ell \in \mathcal{L}_C \subset \mathcal{L}$ are coastal and have direct access to world markets at exogenous prices $\left\{P_g^*\right\}_{g=1}^G, P_M^*$ that depend on external transport costs (e.g. transatlantic freight rates). Other locations $\ell \in \mathcal{L}_I \subset \mathcal{L}$ are interior regions that are connected to the ports through an internal transport network (e.g. the railroad network). We denote the trade cost between any pair of locations $(\ell, \ell') \in \mathcal{L}$ for good g by $\delta_g(\ell, \ell')$. This trade cost is allowed to change over time with improvements in the internal transport network. Each location ℓ has a land area $L(\ell)$ and consists of a continuum of land plots $j \in [0, L(\ell)]$ that are heterogeneous in terms of their productivity for each of the agricultural goods $g = 1, \cdots, G$.[20]

3.1 Preferences and Endowments

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

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

where $c_T(\ell)$ and $c_N(\ell)$ respectively denote consumption of the tradable and non-tradable goods. Following the literature on structural transformation in macroeconomics, we assume inelastic de-

---

[18] A web-based technical appendix contains the derivation of the expressions and results in this section.

[19] For most goods during our sample period, the assumption that Argentina is small relative to world markets is a reasonable approximation. For example, Bennett (1933) estimates that world production of wheat in 1895 (1914) was 2,730.9 (3,617.6) bushels of 60 pounds, whereas Argentina's production of wheat was 46.4 (169.2). Relaxing the assumption of a small open economy leaves the general equilibrium relationships of the model entirely unchanged conditional on prices at the port, but implies that these prices at the port become endogenous. Therefore, our calibration of the model remains unchanged irrespective of whether the economy is small or large, because we recover price-adjusted productivities without taking a stand on their determinants. Only when we use the model to undertake counterfactuals are we required to take a stand on whether or not prices at the port are exogenous.

[20] In the model, all land is used productively. Therefore in our empirical analysis we use geographical land area as our measure of land area in the model. In Section A.3.2 of the web appendix we develop an extension of the model, in which landowners make an endogenous decision whether to leave land wild or convert it to productive use.

mand between tradables and non-tradables ($0 < \sigma < 1$).[21] Tradables consumption is in turn defined over consumption of a composite manufacturing good and the set of agricultural goods $g = 1, \ldots, G$ with the following homothetic price index:

$$E_T(\ell) = E_T\left(\{P_g(\ell)\}_{g=1}^{G}, P_M(\ell)\right), \tag{2}$$

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

Each worker is endowed with one unit of labor that is supplied inelastically with zero disutility. Workers are perfectly mobile across locations and hence arbitrage away real wage differences.[22] The labor market clearing condition for the economy as a whole can be written as:

$$\sum_{\ell \in \mathcal{L}} L(\ell) n(\ell) = N, \tag{3}$$

where $n(\ell) = N(\ell)/L(\ell)$ is population density at location ℓ, to be determined in equilibrium; N is the economy's aggregate labor endowment, which can change over time with both native population growth and foreign migration. Land is owned by immobile landowners who consume where they live and do not own any labor.[23]

Demands for traded and non-traded goods in location ℓ per unit of land are, respectively,

$$c_T(\ell) = \beta_T \left(\frac{E_T(\ell)}{E(\ell)}\right)^{-\sigma} \frac{y(\ell)}{E(\ell)}, \tag{4}$$

$$c_N(\ell) = (1 - \beta_T) \left(\frac{P_N(\ell)}{E(\ell)}\right)^{-\sigma} \frac{y(\ell)}{E(\ell)}, \tag{5}$$

where $y(\ell)$ is income per unit of land (including payments to both labor and land).

3.2 Production Technology

Production in each sector occurs under conditions of perfect competition and constant returns to scale. For simplicity, we assume that the production technology takes the Cobb-Douglas form so that output per unit of land is:

$$\begin{aligned} q_N(\ell) &= z_N(\ell) n_N(\ell)^{1-\alpha_N}, \\ q_M(\ell) &= z_M(\ell) n_M(\ell)^{1-\alpha_M}, \\ q_{g,j}(\ell) &= z_{g,j}(\ell) n_{g,j}(\ell)^{1-\alpha_A}, \end{aligned} \tag{6}$$

---

[21] See, for example, Ngai and Pissarides (2007) and Herrendorf et al. (2013).

[22] Since our sample period is characterized by large-scale population movements to and within Argentina, the assumption of perfect labor mobility appears to be a reasonable approximation. Imperfect labor mobility can be introduced into the model following the approach in Redding (2012).

[23] Under our assumptions of identical and homothetic preferences, equilibrium allocations are invariant with respect to the number of landowners in each location.

where $0 < \alpha_i < 1$ is the land intensity in sector $i = A, M, N$; g indexes agricultural goods; and j indexes land plots within each location ℓ. We make the natural assumption that agriculture is land intensive relative to manufacturing and non-tradables ($\alpha_A > \alpha_M$ and $\alpha_A > \alpha_N$), but that non-tradables and manufactures still require some land ($\alpha_M > 0$ and $\alpha_N > 0$), which is consistent with the high land prices observed in urban areas.

Although, in developing the model, we treat sectoral productivities as exogenous primitives, it is straightforward to allow them to depend on agglomeration forces through external economies of scale (for example, $z_i = Z_i n_i^{\eta_i}$ for sector i, where $\eta_i \geq 0$ parameterizes the strength of agglomeration and Z_i captures the exogenous component of productivity). An advantage of our quantitative approach is that we use utility maximization, profit maximization and population mobility to recover the values that sectoral productivities (z_i) must take in order for the data to be consistent with an equilibrium the model, without taking a stand on the determinants of sectoral productivities. Only when we use the model to undertake counterfactuals are we required to make assumptions about whether sectoral productivities are themselves exogenous ($z_i = Z_i$) or endogenous ($z_i = Z_i n_i^{\eta_i}$ for $\eta > 0$).

In the non-tradable and manufacturing sectors, productivity is allowed to vary across locations $\ell \in \mathcal{L}$ but is assumed to be the same for all land plots $j \in [0, L(\ell)]$ within a location ℓ. In contrast, in agriculture, land plots within each location can differ in terms of their productivities across agricultural goods (e.g. as a result of soil and weather). The realizations of productivity for each agricultural good and land plot $\{z_{g,j}(\ell)\}_{g=1}^{G}$ are drawn independently from a Fréchet distribution. In Eaton and Kortum (2002), the properties of this functional form for productivity are applied across a continuum of goods. In contrast, we use these properties across a continuum of land plots for each good, which enables us to consider a discrete number of agricultural goods and obtain determinate predictions for the shares of agricultural land allocated to each good:

$$Pr\left[z_{g,j}(\ell) < z\right] = e^{-T_g(\ell) z^{\theta(\ell)}}, \tag{7}$$

where $T_g(\ell)$ controls the average productivity of good g in location ℓ and $\theta(\ell)$ controls the dispersion of agricultural productivity in location ℓ.

Our specification with a continuum of land plots and a discrete number of agricultural goods allows for zero land shares for some agricultural goods in some locations, because the Fréchet scale parameter $T_g(\ell)$ can vary by both good g and location ℓ. Therefore we interpret a zero land share for agricultural good g in location ℓ as corresponding to the limiting case in which $\lim_{T_g(\ell) \to 0}$. Similarly, our framework can accommodate the zero populations observed for some locations in the data, which are rationalized in the model by zero productivities in tradeables: $z_A(\ell) = z_M(\ell) = 0$.[24]

---

[24] In the extension of the model in Section A.3.2 of the web appendix, in which landowners choose whether to convert land from wild to productive use, a location also may have zero population because it is not profitable to convert land to productive use.

3.3 Profit Maximization

Markets are perfectly competitive. In each sector, 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. Firms make zero profits in each location in each sector with positive production. If a plot of land is allocated to manufacturing or non-tradables, land rents are equal to revenue per unit of land minus labor costs per unit of land at the equilibrium value of employment density:

$$r_i(\ell) = \max_{n_i(\ell)} \{P_i q_i(n_i(\ell)) - w(\ell) n_i(\ell)\} \qquad \text{for } i = M, N. \tag{8}$$

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:

$$\begin{aligned} r_j(\ell) &= \max_{g=1,..,G} \{r_{g,j}(\ell)\}, \\ r_{g,j}(\ell) &= \max_{n_{g,j}(\ell)} \{P_g(\ell) q_{g,j}(n_{g,j}(\ell)) - w(\ell) n_{g,j}(\ell)\}. \end{aligned} \tag{9}$$

In recognition of the importance of shocks to agricultural productivity (e.g. due to weather), we assume that the landowner of each plot $j \in [0, L(\ell)]$ must choose whether to allocate that plot to agriculture, manufacturing or non-tradables before observing the realization of the productivity shocks $\{z_{g,j}(\ell)\}_{g=1}^{G}$. After a plot of land has been allocated to agriculture, the landowner observes these realizations of productivity, and decides which of the individual agricultural goods to produce. Therefore, the decision whether to allocate the plot of land to manufacturing, non-tradables or agriculture depends on expected land rents in agriculture,

$$r_A(\ell) = \mathbb{E}[r_j(\ell)],$$

which in turn depend on the probability distribution of agricultural productivities defined in (7).

3.4 Sectoral Employment and Wage-Rental Ratio

Using profit maximization and zero profits, equilibrium sectoral variables in each sector and location can be written in terms of the wage-rental ratio $\omega_i(\ell) = w(\ell)/r_i(\ell)$, which itself can be written in terms of wages, productivity and prices. For the manufacturing and non-tradables sectors $i \in \{M, N\}$, employment per unit of land and the wage-rental ratio are

$$n_i(\ell) = \frac{1-\alpha_i}{\alpha_i} \frac{1}{\omega_i(\ell)}, \tag{10}$$

$$\omega_i(\ell) = \left(\frac{w(\ell)}{P_i(\ell) z_i(\ell)} \right)^{\frac{1}{\alpha_i}}. \tag{11}$$

In the agricultural sector, once a plot of land j in location ℓ has been assigned to the production of the agricultural good g, equilibrium variables take exactly the same form as for the composite sectors in (10) to (11), but with wage-rental ratio $\omega_{g,j}(\ell)$, price $P_g(\ell)$, and productivity $z_{g,j}(\ell)$.

The properties of the Fréchet distribution and the continuum of land plots imply that the heterogeneous goods within the agricultural sector aggregate to a representative agricultural good $i = A$ with the following revenue productivity:

$$z_A(\ell) = \Gamma\left(\frac{\alpha_A \theta(\ell) - 1}{\alpha_A \theta(\ell)}\right)^{\alpha_A} \left[\sum_{g=1}^{G} T_g(\ell) P_g(\ell)^{\theta(\ell)}\right]^{1/\theta(\ell)}, \tag{12}$$

where $\Gamma(\cdot)$ is the Gamma function. Therefore, as for manufacturing and non-tradables, 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 price equal to $P_A(\ell) = 1$.[25] Employment density and the wage-rental ratio in the aggregate agricultural sector take the same form as for the manufacturing and non-tradable sectors in (10) and (11), 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 $P_A(\ell) = 1$.

The aggregate agricultural revenue productivity $z_A(\ell)$ completely summarizes the impact of the production technology for each agricultural good $T_g(\ell)$ and the local price of each agricultural good $P_g(\ell)$ on the agricultural sector in each location; these local prices, in turn, depend on world prices and the internal transport network.

3.5 Definition of Equilibrium

We are ready to define the general equilibrium of the economy.

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,

$$u(\ell) \leq u^* \qquad \text{and} \qquad u(\ell) = u^* \text{ if } n(\ell) > 0;$$

(ii) land is allocated optimally across sectors,

$$r(\ell) = \max\{r_A(\ell), r_M(\ell), r_N(\ell)\};$$

---

[25] Note that this is not a price normalization. The entire distribution of agricultural prices is contained in the distribution of $z_A(\ell)$.

(iii) the land market clears in each location,

$$\sum_{i=M,N,A} L_i(\ell) = L(\ell);$$

(iv) population density and sectoral employment adjust to clear the labor market in each location,

$$\sum_{i=M,N,A} \frac{L_i(\ell)}{L(\ell)} n_i(\ell) = n(\ell);$$

(v) the non-tradable 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, i.e.

- 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 (3) holds.

Under our neoclassical assumptions, there exists an unique allocation and set of prices that satisfies these equilibrium conditions.

Proposition 1. *There exists a unique general equilibrium.*

Proof. See the web-based technical appendix $\square$

Two features of the equilibrium definition are worthy of discussion. First, conditions (i) to (v) define a "local" equilibrium for each location ℓ taking the prices of tradable goods and the real wage u^* as given. The distinction between this "local" equilibrium and the general equilibrium proves useful in characterizing the properties of the model. Second, there is potentially rich heterogeneity in agricultural outcomes within each location ℓ as captured in the definition of $z_A(\ell)$. For example, two locations may have the same aggregate employment and land use in agriculture, but have different allocations of employment and land across individual goods within the agricultural sector.

The model determines the structure of each local economy ℓ (the pattern of specialization across sectors and goods with their respective employment and labor shares), the level of economic activity (population density and income per worker), and the distribution of income between labor and land. Next, we use the model to characterize how these outcomes vary with internal and external trade costs. We then move to an empirical and quantitative evaluation of these predictions using the historical Argentinean data.

3.6 Sectoral Specialization Pattern

Condition (i) of the general equilibrium implies that real wages are equalized across all populated locations,[26]

$$u^* = \frac{w(\ell)}{E(\ell)} = \frac{w(\ell)}{\left[\beta_T E_T(\ell)^{1-\sigma} + (1-\beta_T) P_N(\ell)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}}. \tag{13}$$

Since preferences satisfy the Inada conditions, each location necessarily produces non-tradables N and goods in at least one tradable sector, A or M. Assuming that there is positive production in tradable sector i, condition (ii) of the general equilibrium implies that the equilibrium wage-rental ratio must be the same across sectors, $\omega_N(\ell) = \omega_i(\ell)$. Combining this with expression (11) and using the labor mobility condition (13), we implicitly obtain the wage-rental ratio when there is positive production in tradable sector i, $\omega_i(\ell)$, as the unique solution to[27]

$$\left[\beta_T \left(\frac{P_i(\ell)}{E_T(\ell)} z_i(\ell)\,\omega_i(\ell)^{\alpha_i}\right)^{\sigma-1} + (1-\beta_T)\left(z_N(\ell)\omega_i(\ell)^{\alpha_N}\right)^{\sigma-1}\right]^{\frac{1}{\sigma-1}} = u^*. \tag{14}$$

Under autarky, goods in the three sectors $\{N, M, A\}$ are produced. Therefore, (14) must hold for $i = A, M$. Using (11) and imposing $\omega_A(\ell) = \omega_M(\ell)$ we obtain the autarkic wage-rental ratio in location ℓ,

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

Using this expression we can characterize the sectoral pattern of specialization.

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 $z_A(\ell)$.*

Proof. See the web-based technical appendix. □

If a region trades with the rest of the world, constant returns to scale and population mobility imply that goods in only *one* traded sector are produced, while nontraded goods are necessarily produced. Consistent with our data and the historical evidence discussed above, we assume that all Argentinian regions have a comparative advantage relative to the rest of the world in agricultural goods (i.e., $z_A(\ell)$ is sufficiently large in every location). Therefore, when a region trades, it specializes in agriculture and non-traded goods $\{N, A\}$. In turn, whether or not the region trades is determined by comparative advantage and transport costs. Trade takes place if the relative price

---

[26] In an extension of the model with imperfect labor mobility as in Redding (2012), real wages would no longer be equalized, but an analagous population mobility would regulate the size of real wage differences across locations.

[27] To reach 14, first rewrite (13) 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 (11).

of the imported manufacturing good net of transport costs is less than the relative price of the manufacturing good under autarky, which is the case for high enough values of $z_A(\ell)/z_M(\ell)$.[28]

3.7 Specialization across Goods Within the Agricultural Sector

Under our empirically-motivated assumption that all Argentine locations have a comparative advantage relative to the rest of the world in agriculture, and using (14), the wage-rental ratio in location ℓ satisfies:

$$\left[\beta_T \left(\widetilde{z}_A\left(\ell\right)\omega\left(\ell\right)^{\alpha_A}\right)^{\sigma-1} + \left(1-\beta_T\right)\left(z_N(\ell)\omega\left(\ell\right)^{\alpha_N}\right)^{\sigma-1}\right]^{\frac{1}{\sigma-1}} = u^*, \tag{15}$$

where we define

$$\widetilde{z}_A\left(\ell\right) = \frac{z_A\left(\ell\right)}{E_T\left(\ell\right)} \tag{16}$$

as a measure of productivity in the agricultural sector adjusted for the tradables price index.

The population mobility condition (15) determines the endogenous equilibrium wage-rental ratio in each trading location $\omega\left(\ell\right)$ given exogenous aggregate productivites $\left\{z_A\left(\ell\right),\, z_N\left(\ell\right)\right\}$ and the tradables price index $E_T\left(\ell\right)$ that is determined by no-arbitrage. Changes in external and internal transport costs influence the general equilibrium of the model through $z_A\left(\ell\right)$ and $E_T\left(\ell\right)$. It follows that adjusted aggregate agricultural productivity $\widetilde{z}_A\left(\ell\right)$ is a sufficient statistic for the impact of external and internal trade costs on the equilibrium in each location through the tradables sector.[29]

Within the agricultural sector, a share $l_g\left(\ell\right)$ of the land allocated to agriculture is allocated to good g. This share equals the probability that the solution to the landowner's discrete choice problem in (9) yields good g as an outcome. It depends on relative productivites $\left\{T_g\left(\ell\right)\right\}$, relative local prices $\left\{P_g\left(\ell\right)\right\}$, and the Fréchet shape parameter $\theta\left(\ell\right)$:

$$l_g\left(\ell\right) = \frac{T_g\left(\ell\right)P_g(\ell)^{\theta(\ell)}}{\sum_{g'} T_{g'}\left(\ell\right)P_{g'}(\ell)^{\theta(\ell)}}. \tag{17}$$

Assuming that the tradables price index $E_T\left(\ell\right)$ takes the Cobb-Douglas form with share γ_g on each agricultural good, we can relate these land shares to patterns of trade. Given the following tradables price index:

$$E_T\left(\ell\right) = P_M\left(\ell\right)^{1-\gamma_A}\prod_{g=1}^{G} P_g\left(\ell\right)^{\gamma_g}, \qquad \text{where} \qquad \sum_{g=1}^{G}\gamma_g = \gamma_A, \tag{18}$$

---

[28] For sufficiently large transport costs, the model features a "trade frontier" beyond which regions further inland are in autarky. As transport costs fall, there is an expansion of this frontier further inland as additional regions are integrated into world markets.

[29] Aggregate agricultural productivity $z_A\left(\ell\right)$ defined in (12) depends on the following exogenous parameters: productivities for each agricultural good $\left\{T_g\left(\ell\right)\right\}$, prices at the port $\left\{P_g^*\left(\ell'\right) \text{ for } \ell' \in \mathcal{L}_C\right\}$ and transport cost to the nearest port $\left\{\delta_g\left(\ell\right)\right\}_g$. The tradables consumption index $E_T\left(\ell\right)$ depends on the exogenous parameters of prices at the port $\left\{P_g^*\left(\ell'\right), P_M^*\left(\ell'\right) \text{ for } \ell' \in \mathcal{L}_C\right\}$ and internal transport costs $\left\{\delta_g\left(\ell\right)\right\}_g$. External transport costs determine prices at the port, while internal transport costs dictate the wedge between prices at the port and local prices.

the share of each good's exports $x_g(\ell)$ in the total exports $x_A(\ell)$ of each region can be expressed as follows:

$$\frac{x_g(\ell)}{x_A(\ell)} = \frac{l_g(\ell) - \gamma_g}{1 - \gamma_A}. \tag{19}$$

Therefore a location exports an agricultural good g if its land share $(l_g(\ell))$ exceeds its expenditure share (γ_g). This yields a simple chain of comparative advantage within agriculture: if location ℓ exports good g, it necessarily exports all goods g' such that $l_{g'}(\ell)/l_g(\ell) > \gamma_{g'}/\gamma_g$.[30] Furthermore, from (17), relative land shares depend solely on relative prices and technologies:

$$\frac{l_{g'}(\ell)}{l_g(\ell)} = \frac{T_{g'}(\ell)\, P_{g'}(\ell)^{\theta(\ell)}}{T_g(\ell)\, P_g(\ell)^{\theta(\ell)}}. \tag{20}$$

Hence the chain of agricultural comparative advantage is determined by a composite of relative technologies and a power function of relative prices that can be recovered from observed data on relative land shares. This in turn implies a tight connection between relative exports and relative land shares for agricultural goods (from (19) and (20)) that we examine empirically below.

3.8 Spatial Balassa-Samuelson Effect

The model features a *spatial Balassa-Samuelson effect* that determines the impact of remoteness from world markets on the structure of economic activity within each location. Locations closer to world markets (lower $\delta_g(\ell, \ell')$ for all goods g and other locations ℓ') have higher adjusted productivity in agriculture (higher $\widetilde{z}_A(\ell)$), which in turn implies a *higher relative price of non-traded goods* (lower $E_T(\ell)/E(\ell)$) and a *lower wage-rental ratio* (lower $\omega(\ell)$). This spatial variation in relative prices in turn affects population density, income, and structural transformation across sectors: Locations closer to world markets have higher shares of employment in non-tradables, higher population density, and higher aggregate income.

The intuition for this *spatial Balassa-Samuelson effect* is as follows. A lower value of trade costs increases the prices of exported goods and decreases the price of imported goods. On the one hand, the reduction in import prices reduces the tradables consumption price index $(E_T(\ell))$ and the relative price of tradables $(E_T(\ell)/E(\ell))$. On the other hand, the increase in export prices has the direct effect of increasing the tradables consumption price index $(E_T(\ell))$ and the relative price of tradables $(E_T(\ell)/E(\ell))$. But this increase in export prices also indirectly raises wages $(w(\ell))$ and land prices $(r(\ell))$, which in turn raises the prices of non-tradables $(P_N(\ell))$.

The net effect of lower trade costs is to reduce the relative price of tradables and increase real incomes, which has to be compensated by an increase in population to bid up the prices of non-

---

[30] Inspection of (17) reveals that, if transport costs take the form of common iceberg costs across goods, agricultural land and export shares are independent of the level of transport costs. More generally, if transport costs differ across goods and increase with distance to ports, the model implies that more remote regions export a narrower range of products than more centrally-located regions, because transport costs are a source of comparative advantage. In Section A.3.1 of the web appendix, we derive this prediction formally, and in Section A.4 of the web appendix we show that customs in more remote provinces do indeed export a narrower range of products in the data.

tradables and land until real wages are equalized across all populated locations. With inelastic demand between sectors, the lower relative price of tradables implies a higher expenditure share of non-tradables, which in turn implies a higher employment share in non-tradables. Together the higher employment share in labor-intensive non-tradables and the higher population density imply that both sectors must use more labor-intensive techniques in order to satisfy land market clearing, which implies a lower wage-rental ratio (lower $\omega(\ell)$).

This *spatial Balassa-Samuelson effect* also can be seen more formally. From population mobility (13), zero profits, and market clearing, equilibrium population density $n(\ell)$ is:

$$n(\ell) = \frac{N(\ell)}{L(\ell)} = \left(\frac{1}{\alpha_N + (\alpha_A - \alpha_N)\,\beta_T \left(\frac{E_T(\ell)}{E(\ell)}\right)^{1-\sigma}} - 1 \right) \frac{1}{\omega(\ell)}, \tag{21}$$

and the share of labor used in agriculture is:

$$\nu_A(\ell) = \frac{N_A(\ell)}{N(\ell)} = \frac{(1-\alpha_A)\,\beta_T \left(\frac{E_T(\ell)}{E(\ell)}\right)^{1-\sigma}}{1 - \left[\alpha_N + (\alpha_A - \alpha_N)\,\beta_T \left(\frac{E_T(\ell)}{E(\ell)}\right)^{1-\sigma}\right]}. \tag{22}$$

With inelastic demand between tradables and non-tradables ($0 < \sigma < 1$) and labor intensive non-tradables ($\alpha_A > \alpha_N$), the above relationships imply that $n(\ell)$ is decreasing with $\nu_A(\ell)$ and $\omega(\ell)$, while $\nu_A(\ell)$ is decreasing in $E_T(\ell)/E(\ell)$. Furthermore, both $\omega(\ell)$ and $E_T(\ell)/E(\ell)$ are decreasing in $\widetilde{z}_A(\ell)$, where $\widetilde{z}_A(\ell) = z_A(\ell)/E_T(\ell)$ is itself decreasing in $\delta_g(\ell)$ for all goods g.

Proposition 3. ***(Spatial-Balassa Samuelson Effect)*** *If traded and non-traded goods are complements ($\sigma < 1$) and agriculture is more land-intensive than non-traded activities, ($\alpha_A > \alpha_N$), then high trade-cost locations (locations ℓ with higher transport costs $\delta(\ell, \ell')$ and hence lower $\widetilde{z}_A(\ell)$) have* ***(a)*** *higher wage-rental ratios (higher $\omega(\ell)$),* ***(b)*** *lower relative prices of non-traded goods (higher $E_T(\ell)/E(\ell)$),* ***(c)*** *lower population densities (lower $n(\ell)$), and* ***(d)*** *larger shares of labor in agriculture (larger $\nu_A(\ell)$).*

Proof. See the web-based technical appendix. □

This spatial Balassa-Samuelson effect shapes population density, the share of employment in agriculture (structural transformation), and hence the share of population in rural versus urban areas (because agriculture is produced in rural areas and non-traded activities – services and manufacturing for the local market – are undertaken in cities). There are both similarities and differences with the conventional Balassa-Samuelson effect across countries in the macroeconomics literature. In the macroeconomics literature, exogenously higher productivity in tradables in more advanced nations leads to a higher relative price of non-traded goods in these countries. In contrast, in our spatial Balassa-Samuelson effect within countries, differences in relative prices and in the structure of economic activity arise endogenously from differences in internal trade costs.[31]

---

[31] The spatial variation in population density and economic structure implied by the spatial Balassa-Samuelson

3.9 External and Internal Integration

We now use the model to examine the effects of external and internal integration on the pattern of economic activity within countries. The key endogenous variables of interest in the model are population density, sectoral specialization, relative factor prices and the relative price of tradables $\{n(\ell), \nu_A(\ell), \omega(\ell), E_T(\ell)/E(\ell)\}$. The model implies two sufficient statistics for each location, which together with the total population of the economy as a whole and land area for each location, determine all endogenous variables in the model: productivity in agriculture adjusted for the tradables consumption price index and productivity in non-agriculture $\{\widetilde{z}_A(\ell), z_N(\ell)\}$. Totally differentiating population density (21) and the agricultural labor share (22), and using the population mobility condition (11), the change in the endogenous variables of the model can be expressed in terms of changes in these two sufficient statistics and in total population $\left\{\widehat{\widetilde{z}_A}(\ell), \hat{z}_N(\ell), \hat{N}\right\}$:

$$\widehat{n}(\ell) = \frac{(\alpha_A - \alpha_N)\nu_A(\ell)}{\alpha_N(1-\alpha_A) + (\alpha_A - \alpha_N)\nu_A(\ell)}\widehat{\nu}_A(\ell) - \widehat{\omega}(\ell), \tag{23}$$

$$\widehat{\nu}_A(\ell) = \left(1 + \frac{\alpha_A - \alpha_N}{1-\alpha_A}\nu_A(\ell)\right)(1-\sigma)\left(\frac{\widehat{E_T(\ell)}}{E(\ell)}\right). \tag{24}$$

where the wage-rental ratio and price of tradable goods changes according to

$$\widehat{\omega}(\ell) = \frac{(1-\alpha_N)\nu_A(\ell)\left(\widehat{u^*} - \widehat{\widetilde{z}_A}(\ell)\right) + (1-\alpha_A)(1-\nu_A(\ell))\left(\widehat{u^*} - \widehat{z}_N(\ell)\right)}{\alpha_A(1-\alpha_N)\nu_A(\ell) + \alpha_N(1-\alpha_A)(1-\nu_A(\ell))}, \tag{25}$$

$$\left(\frac{\widehat{E_T(\ell)}}{E(\ell)}\right) = \frac{(1-\alpha_A)(1-\nu_A(\ell))\left[\alpha_A\widehat{z}_N(\ell) - \alpha_N\widehat{\widetilde{z}_A(\ell)} - (\alpha_A - \alpha_N)\widehat{u^*}\right]}{\alpha_A(1-\alpha_N)\nu_A(\ell) + \alpha_N(1-\alpha_A)(1-\nu_A(\ell))}, \tag{26}$$

and where the aggregate labor market clearing condition implies

$$\sum_{\ell} \nu(\ell)\,\widehat{n}(\ell) = \widehat{N}, \tag{27}$$

where $\nu(\ell) = N(\ell)/N$ is the share of location ℓ in the national population.

In our empirical analysis below, we use the structure of the model to recover the values of the sufficient statistics $\{\widetilde{z}_A(\ell), z_N(\ell)\}$ that are consistent with the observed data being an equilibrium of the model. External and internal integration affect the equilibrium of the model through these sufficient statistics and the total population of the economy as a whole. Immigration directly affects the economy's total population ($\hat{N}$). The change in adjusted agricultural productivity ($\widehat{\widetilde{z}_A}(\ell)$) depends on the change in agricultural productivity ($\widehat{z}_A(\ell)$) and the change in the tradables

---

effect has correlates in the share of exports in income and the distribution of real income across locations. The share of agricultural exports in income can be expressed as $\frac{x_A(\ell)}{y(\ell)} = (1-\gamma_A)\frac{(1-\alpha_N)\nu_A(\ell)}{(1-\alpha_A)+(\alpha_A-\alpha_N)\nu_A(\ell)} = (1-\gamma_A)\beta_T\left(\frac{E_T(\ell)}{E(\ell)}\right)^{1-\sigma}$. From (21), real income per unit of land is $\frac{y(\ell)}{E(\ell)} = \frac{w(\ell)n(\ell)+r(\ell)}{E(\ell)} = \frac{1-\alpha_A+(\alpha_A-\alpha_N)\nu_A(\ell)}{(1-\alpha_A)(1-\alpha_N)}n(\ell)\,u^*$.

consumption price index ($\widehat{E}_T(\ell)$):

$$\widehat{\widetilde{z}}_A(\ell) = \widehat{z}_A(\ell) - \widehat{E}_T(\ell)\,. \tag{28}$$

Changes in aggregate agricultural productivity ($\widehat{z}_A(\ell)$) (from (12)) depend on the shares of land allocated to each agricultural good ($l_g(\ell)$), changes in agricultural productivity ($\widehat{T}_g(\ell)$), and changes in local prices ($\widehat{P}_g(\ell)$). These changes in local prices ($\widehat{P}_g(\ell)$) can be in turn related to changes in internal transport costs ($\widehat{\delta}_g(\ell)$) and prices at the port ($\widehat{P}_g^*$):

$$\begin{aligned}
\widehat{z_A}(\ell) &= \sum_{g=1}^{G} l_g(\ell)\left(\frac{\widehat{T}_g(\ell)}{\theta(\ell)} + \widehat{P_g}(\ell)\right), \\
&= \sum_{g=1}^{G} l_g(\ell)\left(\frac{\widehat{T}_g(\ell)}{\theta(\ell)} + \widehat{P}_g^* - \widehat{\delta}_g(\ell)\right).
\end{aligned} \tag{29}$$

Assuming a Cobb-Douglas tradables consumption index (18), and maintaining the empirically-motivated assumption that manufactures are imported, changes in the tradables consumption price index also depend on changes in internal transport costs ($\widehat{\delta}_g(\ell)$, $\widehat{\delta}_M(\ell)$) and prices at the port ($\widehat{P}_g^*$):

$$\begin{aligned}
\widehat{E_T(\ell)} &= (1-\gamma_A)\,\widehat{P}_M(\ell) + \sum_{g=1}^{G} \gamma_g \widehat{P}_g(\ell)\,, \\
&= (1-\gamma_A)\left[\widehat{P}_M^* - \widehat{\delta}_M(\ell)\right] + \sum_{g=1}^{G} \gamma_g \left[\widehat{P}_g^* - \widehat{\delta}_g(\ell)\right].
\end{aligned} \tag{30}$$

External integration (e.g. a reduction in transatlantic transport costs) increases the price of exported agricultural goods at the port ($\widehat{P}_g^*$) and reduces the price of imported manufacturing goods at the port ($\widehat{P}_M^*$). Internal integration (e.g. construction of the railroad network) reduces internal trade costs and hence increases the local prices of exported agricultural goods ($\widehat{P}_g(\ell)$) and reduces the local prices of imported manufacturing goods ($\widehat{P}_M(\ell)$) for given prices at the port. Both forms of integration reduce the consumption price index ($E_T(\ell)$) and increase aggregate agricultural productivity ($z_A(\ell)$), and hence increase adjusted agricultural productivity ($\widetilde{z}_A(\ell)$). To the extent that external and internal integration also facilitate the diffusion of technology, they also raise adjusted agricultural productivity ($\widetilde{z}_A(\ell)$) through productivity for each agricultural good ($T_g(\ell)$).

Proposition 4. ***(External and Internal Integration)*** *Reductions in external and internal trade costs that raise a location's adjusted agricultural productivity ($\widehat{\widetilde{z}}_A(\ell)$)* ***(a)*** *reduce its wage-rental ratio (lower $\omega(\ell)$),* ***(b)*** *increase its relative price of the non-traded good (lower $E_T(\ell)/E(\ell)$),* ***(c)*** *raise its population density (higher $n(\ell)$),* ***(d)*** *reduce its share of labor in agriculture (lower $\nu_A(\ell)$), and hence* ***(e)*** *increase its urban population share.*

Proof. See the web-based technical appendix. □

4 Data

Our main data sources are the Argentinian population censuses of 1869, 1895 and 1914 (Republica Argentina 1869, 1895, 1914).[32] Data are reported at the level of provinces and districts, where there are 23 provinces and 386 districts in 1895. For each district, we have information on total population, rural population (which we associate with the agricultural sector), urban population (which we associate with the non-traded sector including manufacturing for the local market and services), the number of natives and immigrants, and geographical land area.[33] Some district boundaries change during our sample period: Between 1869 and 1895, there is an expansion in the geographical boundaries of Argentina, while between 1895 and 1914 a number of districts are subdivided. To overcome these boundary changes, we construct constant spatial units over time based on 1895 districts, aggregating subdivisions using the maps and concordance in Cacopardo (1967).

In addition to data on population and employment structure, the 1895 and 1914 population censuses contain district-level information on agricultural land use, which we use to examine the model's predictions for specialization within the agricultural sector. We construct agricultural land area for the following categories: Cereals (primarily wheat, corn and barley), Textiles (Cotton and Linen), Other Crops (e.g. Tobacco, Sugar and Wine), Cattle and Sheep. The 1895 and 1914 censuses also contain district-level data on the number of agricultural machines, including ploughs, mowers, rakes, threshers, water pumps and wind pumps. Finally, we also have some data on agricultural land prices that we use in an external validation exercise.

We combine our population census data with international trade data for 1870, 1895 and 1914.[34] Information is reported on the quantity and value of Argentina's total exports and imports by product.[35] The number of export products increases over time from 30 in 1870 to 124 in 1895 and 199 in 1914, which is consistent with an increase in the range of products exported by Argentina and an increase in statistical disaggregation induced by this increased export sophistication. The number of import products also increases over time, and is substantially larger than the number of export products in each year, which is consistent with Argentina's narrow specialization in the agricultural sector.[36] We concord export products in each year to an aggregated classification of export sectors that is constant over time, as reported in Panel A of Table 1. We also concord export products in each year to a more disaggregated classification corresponding to the main agricultural goods exported by Argentina (e.g. hides, wool, cereals), as reported in Panel B of Table 1.

---

[32] See the data appendix for further discussion of the data definitions and sources.

[33] 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. At the province-level, data are also available on employment by occupation.

[34] See Cuadro General del Comercio Exterior (1870) and Compañía Sud-Americana de Billetes de Banco (1895, 1914). To address the potential concern that 1914 trade flows could be distorted by the outbreak of the First World War in August of that year, we also have compiled data on 1910 trade flows.

[35] We convert export and import values to U.S. dollars using the historical exchange rates from Della Paolera (1988) and Bordo et al. (2001). See also Denzel (2010). To convert these nominal values into 1869 prices, we use the historical GDP deflators from Carter et al. (2006).

[36] The corresponding numbers for the number of import products are: 124 in 1870, 481 in 1895 and 1,190 in 1914.

In addition to these data on total exports and imports, we have information for each year on the quantity and value of exports of each product from each Customs to each foreign destination and to other customs within Arentina. We use these Customs data to examine patterns of external and internal trade.[37] For 1895 and 1914, we also have data on railroad shipments of agricultural goods (including wheat, corn, linen, wool and hides) by railroad station. Using data on the latitude and longitude of these railroad stations, we assign them to districts and construct railroad shipments for each agricultural good for each district.

We combine our population census and international trade data with information from several other sources. We constructed a Geographical Information Systems (GIS) shapefile of 1895 Argentinian districts based on the maps and concordance in Cacopardo (1967). We also constructed GIS shapefiles of the Argentinian railroad network in 1869, 1895 and 1914, navigable rivers, and historical exploration routes in Argentina using the maps in Randle (1981).

5 Quantitative Analysis

In this section, we use the model as a quantitative tool to decompose the observed changes in population density and sectoral specialization into the contributions of aggregate land area, total population and the model's two sufficient statistics for the internal distribution of economic activity: adjusted agricultural productivity and non-agricultural productivity. First, we calibrate the model's parameters. Second, we solve for the unobserved values of the sufficient statistics $\{\widetilde{z}_A(\ell), z_N(\ell)\}$ for which the observed data are an equilibrium of the model. As in Chari, Kehoe and McGrattan (2007) and Hsieh and Klenow (2009), these sufficient statistics correspond to structural residuals or wedges that ensure the model's predictions are exactly consistent with the observed data. These sufficient statistics capture the economic mechanisms through which external and internal integration affect the spatial distribution of economic activity in the model. Third, we undertake counterfactuals for changes in these sufficient statistics that reveal their respective contributions to aggregate welfare, the aggregate urban population share, and the distribution of population density and the urban population share across locations. In the next section, we provide further evidence on the contributions of external and internal integration through these mechanisms.

5.1 Calibration

We assume central values for the model's parameters $\{\alpha_A, \alpha_N, \beta, \sigma\}$ based on the existing empirical literature. We assume shares of land in factor payments in agriculture and non-tradables of 20 percent and 10 percent respectively, which is line with the values in Caselli and Coleman (2001). We assume inelastic demand between agriculture, manufacturing and non-tradables following a large literature in macroeconomics, including for example Ngai and Pissarides (2007). Specifically,

---

[37] The number of Customs also increases over time from 12 in 1870 to 48 in 1895 and 87 in 1914. As the number of Customs increases, Buenos Aires's share of the value of Argentina's exports declines from 89 to 42 percent. Together the four main customs of Bahia Blanca, Buenos Aires, La Plata and Rosario account for more than 75 percent of the value of Argentina's exports in all years in our sample. Imports are even more concentrated in Buenos Aires.

we assume an elasticity of substitution of 0.5 as a central value in the inelastic range of the parameter space between zero and one. We assume a value for the preference weight β_T equal to 0.3 to ensure that the model is consistent with historical shares of tradables in consumer expenditure.

5.2 Recovering the Sufficient Statistics

We now show that there is a one-to-one mapping from the observed data and the model parameters to the unobserved sufficient statistics $\{\widetilde{z}_A(\ell), z_N(\ell)\}$. We use the model's recursive structure to solve for these sufficient statistics and the unobserved endogenous variables of the model (such as the relative price of tradables and the wage-rental ratio). First, the relative price of tradables $\frac{E_T(\ell)}{E(\ell)}$ can be uniquely determined from observed agricultural employment shares $\nu_A(\ell)$ using (22):

$$\frac{E_T(\ell)}{E(\ell)} = \left[\frac{1}{\beta_T}\frac{(1-\alpha_N)\,\nu_A(\ell)}{(1-\alpha_A)+(\alpha_A-\alpha_N)\,\nu_A(\ell)}\right]^{\frac{1}{1-\sigma}}. \tag{31}$$

Second, the wage-rental ratio $\omega(\ell)$ can be uniquely determined from observed agricultural employment shares $\nu_A(\ell)$ and population densities $n(\ell)$ using (22) and (21):

$$\omega(\ell) \;=\; \frac{(1-\alpha_A)(1-\alpha_N)}{\alpha_N(1-\alpha_A)+(\alpha_A-\alpha_N)\,\nu_A(\ell)}\frac{1}{n(\ell)}. \tag{32}$$

Third, productivity in non-tradables $z_N(\ell)$ and adjusted aggregate agricultural productivity $\widetilde{z}_A(\ell)$ can be uniquely determined up to a normalization from the above solutions for the relative price of tradables $(E_T(\ell)/E(\ell))$ and the wage-rental ratio $(\omega(\ell))$ using population mobility (14):

$$\widetilde{z}_A(\ell) = \frac{u^*}{\omega(\ell)^{\alpha_A}}\frac{1}{(E_T(\ell)/E(\ell))}, \tag{33}$$

$$z_N(\ell) = \frac{u^*}{\omega(\ell)^{\alpha_N}}\left(\frac{1-\beta_T}{1-\beta_T\left(\frac{E_T(\ell)}{E(\ell)}\right)^{1-\sigma}}\right)^{\frac{1}{1-\sigma}}, \tag{34}$$

where the normalization involves a choice of units in which to measure the common level of utility across all locations u^*.

To recover the level of productivities $\{\widetilde{z}_A(\ell), z_N(\ell)\}$ from (33) and (34), we impose the following normalizations. We choose 1914 as our base year and obtain the level of productivities in that year by normalizing the common level of utility across all locations u^* to one. We recover the level of productivities for the two other years of our sample by calibrating the growth in the common level of utility across all locations relative to 1914 to the estimates of real wage growth for Argentina as a whole in Taylor and Williamson (1997).[38]

The distributions of employment and population across locations depend solely on the relative value of these productivities, which can be recovered independently of the normalization:

---

[38] The implied normalizations for utility are as follows: $u^*_{1914} = 1$, $u^*_{1895} = 0.79$ and $u^*_{1869} = 0.65$.

$$
\begin{aligned}
\frac{\widetilde{z}_A(\ell)}{\widetilde{z}_A(0)} &= \left(\frac{\omega(0)}{\omega(\ell)}\right)^{\alpha_A} \frac{E_T(0)/E(0)}{E_T(\ell)/E(\ell)}, \\
\frac{z_N(\ell)}{z_N(0)} &= \left(\frac{\omega(0)}{\omega(\ell)}\right)^{\alpha_N} \left(\frac{1-\beta_T\left(\frac{E_T(0)}{E(0)}\right)^{1-\sigma}}{1-\beta_T\left(\frac{E_T(\ell)}{E(\ell)}\right)^{1-\sigma}}\right)^{\frac{1}{1-\sigma}}.
\end{aligned}
\tag{35}
$$

where we have chosen one region (region 0) as the base region.

In Panel A of Figure 6, we display log population density ($n(\ell)$) in 1869 and 1914 against Great Circle distance from Buenos Aires, which from Figures 3-5 provides a summary measure of remoteness from world markets.[39] We show both the data and the fitted values from locally-weighted linear least squares regressions for each year. Despite the reduction in internal trade costs from the large-scale expansion of the railroad network, we find a *steepening* of the gradient of economic activity with remoteness from Buenos Aires over time. In Panel B of Figure 6, we show an analogous figure for the urban population share ($\nu_N(\ell) = N_N(\ell)/N(\ell)$). Although Argentina's export boom and economic development were agriculturally based, we find both an increase in the urban population share and a steepening of its gradient with respect to remoteness over time.

Our theoretical model rationalizes these features of the data through the spatial Balassa-Samuelson effect. In Panels C and D of Figure 6, we display the log relative price of tradables (from (31)) and the log wage-rental ratio (from (32)) against Great Circle distance from Buenos Aires. From the spatial Balassa-Samuelson effect, the steepening of the population density and urbanization gradients in Panels A and B go hand-in-hand with a steepening of the gradients of the relative price of tradables and the relative wage-rental ratio in Panels C and D.

Panel A of Figure 7 graphs adjusted agricultural productivity ($\widetilde{z}_A(\ell) = z_A(\ell)/E_T(\ell)$) against Great Circle Distance from Buenos Aires and again shows the spatial Balassa-Samuelson effect. In the cross-section, the higher population densities and urban population shares of locations close to world markets are rationalized by higher adjusted agricultural productivities (from (31)-(33)). Similarly, over time, the steeper gradients in population density and the urban population share are captured by a steeper gradient in adjusted agricultural productivity. Panel B shows the analogous figure for non-agricultural productivity ($z_N(\ell)$). Given the gradients of the relative price of tradables and the wage-rental ratio implied by the observed data (Panels C and D of Figure 6), the model requires somewhat higher non-agricultural productivity in more remote locations (from (34)), though the gradient is shallower than for adjusted agricultural productivity. Although the calibrated producivities for each sector depend on both population density and the urban popu-

---

[39] As discussed above in subsection 2.3 and footnote 14, Buenos Aires accounts for a disproportionate share of trade, and its development into the main port was driven by its status as the seat of the Viceroyalty of the Río de la Plata rather than its natural suitability as a port. We find similar results using the minimum Great Circle distance to the four main customs of Buenos Aires, Rosario, Bahia Bahia and La Plata as an alternative measure of remoteness. The correlation between the log Great Circle distances to Buenos Aires and to the closest of the four main customs is 0.92 (significant at the 1 percent level). We also find similar results using a measure of lowest-cost effective distance, in which distance by land, water and rail is assigned the weights from Donaldson (2013).

lation share, together with the parameters and equilibrium conditions of the model, we find that adjusted agricultural productivity closely reflects population density (Panel C of Figure 7), while non-agricultural productivity closely reflects the urban population share (Panel D of Figure 7).

5.3 Counterfactuals

We now use the structure of the model to undertake counterfactuals to show the contribution of land area, total population and the two sufficient statistics for the distribution of economic activity $\{\widetilde{z}_A(\ell), z_N(\ell)\}$. To isolate the economic mechanisms in the model, we consider a change in each of these variables in turn. We emphasize that changes in one variable can influence another. For example, reductions in tradables consumption price indices and increases in productivity $\{\widetilde{z}_A(\ell), z_N(\ell)\}$ can attract immigration. Or an increase in population (N) can raise productivity $\{\widetilde{z}_A(\ell), z_N(\ell)\}$ through agglomeration effects. Nevertheless, we can use the structure of the model to isolate effects operating through tradables consumption price indices and productivities $\{\widetilde{z}_A(\ell), z_N(\ell)\}$ and effects operating through land area and the total population (N).

Given counterfactual values for land area, total population and the sufficient statistics $\{\widetilde{z}_A(\ell), z_N(\ell)\}$, we solve for the counterfactual levels of the endogenous variables $\{n(\ell), \omega(\ell), E_T(\ell)/E(\ell), u^*\}$ using the following system of equations:

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

$$\frac{E_T(\ell)}{E(\ell)} = \frac{u^*}{\widetilde{z}_A(\ell)\,\omega(\ell)^{\alpha_A}},$$

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

$$\sum_{\ell} \frac{L(\ell)}{L} n(\ell) = \frac{N}{L}.$$

We solve this system of equations using a shooting algorithm, in which we guess an initial vector for the endogenous variables, solve the system of equations for the new vector of the endogenous variables, and update our guess using a weighted average of the initial and new values. Since the equilibrium of the model is unique (Proposition 1), the solution of this system of equations converges rapidly to that unique equilibrium. Having solved for $\{n(\ell), \omega(\ell), E_T(\ell)/E(\ell), u^*\}$, the equilibrium value of all other endogenous variables can be determined, including the share of agriculture in employment for each location $\nu_A(\ell)$.

Table 3 reports the results of these counterfactuals using 1914 as the base year. In Counterfactual I, we exclude districts that were unpopulated in 1869, but hold productivities in all other districts constant at their 1914 values, and hold the total Argentine population constant at its 1914 value.[40] This first counterfactual captures the effects of the expansion of the frontier of settlement on welfare and the structure of economic activity. Concentrating the 1914 population in a smaller geographical land area reduces real wages by 7 percent (through a higher price of land) and increases the aggregate urban population share by 2 percent (since regions that were settled earlier tend to have relatively higher adjusted agricultural productivity, which reallocates employment towards the urban non-tradable sector).

In Counterfactual II, we exclude districts that were unpopulated in 1869, and adjust the total Argentine population by net immigration from 1869-1914, but hold productivities in all other districts constant at their 1914 values. Therefore the comparison of Counterfactuals I and II captures the effects of immigration. Reducing the total population increases real wages by 10 percentage points (through a lower price of land) and reduces the aggregate urban population share by 1 percent. This small effect on the aggregate urban population share reflects two offsetting effects. On the one hand, a lower population reduces the relative price of land, which favors the land-intensive agricultural sector (reducing the aggregate urban population share). On the other hand, a lower total population alleviates congestion costs in the most-densely populated locations, which redistributes population towards more-densely populated urban areas.

In Counterfactual III, we exclude districts that were unpopulated in 1869, adjust the total Argentine population by net immigration from 1869-1914, and set adjusted agricultural productivities and non-agricultural productivities equal to their 1869 values. This counterfactual comes close to replicating the distribution of economic activity in 1869 but includes native population growth from 1869-1914. Real wages fall by 38 percent and the aggregate urban population share falls by 21 percent. Comparing Counterfactual III with our earlier counterfactuals reveals the importance of changes in productivity and the tradables consumption price index relative to changes in the frontier of settlement and total population.

In Counterfactual IV, we exclude districts that were unpopulated in 1869, adjust the total Argentine population by net immigration from 1869-1914, set adjusted agricultural productivities equal to their 1869 value, but hold non-agricultural productivities constant at their 1914 values. In this counterfactual, real wages fall by 8 percent and the aggregate urban population share falls by around 9 percent. Therefore the effects of changes in productivity and the consumption price index within the tradables sector are of around the same magnitude as changes in the total population through immigration (around 2.94 million people from 1869-1914 or around one third of the 1914 Argentine population). The difference in predicted effects between Counterfactuals III and IV reflects both the impact of changes in non-agricultural productivities and their covariance with changes in agricultural productivities across locations.

While we have so far focused on the aggregate predictions of these counterfactuals for the real wage and the urban population share, we now turn to their disaggregate predictions for the spatial distribution of economic activity. In Panel A of Figure 8, we graph actual and counterfactual

---

[40] As discussed above, zero population in a district is rationalized by zero productivity in tradables. In the extension of the model in the web appendix, in which we introduce an endogenous conversion of land to productive use, a location also may have zero population because it is not profitable to convert land to productive use.

population densities against log distance from Buenos Aires for each district in Counterfactuals I-II. In Panel B, we display analogous information for urban population shares against log distance from Buenos Aires. Panels C and D display the same information for Counterfactuals III-IV. In Panels A-B, we find that foreign migration increases population density and leads to a modest steepening of the gradient of population density with respect to geographical remotenesss, but leaves the urban population share unchanged. In Panels C-D, we find that changes in adjusted agricultural productivity ($\widetilde{z}_A \ell$) and non-agricultural productivity ($z_N(\ell)$) account for most of the steepening of the gradient of population density and nearly all of the steepening of the gradient of the urban population share. Comparing Counterfactuals III-IV in Panels C-D, changes in adjusted agricultural productivity make a substantial contribution to the overall productivity effect.

6 Further Evidence

The results of the previous section establish the quantitative relevance of changes in adjusted agricutural productivity ($\widetilde{z_A}(\ell) = z_A(\ell)/E_T(\ell)$) for aggregate welfare, urbanization and the distribution of economic activity across sectors and regions. Furthermore, we find a steepening of the gradient of adjusted agricultural productivity with respect to geographical remoteness over time. In this section, we provide further evidence on the role of the economic mechanisms in the model in explaining these findings.

6.1 Railroads, Technology and Agricultural Specialization

In the model, adjusted agricultural productivity ($\widetilde{z_A}(\ell)$) depends on the prices of agricultural goods ($P_g(\ell)$), the technologies for agricultural goods ($T_g(\ell)$), and the prices of imported manufactures (P_M). Combining (28) and (29), we can decompose changes in adjusted agricultural productivity into the contributions of changes in each of these components:

$$\widehat{\widetilde{z}}_A(\ell) = \underbrace{-(1-\gamma_A)\widehat{\delta}_M(\ell) - \sum_{g=1}^{G}(l_g-\gamma_g)\widehat{\delta}_g(\ell)}_{\textit{Trade Costs}} + \underbrace{\sum_{g=1}^{G}(l_g(\ell)-\gamma_g)\widehat{P}_g^* - (1-\gamma_A)\widehat{P}_M^*}_{\textit{Terms of Trade}} + \underbrace{\sum_{g=1}^{G} l_g(\ell)\left(\frac{\widehat{T}_g(\ell)}{\theta(\ell)}\right)}_{\textit{Technology}},$$

which, in the special case of common trade costs for all goods $\widehat{\delta}_g(\ell) = \widehat{\delta}_M(\ell) = \widehat{\delta}(\ell)$, simplifies to:

$$\widehat{\widetilde{z}}_A(\ell) = \underbrace{-2(1-\gamma_A)\widehat{\delta}(\ell)}_{\textit{Trade Costs}} + \underbrace{\sum_{g=1}^{G}(l_g(\ell)-\gamma_g)\widehat{P}_g^* - (1-\gamma_A)\widehat{P}_M^*}_{\textit{Terms of Trade}} + \underbrace{\sum_{g=1}^{G} l_g(\ell)\left(\frac{\widehat{T}_g(\ell)}{\theta(\ell)}\right)}_{\textit{Technology}}.$$

Reductions in internal trade costs from the expansion of the railroad network would be expected to increase the price of exported agricultural goods and reduce the price of imported manufacturing goods, which would *flatten* the spatial gradient in adjusted agricultural productivity over time. In contrast to these predictions, we observe a *steepening* of the spatial gradient of adjusted agricultural productivity over time. One potential explanation for this steepening is that the expansion of the railroad network is geographically uneven and concentrated in locations close to Buenos Aires. To the extent that these regions experience larger increases in the prices of exported agricultural goods ($P_g(\ell)$) and larger reductions in the prices of imported manufactures (P_M), they would also experience larger increases in adjusted agricultural productivity. In Panels A-B of Figure 9, we examine the geographical diffusion of the railroad network, measured in terms of either stations per square kilometer (Panel A) or the length of railroad per square kilometer (Panel B). Consistent with this explanation, the regions closest to Buenos Aires indeed experience the largest increases in both station and railroad density.

Another potential explanation for the observed steepening of the spatial gradient of adjusted agricultural productivity is that technological change ($T_g(\ell)$) is geographically uneven and concentrated in the regions closest to Buenos Aires. In Panel C of Figure 9, we examine the geographical diffusion of technology using data on the number of agricultural machines per square kilometer (ploughs, mowers, rakes, threshers, water machines and wind machines), which are available for 1895 and 1914. Consistent with this explanation, the regions closest to Buenos Aires experience the largest increases in the density of agricultural machines. As shown in Panel of D of Figure 9, we also find that the geographical diffusion of technology and railroads are closely related to one another. This close relationship is consistent with railroad construction enhancing the incentives for technology adoption and reducing the cost of importing agricultural machinery (causality running from railroads to machnery adoption). But it is also consistent with the construction of the railroad network being endogenous to agricultural productivity (causality running from a third variable, agricultural productivity, to both railroads and machinery adoption). We provide further evidence on these two potential explanations below.

To the extent that changes in prices ($P_g(\ell)$) and technology ($T_g(\ell)$) are uneven across agricultural goods, the model predicts changes in comparative advantage and specialization across goods within the agricultural sector. From (19) and (20), relative exports and relative land shares for agricultural goods depend on a combination of relative technologies and a power function of relative prices. In Panels A-E of Figure 10 we examine agricultural specialization across locations and over time using data on the shares of agricultural land area allocated to cereals, textiles, other crops, cattle and sheep. Consistent with the comparative-advantage-based mechanism in the model, we find substantial changes in agricultural specialization over time, with a reallocation towards cattle (Panel A) and Cereals (Panel B) and away from sheep (Panel C). The increase in cattle farming is largest in locations close to Buenos Aires (Panel A), which is consistent with the new technologies for freezing and chilling beef reorientating the cattle industry towards meat as opposed to hides.[41] The increase in cereals cultivation is largest at intermediate distances from Buenos Aires (Panel B), which is in line with the historical literature on the role of the "Conquest of the Desert" and

---

[41] As emphasized in the historical literature, cattle farming for meat implies a greater incentive to reduce the distance travelled by each animal (to reduce the loss of muscle and fat), and hence to concentrate cattle farming close to rail connections and meat processing plants. Consistent with this hypothesis, we find a reallocation towards pure-breed cattle (with higher quality meat and higher ratios of meat to animal weight) and away from mixed and native cattle in the locations closest to Buenos Aires and the meat processing plants reported in the 1914 census.

the expansion of the railroad network in opening up the agricultural hinterland for cereals production.[42] There is also an increase in textiles cultivation at intermediate distances from Buenos Aires (Panel D) and a decline in the cultivation of other crops (Panel E) in response to these changes in comparative advantage.

6.2 External Validity

The model imposes additional discipline on these economic mechanisms that provides the basis for external validity checks. From agricultural land shares (17) and agricultural export shares (19), the model predicts an approximately log linear relationship between relative export values and relative land shares, which becomes exact as the domestic consumption share for each good converges towards zero ($\gamma_g, \gamma'_g \to 0$):

$$\frac{x_g(\ell)}{x'_g(\ell)} = \frac{l_g(\ell) - \gamma_g}{l'_g(\ell) - \gamma'_g}.$$

In Panels A-D of Figure 11, we examine these predictions by combining data on shipments from railroad stations within each district of Wheat, Corn, Linen, Wool and Hides with data on the shares of agricultural land used for Wheat, Corn, Linen, Sheep and Cattle. The shipments data are only available for districts containing railroad stations and correspond to the quantity rather than the value of each good. Nonetheless, we find a tight relationship between log relative railroad shipments and log relative land shares, with for example an elasticity (standard error) of 0.9515 (0.0721) and an R-squared of 0.609 in Panel A. The relationship is less tight for hides and the agricultural land share for cattle in Panel D, which is consistent with some hides coming from animals other than cattle (in particular sheep). Summing railroad shipments across districts, we find that total railroad shipments are substantial relative to export quantities for each agricultural good, confirming the external orientation of agricultural production.

From the Cobb-Douglas production technology, the model also implies a close connection between land prices and agricultural productivity:[43]

$$r(\ell) = z_A(\ell)\, \omega(\ell)^{\alpha_A - 1}. \tag{36}$$

In Panel A of Figure 12, we examine this prediction using data on land prices (available for 1895) and our measure of adjusted agricultural productivity ($\widetilde{z}_A(\ell) = z_A(\ell)/E_T(\ell)$) to proxy for agricultural productivity ($z_A(\ell)$). Although this proxy also captures the tradables consumption price index, we find an approximately log linear relationship between the two variables, with an elasticity (standard error) of 1.4508 (0.2343) and an R-squared 0.2713. Therefore the higher adjusted agricultural productivity of locations close to world markets is indeed reflected in a higher price of the immobile factor land.

---

[42] For a comparison of the expansion of the wheat frontier in Argentina and Canada, see Adelman (1994).

[43] This expression for $r(\ell)$ follows from evaluating (11) at $i = A$.

6.3 Internal Investments

We now provide further evidence on the role of internal investments – in the form of railroad construction and the adoption of agricultural machinery – in understanding the economy's response to external integration. As shown in Panel D of Figure 9, station density and machinery density are highly correlated with one another. Therefore we use a Principal Components Analysis to construct a composite measure that incorporates both of these dimensions of internal investments. In Panels B-C of Figure 12, we display station density and machinery density against the first component from this Principal Components Analysis. Although the first component is more strongly related to station density than to machinery density, it captures the variation in both of these measures of internal investments, and accounts for around 80 percent of the overall variance.

We first show that much of the steepening of the spatial gradient of adjusted agricultural productivity can be explained statistically by this measure of internal investments. In Column (1) of Table 4, we regress adjusted agricultural productivity growth from 1869-1914 on log Great Circle distance from Buenos Aires. We find a negative estimated coefficient that is statistically significant at the one percent level, which captures the steepening of the spatial gradient of adjusted agricultural productivity over time. In Column (2), we include the initial value of adjusted agricultural productivity in 1869 to control for persistent differences across locations in agricultural productivity and initial levels of economic development. We continue to find a negative and statistically significant coefficient, which if anything increases in magnitude now that we have controlled for initial adjusted agricultural productivity.[44]

In Column (3), we augment the regression with our measure of investment density for 1914. As shown in Figure 3, the railroad network was of limited extent in 1869 and agriculture at that time was dominated by cattle and sheep ranching with limited machinery requirements (as measured by ploughs, mowers, rakes, threshers, water pumps and wind pumps). Therefore the level of investment density for 1914 captures the growth of internal investments from 1869-1914. We find a positive estimated coefficient on investment density that is significant at the one percent level, which is consistent with railroad expansion reducing the tradables consumption price index, and with both railroad expansion and agricultural technology adoption raising aggregate agricultural productivity. Furthermore, the coefficient on distance from Buenos Aires halves in magnitude, suggesting that a substantial part of the steepening of the gradient of adjusted agricultural productivity can be statistically explained in terms of internal investments.

This specification shows that internal investments help to statistically explain the steepening gradient but does not have a causal interpretation. We expect causality to run in both directions between adjusted agricultural productivity and internal investments. On the one hand, internal investments in railroads can reduce internal transport costs, which raises local prices and increases adjusted agricultural productivity. Or internal investments in machinery adoption can directly

---

[44] We experimented with including measures of agricultural suitability based on crop growing cycles, soil characteristics and climatic conditions from the United Nations, but these variables are not significant after controlling for initial adjusted agricultural productivity.

raise agricultural productivity (causality running from internal investments to adjusted agricultural productivity). On the other hand, external transport cost reductions that increase export prices raise adjusted agricultural productivity, which increases the return to internal investments (causality running from adjusted agricultural productivity to internal investments). In the remaining columns of the table, we provide further evidence on the timing of the relationship between adjusted agricultural productivity growth and internal investments. In Columns (4) and (5), we re-estimate the specification from Column (3) for the sub-periods 1869-1895 and 1895-1914. In both cases, internal investments for a given sub-period help to explain the steepening of the gradient of adjusted agricultural productivity within that sub-period. For the second sub-period, the coefficient on distance to Buenos Aires becomes statistically insignificant from 1895-1914, so that all of the steepening gradient is statistically explained by internal investments. In Column (6), we consider a placebo specification, in which we regress adjusted agricultural productivity growth from 1869-1895 on future internal investments from 1895-1914. We find that future internal investments have no predictive power for past adjusted agricultural productivity growth, which casts doubt on the existence of pre-trends in adjusted agricultural productivity growth between locations that subsequently differ in their internal investments.

Even in the absence of pre-trends, there could be omitted third variables (e.g. shocks to agricultural productivity) that both increase internal investments and raise adjusted agricultural productivity. To address this concern and provide causal evidence on the contribution of internal investments towards changes in adjusted agricultural productivity, we use an instrumental variable for railroad construction, which directly affects station density and hence indirectly raises the return to adopting agricultural machinery. To address the non-random assignment of railroads, our instrument uses the idea that locations can be treated with transport infrastructure, not because of their own unobserved characteristics, but because they happen to lie along the route between other locations (see Chandra and Thompson 2000 and Michaels 2008).

In particular, the development of 16th-century colonial cities within the Spanish empire was shaped by trading routes towards the North-West through Bolivia. However, once these cities existed, there was an incentive to connect them to the 19th-century railroad network. As a result, locations that happened to be on the shortest route between other locations and 16th-century cities were disproportionately likely to obtain a railroad connection, independently of their own unobserved characteristics. Based on this idea, we construct our instrument as follows. We discretize the map of Argentina into 136,795 pixels, and compute the shortest route across these pixels from the center of each district to the nearest 16th-century city. For each district, we define "Route C16" as the fraction of pixels within that district that lie along these shortest routes between the centers of districts and 16th-century cities. This measure has a median of 0.17, a mean of 0.28, and a standard deviation of 0.27 across districts.[45]

In Column (1) of Table 5, we reproduce our OLS specification from Column (3) of the previous table, which includes distance to Buenos Aires, initial adjusted agricultural productivity and our

---

[45] See the data appendix for the Spanish cities that were founded in the 16th century (including Buenos Aires).

measure of investment density. In Column (2) of Table 5, we re-estimate this specification instrumenting internal investments with our Route C16 measure. After controlling for historical levels of economic development and geographical location, our identifying assumption is that the frequency with which a district lies on the shortest route to a 16th-century city only affects adjusted agricultural productivity growth from 1869-1914 through internal investments in railroad construction and agricultural technology adoption.

As shown in Column (2), we again find a positive estimated coefficient on investment density that is significant at the one percent level. The IV coefficient is somewhat larger than the OLS coefficient, which is consistent with larger effects from quasi-experimental assignment of railroads than from the assignment of railroads based on the existing political economy process.[46] Furthermore, the coefficient on distance to Buenos Aires is no longer statistically significant. Therefore, using quasi-experiment variation in internal investments and including our controls, there is no longer any evidence of a steepening in the gradient of adjusted agricultural productivity. In Column (3), we report the first-stage regression. We find that districts closer to Buenos Aires, districts with higher initial adjusted agricultural productivity, and districts that frequently lie along the shortest route to 16th-century cities have statistically significantly higher investment densities. To provide evidence of the power of the instrument, Column (2) reports the p-value for the Kleibergen-Paap LM test, in which we reject the null hypothesis of underidentification. Column (3) reports the F-statistic for the significance of the excluded exogenous variable in the first-stage regression, which exceeds the recommended threshold of 10 from Stock, Wright and Yogo (2002).[47]

In Columns (4)-(5), we consider a more conservative specification, in which we exclude the districts that contain 16th-century cities. Therefore, in this specification, we focus solely on differences in the frequency with which other districts lie along the route to these 16th-century cities. After controlling for historical levels of economic development and geographical location, we again we find a positive estimated coefficient on investment density that is significant at the 1 percent level, and the coefficient on distance to Buenos Aires is again no longer significant.

To illustrate the quantitative relevance of these internal investments in transport infrastructure and technology adoption for aggregate economic development, we return to our framework for undertaking counterfactuals in subsection 5.3. We consider a modified version of Counterfactual IV, in which we exclude districts that were unpopulated in 1869, adjust the total Argentine population by net immigration from 1869-1914, change adjusted agricultural productivities, but hold non-agricultural productivities constant at their 1914 values. Instead of setting adjusted agricultural productivities equal to their 1869 values, we remove the predicted impact of internal investments from 1869-1914 from adjusted agricultural productivities, using the estimated coefficient on internal investments from our IV specification in Column (4) of Table 5. Real wages and the aggregate urban share fall by 6 and 5 percent respectively, which are substantial effects relative to the overall

---

[46] For example, there was political pressure for non-economic reasons to construct railroads to open up the interior regions of Argentina and improve transport connections with Chile, as discussed in Lewis (1983).

[47] We find that the instrument has power for both components of investment density. Estimating separate first stage regressions for station and machinery density, we find first-stage F-statistics of 10.36 and 10.58 respectively.

reductions for Counterfactual IV of 8 and 9 percent respectively in Table 3.

Taking the results of this section together, we find evidence in support of the mechanisms in the model, and find that changes in adjusted agricultural productivity are indeed related to internal investments in transport infrastructure and technology adoption.

7 Conclusions

We provide new theory and evidence on the relationship between external trade, structural transformation and economic development. We develop a tractable general equilibrium model of the distribution of economic activity across regions and sectors within countries that is amenable to quantitative analysis. Instead of viewing countries in the aggregate, we model the internal reallocations of resources across regions and sectors that are central to economic development. We combine this quantitative model with a large-scale source of exogenous variation in external integration from the natural experiment of Argentina's integration into world markets in the late-nineteenth century.

The model highlights a *spatial Balassa-Samuelson effect*, in which locations close to world markets have high population densities, high shares of employment in the non-traded sector, high relative prices of non-traded goods, and high land prices relative to wages. These predictions emerge from a model with standard neoclassical ingredients, because the tradables sector is land intensive and demand between tradables and non-tradables is inelastic. The intuition for the spatial Balassa-Samuelson effect is that locations close to world markets have a low relative price of tradables, which attracts population and increases the share of expenditure on non-traded goods if demand is inelastic. This combination of a higher population and a higher share of expenditure on non-traded goods increases the demand for non-traded goods, which implies a high relative price of non-tradables and a high share of employment in non-traded sector. Therefore both the tradable and non-tradable sectors must use more labor-intensive techniques to satisfy land market clearing, which implies high land prices relative to wages.

We find strong empirical confirmation of this spatial Balassa-Samuelson effect. Although Argentina's late 19th-century export boom is based on agriculture, there is a rise in the share of the urban population over time. Despite the reduction in internal transport costs from the large-scale expansion of the railroad network, there is a steepening of the gradients of both population density and the urban population share with geographical remoteness over time. We show how the structure of the model can be used together with observed data on population density and the urban population share to recover two sufficient statistics for the distribution of economic activity: (a) productivity in the export sector (agriculture) adjusted by the tradables consumption price index (including imported manufactures) and (b) productivity in non-tradables. We undertake counterfactuals to show that the effects of changes in adjusted agricultural productivity are quantitatively large relative to changes in land area and total population in understanding changes in the spatial distribution of economic activity within Argentina. We provide evidence connecting the changes in adjusted agricultural productivity to the underlying economic mechanisms in the model.

Our findings highlight the role of internal trade costs in hampering the ability of interior regions to participate in world markets. Furthermore, our results suggest that reductions in external and internal trade costs need not reduce spatial disparities within countries. We link the steepening of the spatial gradient of economic activity to geographically uneven investments in transport infrastructure and technology adoption that were concentrated in locations close to world markets. Thus our analysis points towards the role of these complementary internal investments in mediating the economy's response to external integration.

References

[1] Adamopoulos, Tasso (2011) "Transportation Costs, Agricultural Productivity and Cross-Country Income Differences," International Economic Review, 52(2), 489-521.

[2] Adelman, Jeremy (1994) *Frontier Development: Land, Labor and Capital on the Wheatlands of Argentina and Canada 1890-1914*, Oxford: Oxford University Press.

[3] Amaral, Samuel (2002) *The Rise of Capitalism on the Pampas: The Estancia*, Cambridge: Cambridge University Press.

[4] Arkolakis, Costas and Treb Allen (2013) "Trade and the Topography of the Spatial Economy," Yale University, mimeograph.

[5] Atkin, David and David Donaldson (2012) "Who's Getting Globalized? The Size and Nature of Intranational Trade Costs," mimeograph.

[6] Banerjee, Abhijit, Esther Duflo and Nancy Qian (2012) "On the Road: Transportation Infrastructure and Economic Development," *NBER Working Paper*, 17897.

[7] Baum Snow, Nathaniel (2007) "Did Highways Cause Suburbanization?" *Quarterly Journal of Economics*, 122(2), 775-805.

[8] Balassa, Bela (1964) "The Purchasing Power Parity Doctrine: A Reappraisal," *Journal of Political Economy,* 72(6), pp. 584-96.

[9] Berlinski, Samuel, Sebastian Galiani and Laura Jaitman (2011) "Impact of Railways: Expansion on Spatial Population Patterns," *Lincoln Institute of Land Policy Working Paper*, WP11SB1.

[10] 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.

[11] Bernhofen, Daniel M. and Brown, John C. (2004) "A Direct Test of the Theory of Comparative Advantage: The Case of Japan," *Journal of Political Economy*, 112(1), 48-67.

[12] Bernhofen, Daniel M. and Brown, John C. (2005) "An Empirical Assessment of the Comparative Advantage Gains from Trade: Evidence from Japan," *American Economic Review*, 95(1), 208-225.

[13] Bordo, Michael, Barry Eichengreen, Daniela Klingebiel and Soledad Maria Martinez Peria (2001) "Is the Crisis Problem Growing More Severe?" *Economic Policy*, 24, 51-82.

[14] Bustos, Paula, Bruno Caprettini and Jacopo Ponticelli (2012) "Agricultural Productivity and Structural Transformation: Evidence from Brazil," Universidad Pompeu Fabra, mimeograph.

[15] Cacopardo, María Cristina (1967) *Republica Argentina: Cambios en Los Limites Nacionales, Provinciales y Departmentales, A Traves de Los Censos Nacionales de Poblacion*, Buenos Aires: Instituto Torcuato Di Tella.

[16] Caliendo, Lorenzo, Fernando Parro, Esteban Rossi-Hansberg and Pierre-Daniel Sarte (2013) "The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy," Princeton University, mimeograph.

[17] Carter, Susan B., Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch and Gavin Wright (2006) *Historical Statistics of the United States Millennial Edition*, Cambridge: Cambridge University Press.

[18] Caselli, Francesco and Wilbur John Coleman II (2001) "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation," *Journal of Political Economy*, 109(3), 584-616.

[19] Chandra, Amitabh and Eric Thompson (2000) "Does Public Infractsructure Affect Economic Activity? Evidence from the Rural Interstate Highway System," *Regional Science and Urban Economics*, 30(4), 457-490.

[20] Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan (2007) "Business Cycle Accounting," *Econometrica*, 75 (3), 781–836.

[21] Compañía Sud-Americana de Billetes de Banco (1895) *Anuario de la Dirección General de Estadística Correspondiente al Año 1895*, Buenos Aires.

[22] Compañía Sud-Americana de Billetes de Banco (1914) *Anuario de la Dirección General de Estadística Correspondiente al Año 1914*, Buenos Aires.

[23] Cortés Conde, Roberto (1993) "The Growth of the Argentine Economy, 1870-1914," Chapter 2 in (ed) Leslie Bethell, *Argentina Since Independence*, Cambridge: Cambridge University Press.

[24] Coşar, Kerem and Banu Demir (2014) "Domestic Road Infrastructure and International Trade: Evidence from Turkey," Chicago Booth School of Business, mimeograph.

[25] Coşar, Kerem and Pablo Fajgelbaum (2012) "Internal Geography, International Trade, and Regional Specialization," UCLA, mimeograph.

[26] Cuadro General del Comercio Exterior (1870) *Estadística Comercial de la República Argentina*, Buenos Aires.

[27] Davis, Donald R. and David E. Weinstein (2002) "Bones, Bombs, and Break Points: The Geography of Economic Activity," *American Economic Review*, 92(5), 1269-1289.

[28] Davis, Donald R. and David E. Weinstein (2008) "A Search for Multiple Equilibria in Urban Industrial Structure," *Journal of Regional Science*, 48(1), 29-65.

[29] Della Paolera, Gerardo (1988) *How the Argentine Economy Peformed during the International Gold Standard: A Reexamination*, PhD Dissertation, University of Chicago.

[30] Denzel, Markus A. (2010) *Handbook of World Exchange Rates, 1590-1914*, Farnham: Ashgate.

[31] Donaldson, Dave (2013) "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, forthcoming.

[32] Donaldson, Dave and Richard Hornbeck (2012) "Railroads and American Economic Growth: A Market Access Approach," MIT, mimeograph.

[33] Droller, Federico (2013) "Migration and Long-run Development: Evidence from Settlements in the Pampas," Brown University, mimeograph.

[34] Duranton, Gilles, Peter Morrow and Matthew Turner (2013) "Roads and Trade: Evidence from the US," University of Toronto, mimeograph.

[35] Duranton, Gilles and Matthew Turner (2011) "The Fundamental Law of Highway Congestion: Evidence from the U.S." *American Economic Review*, 101(6), 2616-2652.

[36] Duranton, Gilles and Matthew Turner (2012) "Urban Growth and Transportation," *Review of Economic Studies*, 79(4), 1407-1440.

[37] Eaton, Jonathan and Kortum, Samuel (2002) "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741-1779.

[38] Faber, Ben S. (2013) "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System," London School of Economics, mimeograph.

[39] Foster, Andrew D. and Mark R. Rosenzweig (2007) "Economic Development and the Decline of Agricultural Employment," Chapter 47 in (eds) T. Paul Schultz and John A. Strauss, *Handbook of Development Economics*, Volume 4, Amsterdam: Elsevier, 3051-3083.

[40] Fujita, Masahisa, Paul Krugman, and Anthony J. Venables (1999) *The Spatial Economy: Cities, Regions and International Trade*, Cambridge: MIT Press.

[41] Gollin, Doug, David Lagakos and Michael E. Waugh (2013) "Worker Heterogeneity, Selection, and the Gains from Trade in Agriculture," NYU Stern, mimeograph.

[42] Gollin, Douglas and Richard Rogerson (2012) "Productivity, Transport Costs and Subsistence Agriculture," Princeton, mimeograph.

[43] Gollin, Douglas, Rémi Jedwab and Dietrich Vollrath (2012) "Urbanization with and without Structural Transformation," Williams College, mimeograph.

[44] Hanson, Gordon H. (1996a) "Localization Economies, Vertical Organization, and Trade," *American Economic Review*, 86(5), 1266-78.

[45] Hanson, Gordon H. (1996b) "Economic Integration, Intra-industry Trade, and Frontier Regions," *European Economic Review*, 40(3-5), 941-49.

[46] Herrendorf, Berthold, Richard Rogerson and Ákos Valentinyi (2013) "Growth and Structural Transformation," *Handbook of Economic Growth*, forthcoming.

[47] Herrendorf, Berthold, James A. Schmitz and Arilton Teixeira (2012) "Transportation and Development: Insights from the U.S., 1840-1860," *International Economic Review*, 53(3), 693-716.

[48] Hsieh, Chang-Tai and Peter Klenow (2009) "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, November, 1403-1448.

[49] Inter-American Development Bank (2013) Too Far To Export: Domestic Transport Costs and Regional Export Disparities in Latin America and the Caribbean, Washington DC: Inter-American Development Bank.

[50] Karádi, Peter and Miklós Koren (2013) "Cattle, Steaks and Restaurants: Development Accounting when Space Matters," Central European University, mimeograph.

[51] Kovak, Brian (2013) "Regional Efects of Trade Reform: What is the Correct Measure of Liberalization?" *American Economic Review*, 103(5), 1960-1976.

[52] Lagakos, David and Michael E. Waugh (2013) "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economic Review*, 103(2), 948-980.

[53] Lewis, Colin M. (1983) *British Railways in Argentina 1857-1914: A Case Study of Foreign Investment*, Institute for Latin American Studies: London.

[54] Martincus, Christian Volpe, Jerónimo Carballo Ana Cusolito (2012) "Routes, Exports, and Employment in Developing Countries: Following the Trace of the Inca Roads," Inter-American Development Bank, mimeograph.

[55] Matsuyama, Kiminori (1992) "Agricultural Productivity, Comparative Advantage and Economic Growth," *Journal of Economic Theory*, 58, 317-334.

[56] Michaels, Guy (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.

[57] Michaels, Guy, Ferdinand Rauch and Stephen Redding (2012) "Urbanization and Structural Transformation," *Quarterly Journal of Economics*, 127, 535-586.

[58] Ngai, Rachel and Chris Pissarides (2007) "Structural Change in a Multisector Model of Growth," *American Economic Review*, 97(1), 429-443.

[59] North, Douglass (1958) "Freight Rates and Economic Development 1750-1913," *Journal of Economic History*, 18(4), 537-555.

[60] O'Rourke, Kevin H. and Jeffrey G. Williamson (1991) *Globalization and History: The Evolution of a Nineteenth-Century Atlantic Economy*, MIT Press.

[61] Ramondo, Natalia, Andres Rodriguez-Clare and Milagro Saborio (2012) "Scale Effects and Productivity: Does Country Size Matter?" U.C. Berkeley, mimeograph.

[62] Randle, Patricio H. (1981) *Atlas del Desarrollo Territorial de la Argentina*, Buenos Aires: Oikos.

[63] República Argentina (1869) *Primer Censo de la República Argentina*, Buenos Aires: Superintendente del Censo.

[64] República Argentina (1895) *Segundo Censo de la República Argentina*, Buenos Aires: Comision Directiva.

[65] República Argentina (1914) *Tercer Censo Nacional*, Buenos Aires: Comisión Nacional.

[66] Redding, Stephen J. (2012) "Goods Trade, Factor Mobility and Welfare," *NBER Working Paper*, 18008, April.

[67] Redding, Stephen J. and Daniel M. Sturm (2008) "The Costs of Remoteness: Evidence from German Division and Reunification," *American Economic Review*, 98(5), 1766-1797.

[68] Rocchi, Fernando (2006) *Chimneys in the Desert: Industrialization in Argentina During the Export Boom Years 1870-1930*, Stanford: Stanford University Press

[69] Samuelson, Paul (1964) "Theoretical Notes on Trade Problems," *Review of Economics and Statistics*, 46(2), pp. 145-154.

[70] Scobie, James R. (1971) *Argentina: A City and a Nation*, New York: Oxford University Press.

[71] Sotelo, Sebastian (2014) "Trade Frictions and Agricultural Productivity," University of Michigan, mimeograph.

[72] Stock, James H., Jonathan H. Wright and Motohiro Yogo (2002) "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business and Economic Statistics*, 20(4), 518-529.

[73] Swiecki, Tomasz (2013) "Intersectoral Distortions, Structural Change and the Welfare Gains from Trade," Vancouver School of Economics, mimeograph.

[74] Taylor, Alan (1992) "External Dependence, Demographic Burdens, and Argentine Economic Decline After the Belle Époque," *Journal of Economic History*, 52(4), 907-936.

[75] Taylor, Alan (1997) "Peopling the Pampa: On the Impact of Mass Migration to the River Plate, 1870–1914," *Explorations in Economic History*, 34, 100-32.

[76] Taylor, Alan and Jeffrey G. Williamson (1997) "Convergence in the Age of Mass Migration," *European Review of Economic History*, I, 27-63.

[77] Topalova, Petia (2010) "Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India," *American Economic Journal: Applied Economics*, 2(4), 1-41.

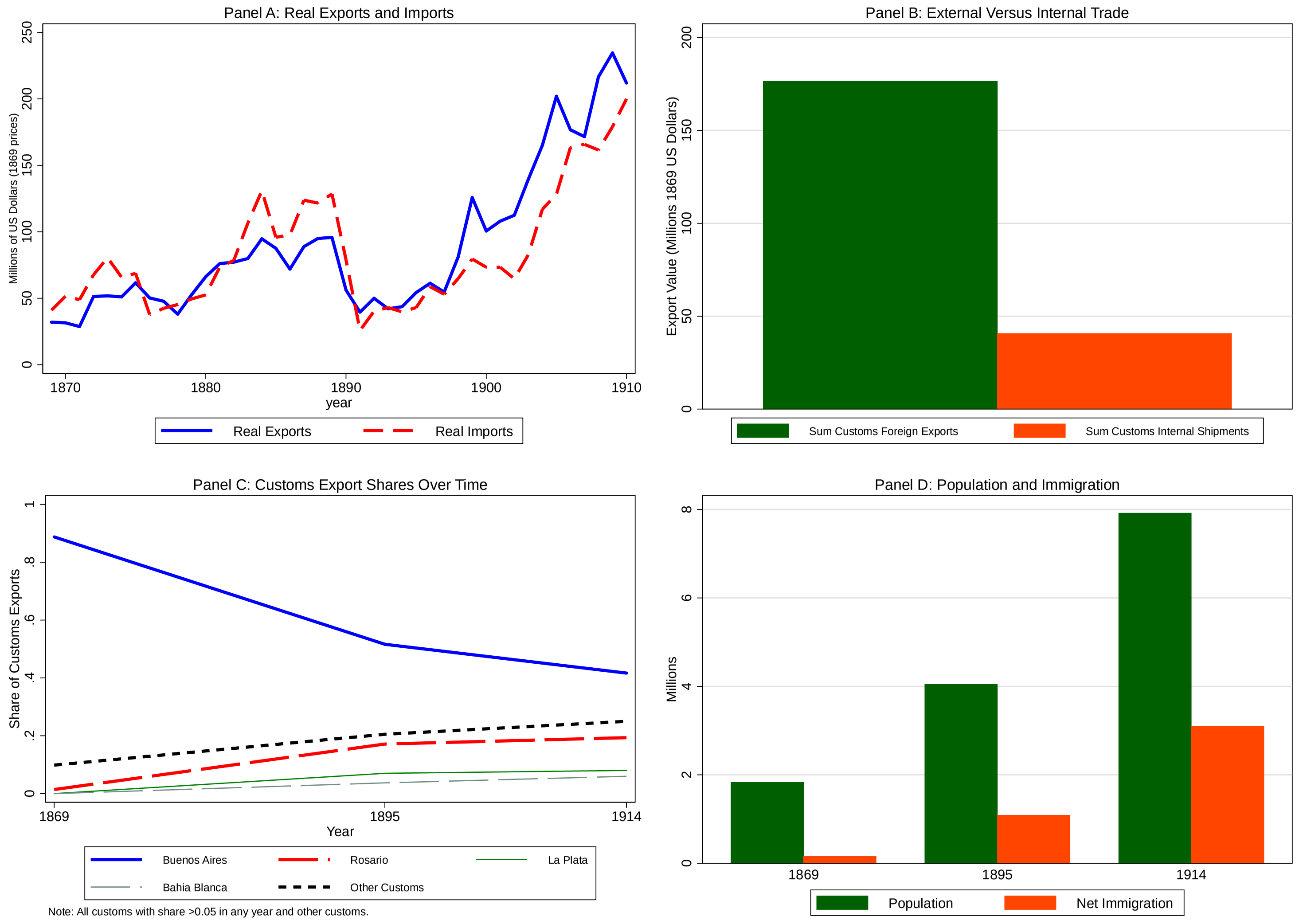
[78] United Nations Economic and Social Council and United Nations Economic Commission for Africa (2009) *The Transport Situation in Africa*, Addis Ababa: Ethiopia, Sixth Session of the Committee on Trade, Regional Cooperation and Integration.

[79] Uy, Timothy, Kei-Mu Yi, and Jing Zhang (2012) "Structural Change in an Open Economy," *Journal of Monetary Economics.*

[80] World Bank (2009) *World Development Report: Reshaping Economic Geography*, Washington DC: World Bank.

[81] World Bank (2001) *Africa's Transport Infrastructure Mainstreaming Maintenance and Management*, Washington DC: World Bank.

[82] World Trade Organization (2004) *World Trade Report: Exploring the Linkage Between the Domestic Policy Environment and International Trade*, Geneva: World Trade Organization.



Note: All customs with share >0.05 in any year and other customs.

Figure 1: Trade and Immigration

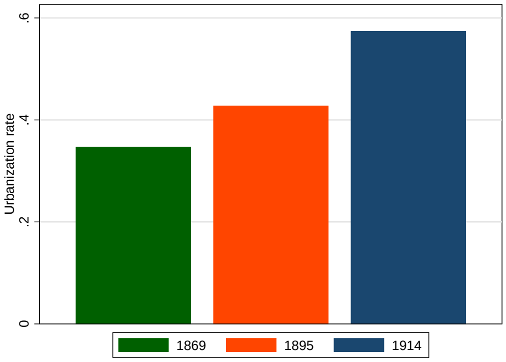


Figure 2: Share of Population Living in Cities and Towns

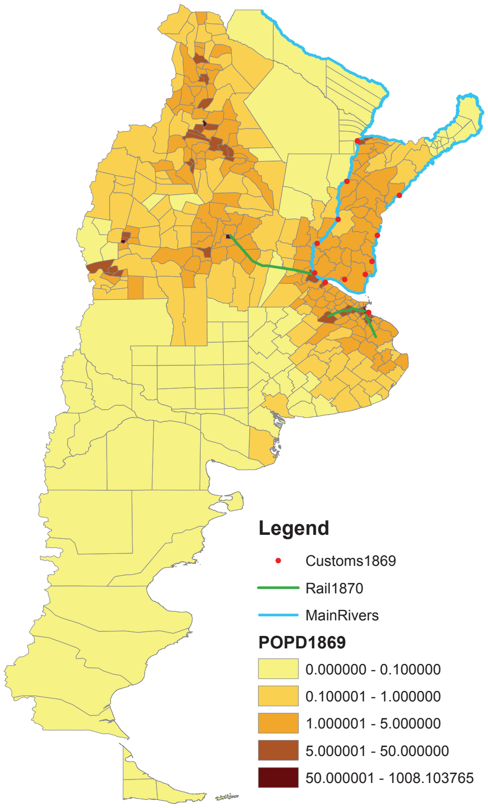


Figure 3: Population Density 1869

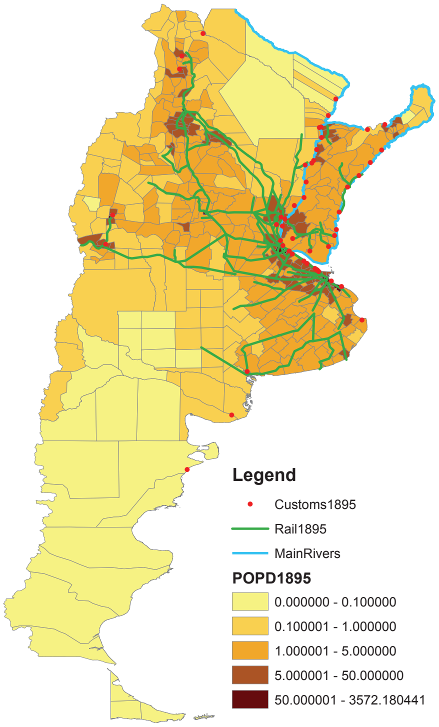


Figure 4: Population Density 1895

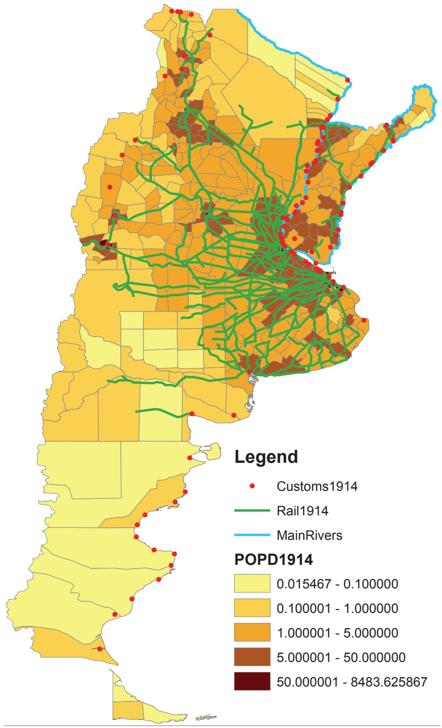
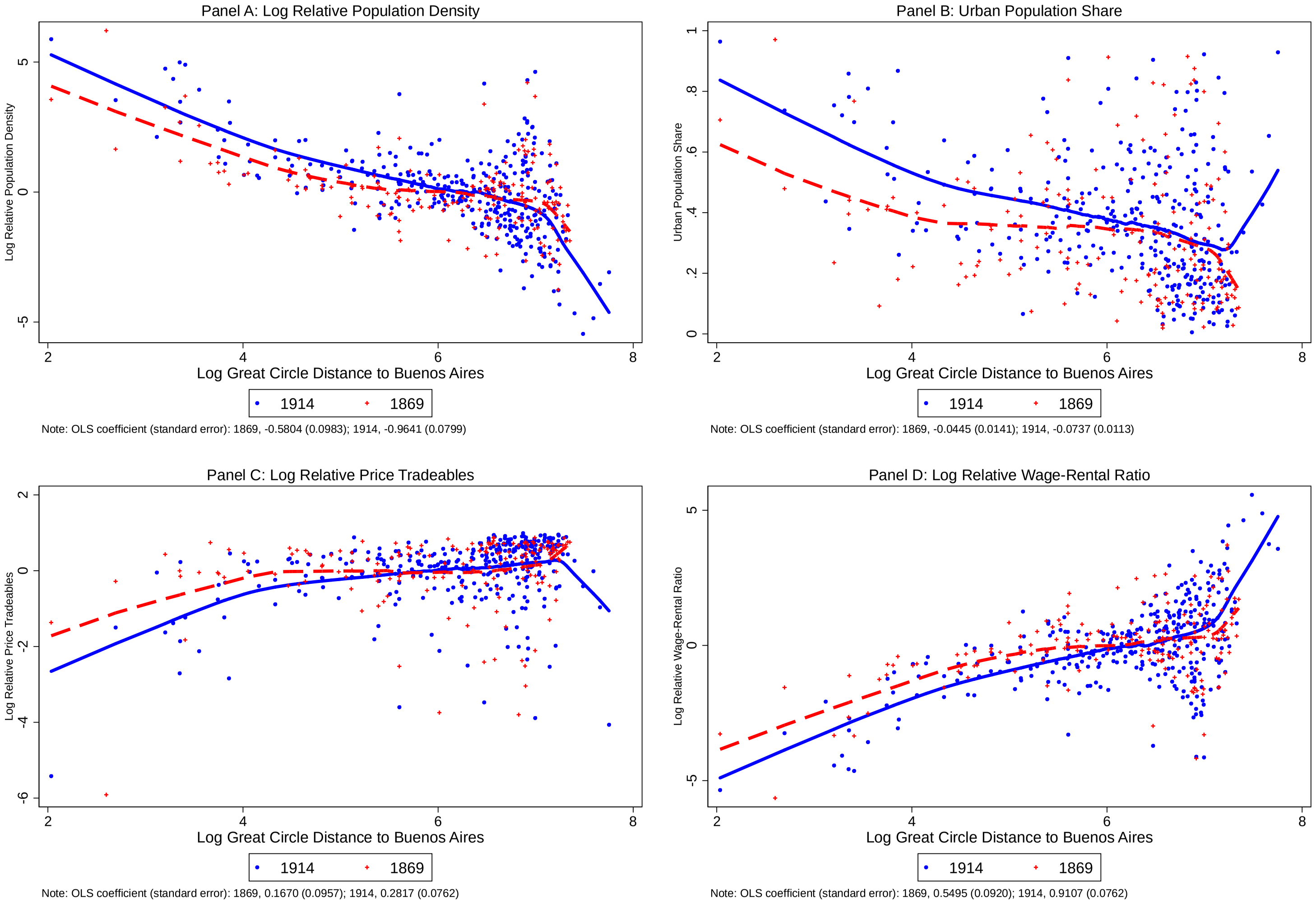


Figure 5: Population Density 1914



Panel A: Log Relative Population Density

Note: OLS coefficient (standard error): 1869, -0.5804 (0.0983); 1914, -0.9641 (0.0799)

Panel B: Urban Population Share

Note: OLS coefficient (standard error): 1869, -0.0445 (0.0141); 1914, -0.0737 (0.0113)

Panel C: Log Relative Price Tradeables

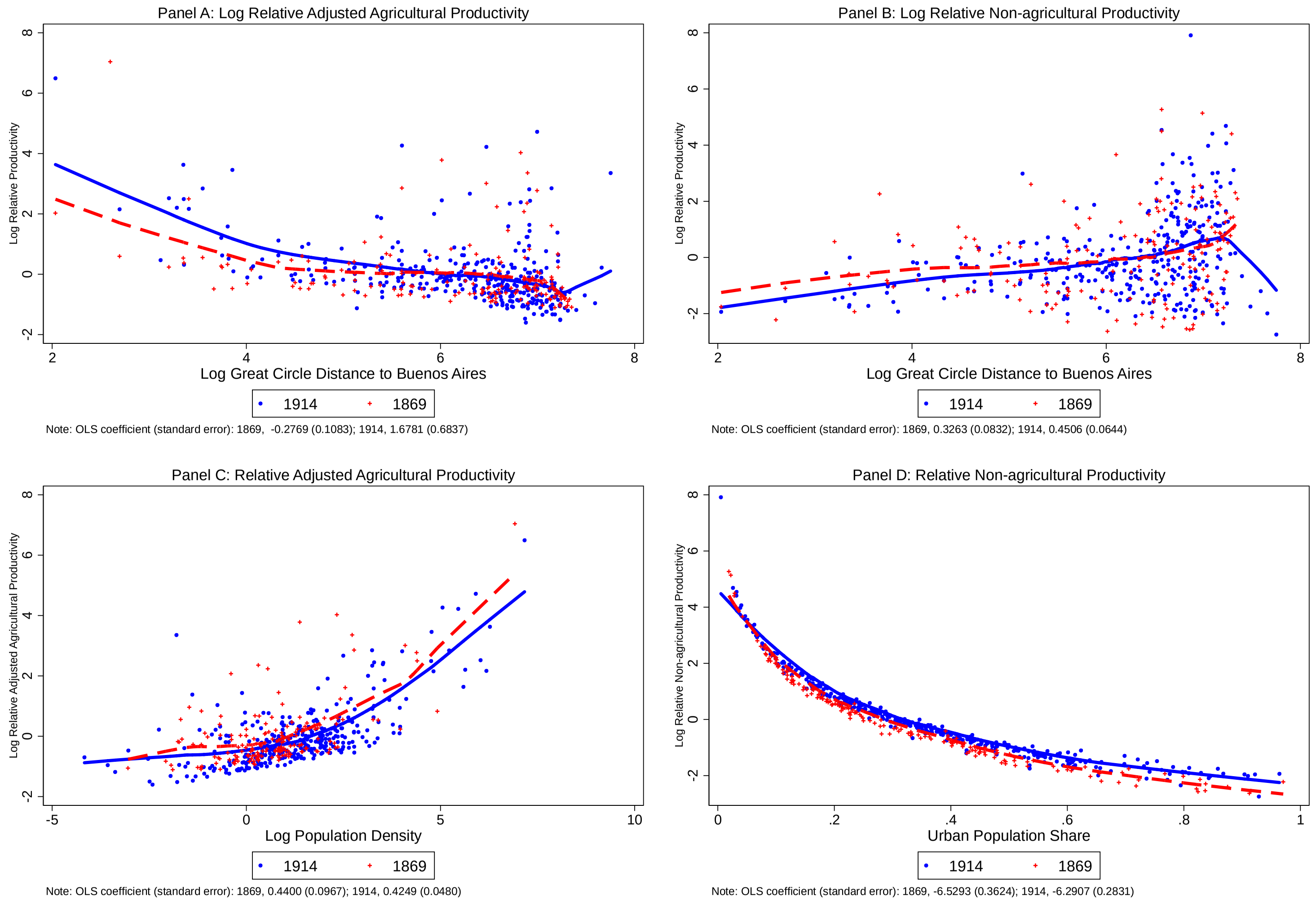
Note: OLS coefficient (standard error): 1869, 0.1670 (0.0957); 1914, 0.2817 (0.0762)

Panel D: Log Relative Wage-Rental Ratio

Note: OLS coefficient (standard error): 1869, 0.5495 (0.0920); 1914, 0.9107 (0.0762)

Note: Relative values in Panels A, C and D normalized to have a mean of zero across districts within each year.

Figure 6: Spatial Gradients in Economic Activity 1869-1914



Note: OLS coefficient (standard error): 1869, -0.2769 (0.1083); 1914, 1.6781 (0.6837)

Note: OLS coefficient (standard error): 1869, 0.3263 (0.0832); 1914, 0.4506 (0.0644)

Note: OLS coefficient (standard error): 1869, 0.4400 (0.0967); 1914, 0.4249 (0.0480)

Note: OLS coefficient (standard error): 1869, -6.5293 (0.3624); 1914, -6.2907 (0.2831)

Note: Relative values in Panels A-D normalized to have a mean of zero across districts within each year.

Figure 7: Adjusted Agricultural Productivity and Non-Agricultural Productivity 1869-1914

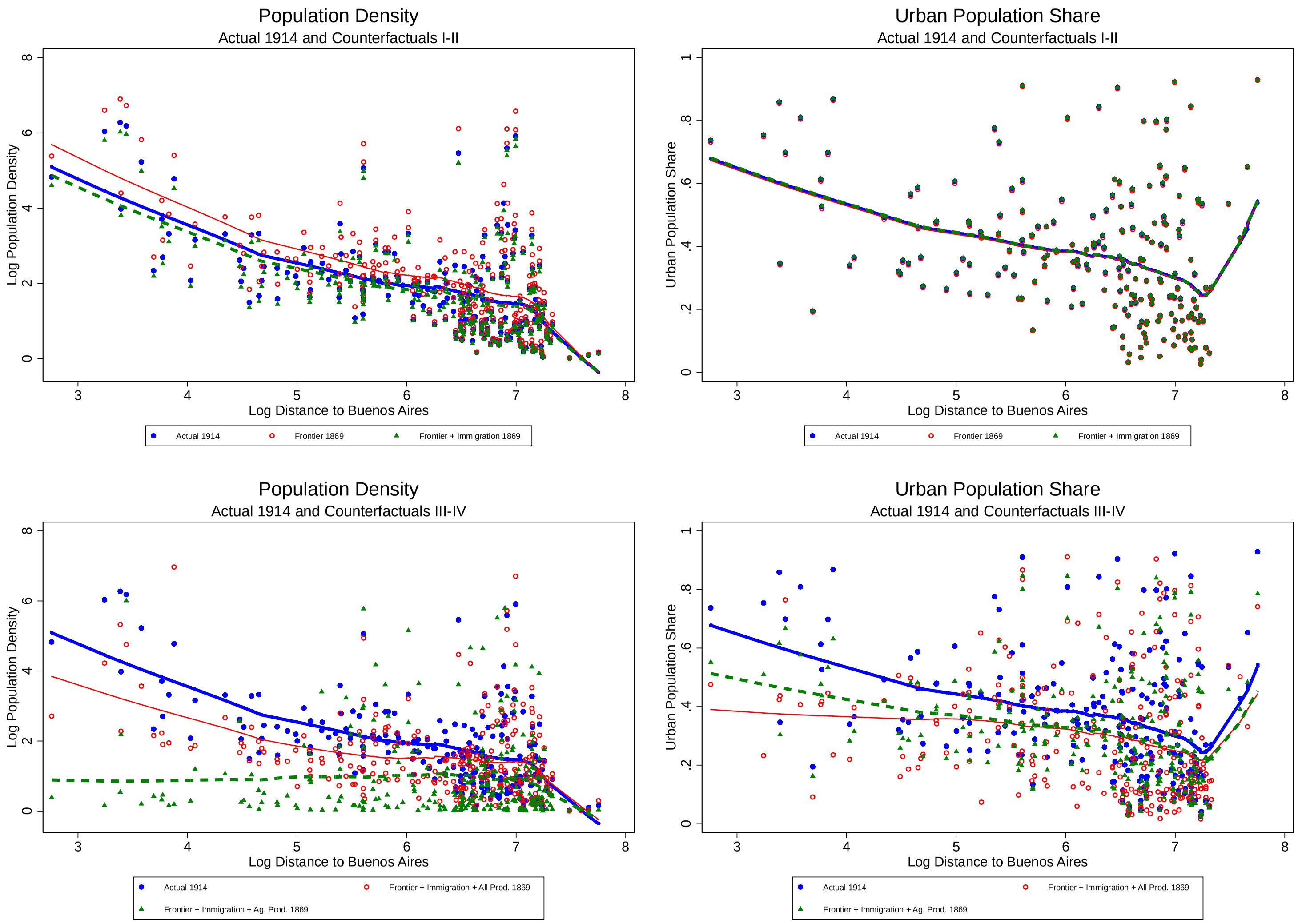
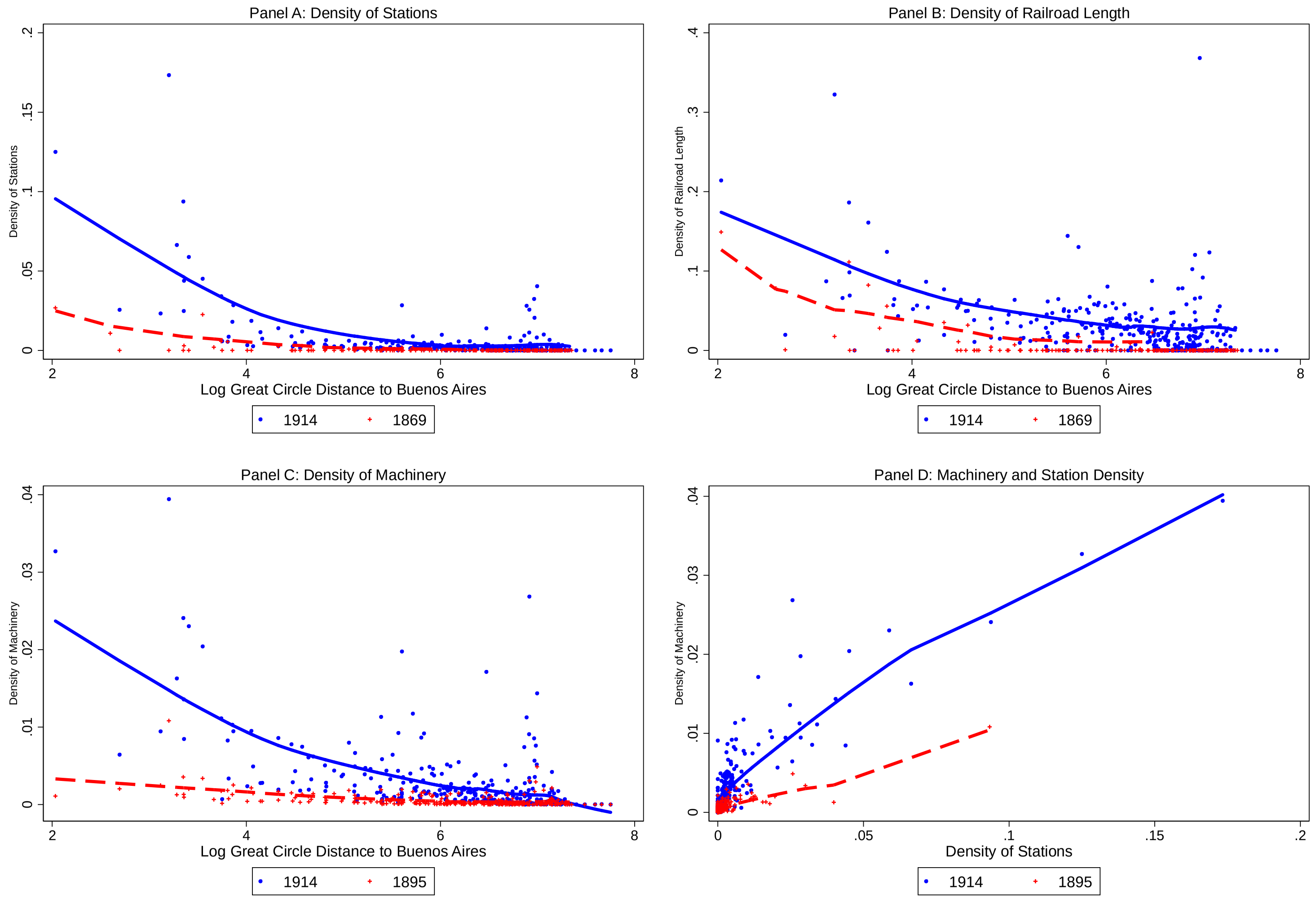
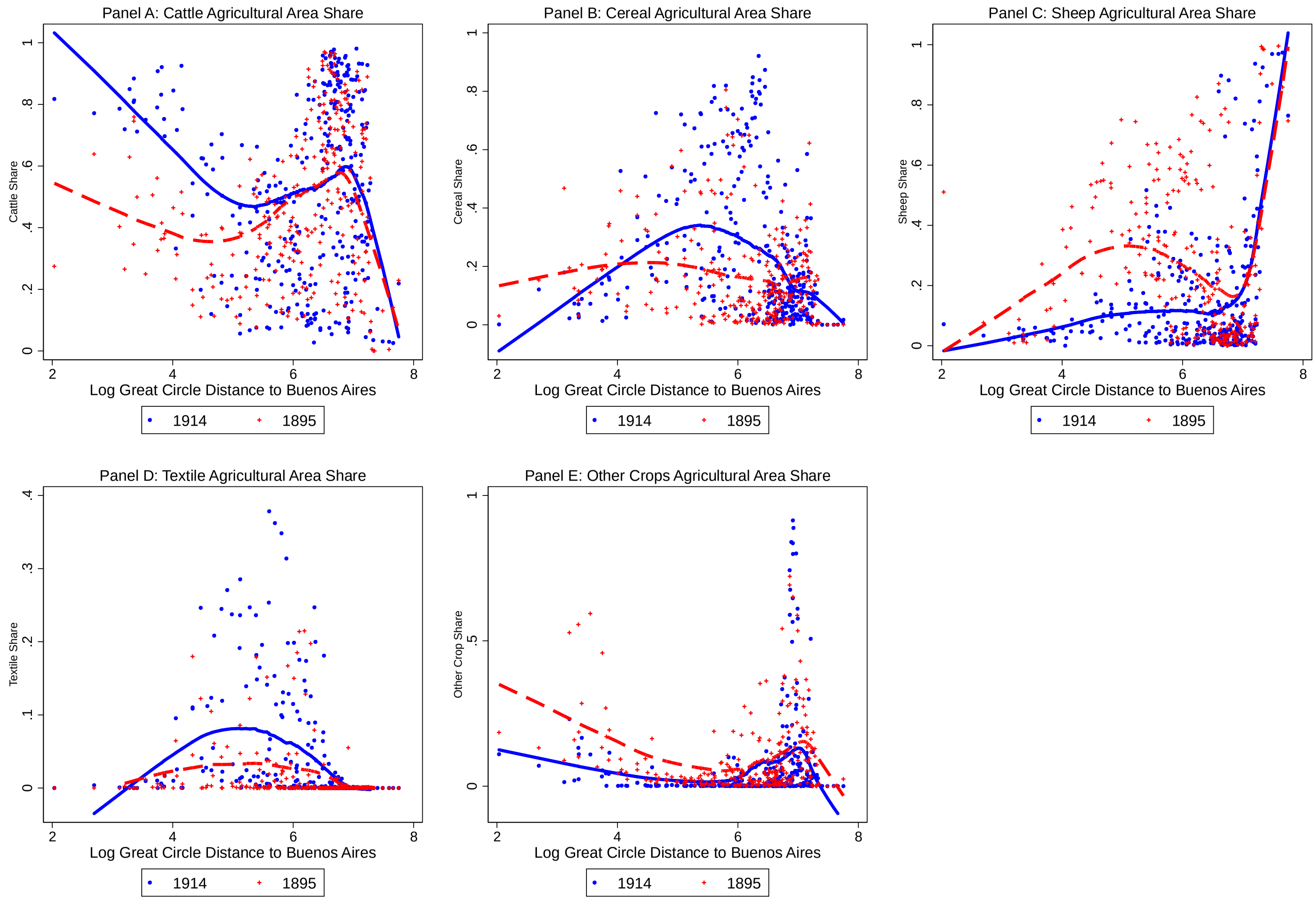


Figure 8: Counterfactuals for Sufficient Statistics $\{\widetilde{z}_A(\ell), z_N(\ell), N\}$



Note: Densities measured per kilometer squared of geographical land area.

Figure 9: Railroad and Machinery Diffusion



Panel A: Cattle Agricultural Area Share

Panel B: Cereal Agricultural Area Share

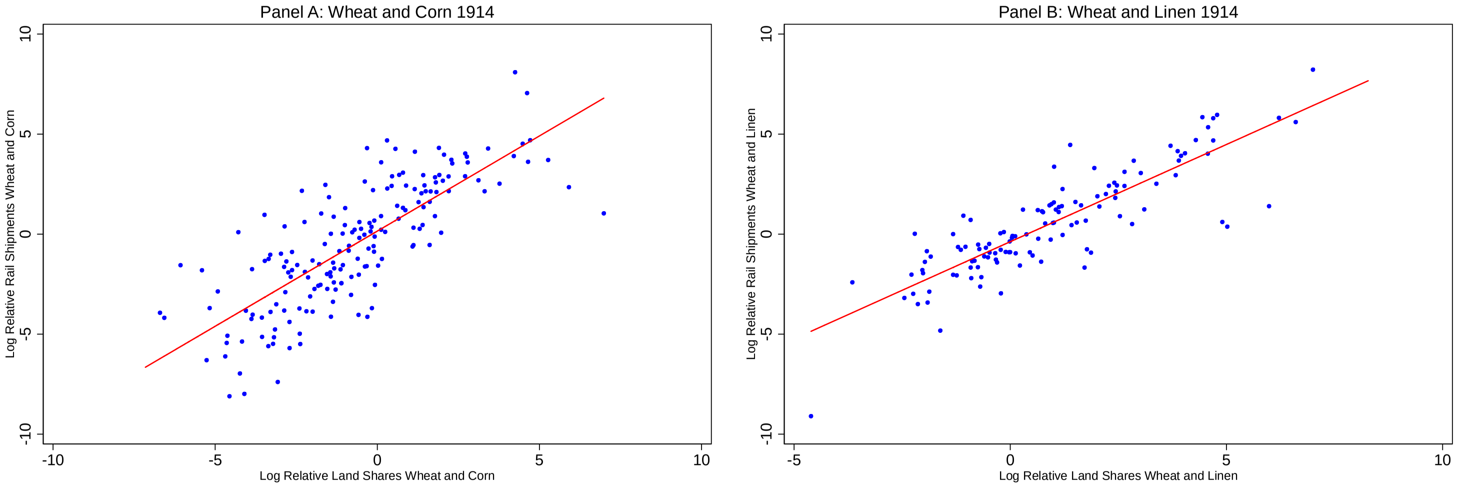
Panel C: Sheep Agricultural Area Share

Panel D: Textile Agricultural Area Share

Panel E: Other Crops Agricultural Area Share

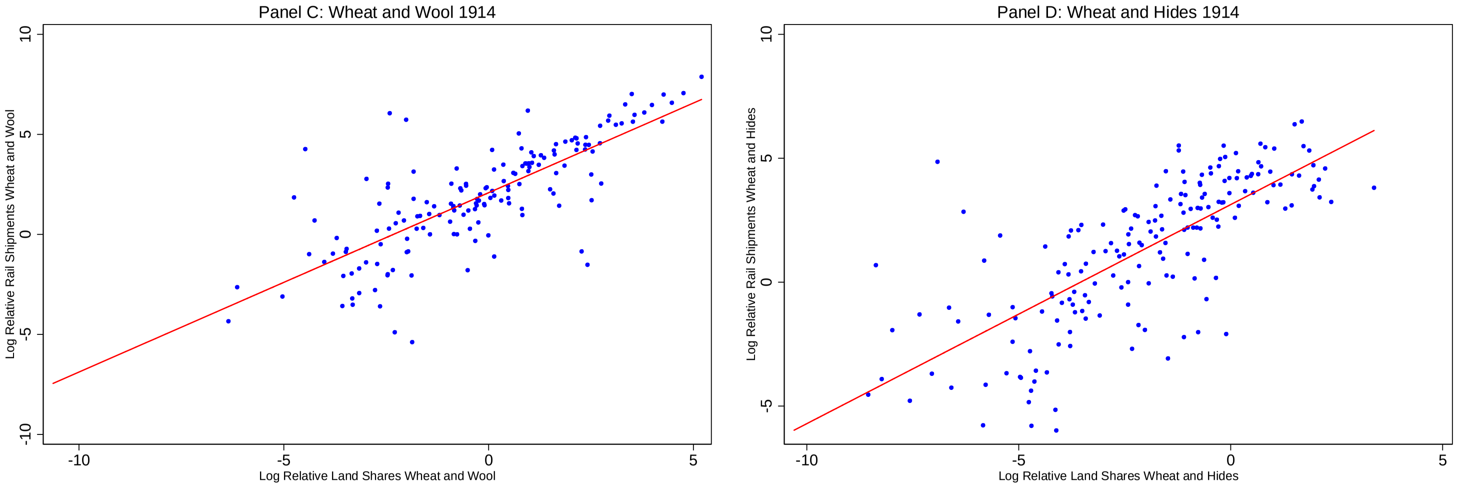
Note: Agricultural area shares sum to one.

Figure 10: Shares of Agricultural Land Area 1895-1914



Note: OLS coefficient (standard error) 0.9515 (0.0721) and R-squared 0.609.

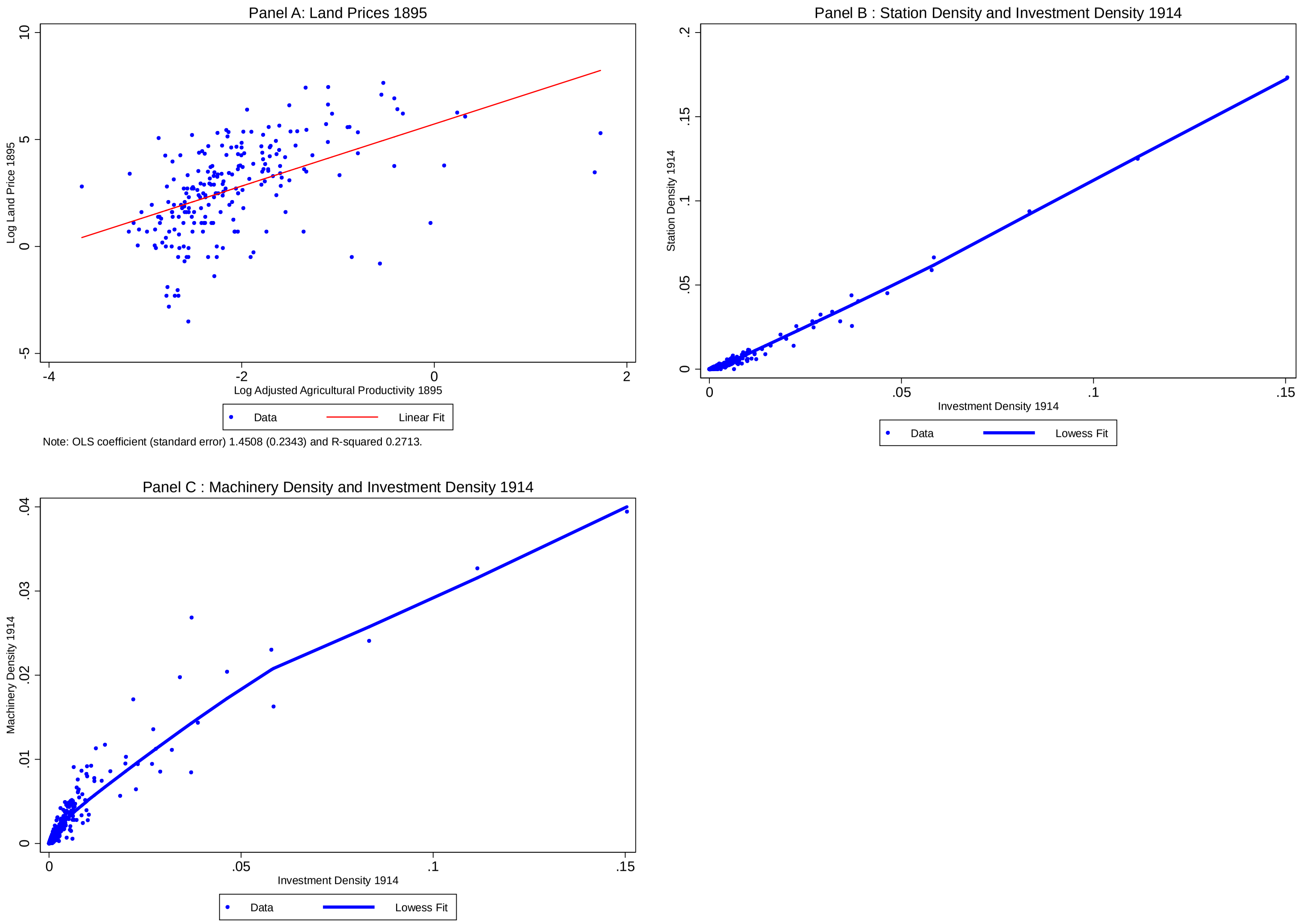
Note: OLS coefficient (standard error) 0.9720 (0.0769) and R-squared 0.733.



Note: OLS coefficient (standard error) 0.8965 (0.0647) and R-squared 0.5934.

Note: OLS coefficient (standard error) 0.8840 (0.0706) and R-squared 0.5534.

Figure 11: Relative Rail Shipments and Relative Land Shares



Note: OLS coefficient (standard error) 1.4508 (0.2343) and R-squared 0.2713.

Figure 12: Land Prices, Station Density and Machinery Density

Panel A **Sector**	 **Sector Name**	 **1869**	 **1895**	 **1910**	 **1914**
I	Livestock Products	100	62.75	43.21	43.45
II	Agriculture Products	0	34.85	52.76	52.79
III	Forest Products	0	1.82	2.84	2.65
IV	Mining Products	0	0.29	0.14	0.02
V	Products of Hunting and Fishing	0	0.15	0.38	0.38
VI	Miscellaneous Products	0	0.16	0.67	0.72
Panel B **Product**	 **Product Name**	 **1869**	 **1895**	 **1910**	 **1914**
1	Beef	4.82	0.05	6.81	10.56
2	Live Cattle	0	5.83	1.09	1.00
3	Cereals	0	33.99	52.52	50.81
4	Hides, Bones and Animal Parts	66.69	26.58	18.02	16.93
5	Mutton	0	1.40	1.61	1.34
6	Live Sheep	0	1.08	0.06	0.04
7	Other	25.30	25.84	15.79	13.53

Notes: Percentage shares of sectors and products in the value of total exports.

Table 1: Export Composition over Time

Panel A **Sector**	 **Sector Name**	 **1895**
I	Live Animals	0.64
II	Foodstuffs	12.17
III	Beverages	9.30
IV	Tobacco	1.27
V	Textiles, Yarns, Fabrics and Cordage	35.97
VI	Clothing and Apparel	3.36
VII	Chemicals and Pharmaceuticals	5.09
VIII	Wood and Wood Products	4.02
IX	Paper and Paper Products	2.36
X	Leather and Leather Products	0.68
XI	Iron and Iron Products	10.33
XII	Miscellaneous Construction Materials	2.03
XIII	Other Metals and Metal Products	1.51
XIV	Stones, Earth, Glassware and Ceramics	2.38
XV	Fuel and Lighting Supplies	6.65
XVI	Miscellaneous Manufacturing Products	2.25
Panel B **Product**	 **Product Name**	 **1895**
1	Mowers	0.23
2	Plows	0.13
3	Rakes	0.02
4	Steam Machines	0.24
5	Threshers	0.51
6	Wire	2.02

Notes: Percentage shares of sectors in the value of total imports.

Table 2: Import Composition 1895

Counterfactual (relative to base of 1914)	**Counterfactual Real Wage Relative to Base**	**Counterfactual Aggregate Urban Share Relative to Base**
I. Frontier 1869	93%	2%
II. Frontier + Immigration 1869	103%	-1%
III. Frontier + Immigration + $\{\tilde{z}_A(\ell), z_N(\ell)\}$ 1869	62%	-21%
IV. Frontier + Immigration + $\{\tilde{z}_A(\ell)\}$ 1869	92%	-9%

Notes: For each counterfactual the base is the 1914 distribution of employment across districts and sectors. In Counterfactual I, we exclude districts that were unpopulated in 1869, but hold productivities in all other districts constant at their 1914 values, and hold the total Argentine population constant at its 1914 value. In Counterfactual II, we exclude districts that were unpopulated in 1869, and adjust the total Argentine population by net immigration from 1869 to 1914, but hold productivities in all other districts constant at their 1914 values. In Counterfactual III, we exclude districts that were unpopulated in 1869, and adjust the total Argentine population by net immigration from 1869 to 1914, and set adjusted agricultural productivities and non-agricultural productivities to their 1869 values. In Counterfactual IV, we exclude districts that were unpopulated in 1869, adjust the total Argentine population by net immigration from 1869 to 1914, set adjusted agricultural productivities equal to their 1869 values, but hold non-agricultural productivities constant at their 1914 values.

Table 3: Counterfactuals for 1914 base year

	(1) $\Delta\ln\tilde{z}_A^{69-14}$	(2) $\Delta\ln\tilde{z}_A^{69-14}$	(3) $\Delta\ln\tilde{z}_A^{69-14}$	(4) $\Delta\ln\tilde{z}_A^{69-95}$	(5) $\Delta\ln\tilde{z}_A^{95-14}$	(6) $\Delta\ln\tilde{z}_A^{69-95}$
ln(distBA)	-0.3589***	-0.3953***	-0.1629**	-0.1656***	0.0196	-0.1937***
	(0.0893)	(0.0953)	(0.0797)	(0.0474)	(0.0647)	(0.0527)
$\ln\tilde{z}_A^{1869}$		-0.2167	-0.3356***	-0.3830***	0.0512	-0.3671***
		(0.1313)	(0.1126)	(0.0999)	(0.0834)	(0.1040)
Investment Density^{1914}			27.8867***			
			(9.3823)			
Investment Density^{1895}				18.7926**		
				(7.4532)		
Investment Density$^{1895\text{-}1914}$					34.7098***	3.1632
					(9.6186)	(5.1115)
Estimation	OLS	OLS	OLS	OLS	OLS	OLS
Observations	157	157	157	142	142	142
R-squared	0.1894	0.2347	0.4231	0.3761	0.4669	0.3438

Note: $\Delta\ln\tilde{z}_A^{69-14}$, $\Delta\ln\tilde{z}_A^{69-95}$ and $\Delta\ln\tilde{z}_A^{95-14}$ are the growth in adjusted agricultural productivity from 1869-1914, 1869-1895 and 1895-1914 respectively; ln(distBA) is log Great Circle distance from Buenos Aires; $\ln\tilde{z}_A^{1869}$ is log adjusted agricultural productivity in 1869; Investment Density^{1895} and Investment Density^{1914} are the first component from a principal components analysis for Station Density and Machinery Density for 1895 and 1914 respectively; Investment Density$^{1895\text{-}1914}$ is the change in Investment Density from 1895-1914; Station Density is the number of railroad stations per kilometer squared of geographical land area; Machinery Density is the number of agricultural machines (ploughs, mowers, rakes, threshers, water machines and wind machines) per kilometer squared of geographical land area; Standard errors are heteroscedasticity robust; *** denotes significance at the 1 percent level; ** denotes significance at the 5 percent level; * denotes significance at the 10 percent level.

Table 4: Internal Investments and the Spatial Gradient of Economic Activity

	(1) $\Delta\ln\tilde{z}_A^{69-14}$	(2) $\Delta\ln\tilde{z}_A^{69-14}$	(3) Investment Density	(4) $\Delta\ln\tilde{z}_A^{69-14}$	(5) Investment Density
ln(distBA)	-0.1629**	-0.0693	-0.0058***	-0.1650	-0.0057***
	(0.0797)	(0.1309)	(0.0019)	(0.1395)	(0.0019)
$\ln\tilde{z}_A^{1869}$	-0.3356***	-0.3835***	0.0033**	-0.5258***	0.0036**
	(0.1126)	(0.1157)		(0.1200)	(0.0015)
Investment Density	27.8867***	39.1298***		34.4048***	
	(9.3823)	(13.4993)		(13.2023)	
Route C16			0.0195***		0.0206
			(0.0056)		(0.0057)
Estimation	OLS	IV	OLS	IV	OLS
First-stage			yes		yes
Exclude C16 cities				yes	yes
Underid (p-value)		0.0034		0.0032	
First-stage F-test			12.04		13.12
Observations	157	157	157	141	141
R-squared	0.4231	-	0.4267	-	0.4348

Note: $\Delta\ln\tilde{z}_A^{69-14}$ is the growth in adjusted agricultural productivity from 1869-1914; ln(distBA) is log Great Circle distance from Buenos Aires; $\ln\tilde{z}_A^{1869}$ is log adjusted agricultural productivity in 1869; Investment Density is the first component from a principal components analysis for Station Density and Machinery Density; Station Density is the number of railroad stations per kilometer squared of geographical land area; Machinery Density is the number of agricultural machines (ploughs, mowers, rakes, threshers, water machines and wind machines) per kilometer squared of geographical land area; Route C16 is the fraction of a district's geographical land area that lies along a shortest route from the centroid of any district to a 16th Century Spanish colonial city; Underid is the p-value from the Kleibergen-Paap LM statistic under-identification test; First-stage F-test is the F-statistic from a test of the significance of the excluded exogenous variable in the first-stage regression; standard errors are heteroscedasticity robust; *** denotes significance at the 1 percent level; ** denotes significance at the 5 percent level; * denotes significance at the 10 percent level.

Table 5: Instrumental Variables Estimates