Free-Riding Yankees: Canada and the Panama Canal*

Sebastian Galiani[1], Luis F. Jaramillo[2], and Mateo Uribe-Castro[3]

[1]University of Maryland and NBER

[2]University of Maryland

[3]Universidad del Rosario

June 25, 2023

Abstract

We study the impact of the Panama Canal on the development of Canada's manufacturing sector in the years from 1900 to 1939. Using newly digitized county-level data from the Census of Manufactures and a market-access approach, we exploit the plausibly exogenous nature of this historical episode to study how changes in transportation costs influence the location of economic activity and productivity dynamics. Our reduced-form estimates show that lowered shipping costs led to greater market integration of marginally productive Canadian counties with key markets both inside and outside of Canada. This development permitted the reallocation of production activity to places whose production levels had been inefficiently low before the Canal opened. A shift from the 25th to the 75th percentile in terms of gains in market access brought about by the opening of the Canal led to a 9% increase in manufacturing revenues and input expenditures. Productivity rose by 13%. These effects persist when general equilibrium effects are considered: the closure of the Canal in 1939 would have resulted in economic losses equivalent to 1.86% of GDP, chiefly as a result of the restriction of the country's access to international markets. Altogether, these results suggest that the Canal substantially altered the economic geography of the Western Hemisphere in the first half of the twentieth century.

JEL: F14, F63, N22.

Keywords: Productivity, Reallocation, Market Access, Panama Canal, Canada.

*We are grateful to John Haltiwanger, Rick Hornbeck, Eunhee Lee, Nuno Limão, Noel Maurer, Chris Meissner, Martin Rotemberg, John Shea, John Wallis and participants at the NBER Summer Institute Seminar - DAE, the Annual Meetings of the EHA, the Annual Meetings of the EEA, RIDGE Historical Development Workshop, the URosario Political Economy Workshop, and the UMD Trade Workshop for their helpful comments and suggestions. Kris Inwood and Ian Keay provided valuable guidance in terms of data sources at an early stage of the paper.

1 Introduction

On March 23, 2021, the Ever Given ran aground in the Suez Canal. Although the Canal was blocked for a relatively short amount of time, the event disrupted supply chains around the world and altered maritime traffic patterns.[1] The episode served as a reminder of the obvious: the world's great man-made waterways, which have not always been there, play a central role in today's economy. Perhaps more importantly, they may have shaped the spatial distribution of economic activity. The construction of the Suez and Panama Canals led to reductions in transportation costs that dramatically widened the range of economic opportunities available to what were, then, remote places. Recent research by Brancaccio et al. (2020) indicates that closing the Panama Canal today would reduce global welfare by 3.28%. This leads us to ask: How do changes in transportation costs affect the location of economic activity and productivity dynamics?

We study how Canada's economic geography changed after the completion of the Panama Canal and estimate the market access gains obtained by different counties after the Canal began commercial operations.[2] Counties in the eastern part of Canada already had suitable transportation infrastructure in place to allow them to reach markets in the United States and Western Europe. However, counties in the western part of the country saw their trade opportunities expand significantly after 1920. We show that changes in market access had led to increases in manufacturing activity by 1939. Firms' revenues and their use of factor inputs rose as a result. Moreover, we find that manufacturing industries in counties that benefited from the opening of the Canal also realized productivity gains. We decompose those gains as coming from either increases in revenue obtained while using the same level of inputs given the existing production technology (total factor revenue productivity (TFPR)) or increases in the use of inputs in locations where they yield a different productive use (allocative efficiency (AE)). Our findings suggest that the bulk of these productivity gains were attributable to increased AE.

Canada is well suited to be used as a case study of the subnational impacts of the opening of the Panama Canal for a variety of reasons. First, shipping costs for the country fell because the Canal provided an alternative to the comparatively expensive all-North-American routes. Those routes typically required cumbersome and onerous transshipments and were not, in some cases, available throughout the year due to harsh weather

[1] See The Economist (2021).

[2] The administrative divisions or districts in Canada that are analogous to counties in the United States are referred to by different names in different Canadian provinces. The term "county" is used in this paper in a generic descriptive sense for ease of reference.

conditions in the winter. Innis (1933) documents shipping rate reductions from Liverpool, England, that ranged from 93% in the case of Vancouver to zero in the case of Regina, located to the north of eastern Montana. Second, since the Canal was not built for the express purpose of changing any outcome in Canada, the variability it generated in transport costs in a panel data model is plausibly strictly exogenous. This is useful, since a key concern that arises when studying the effects of transportation infrastructure on economic activity is whether the placement of roads, railroads and canals is influenced by the potential for economic growth. The United States built and operated the Canal to achieve strategic objectives that had little to do with its vast yet relatively underpopulated northern neighbor (McCullough, 2001).

We show that the Panama Canal impacted Canadian manufacturing through changes in market access. This paper combines newly digitized county-level data from the Canadian Census of Manufactures of 1901, 1911 and 1939; geographic information system (GIS) data for transportation networks in Canada, the United States and the rest of the world before and after the opening of the Canal; and a market access approach (Redding and Venables, 2004; Donaldson and Hornbeck, 2016; Hornbeck and Rotemberg, 2021). Market access, a concept derived from the international trade literature, measures exposure and proximity to consumers and suppliers by weighting the income-normalized population of destination markets by how costly it is to reach them. We compute market access in 1910 and 1920 for each Canadian county and use its variation to estimate how the reduction in transportation costs affected the county's manufacturing activity.

We find that the Canal had substantial effects on several manufacturing outcomes. Moving a county from the 25th percentile of market access gains in 1920 to the 75th percentile would increase manufacturing revenues by 8.9%. Capital expenditures would go up by 7%, labor expenditures by 9.7%, employment by 9.1%, material expenditures by 8.3% and measured productivity by 13.1%. Though the results are noisier when we decompose productivity into TFPR and AE, we show that increases in AE– which measures the extent to which inputs are used in locations where they yield their highest productive use due to imperfections in the economy– seem to be relatively more important to explain changes in measured productivity.

One concern is that changes in market access induced by the construction of the Canal could be correlated with secular changes in outcomes that predated its opening. In other words, the concern is that trends in manufacturing outcomes seen before the opening of the Canal would have given rise to the same outcomes, at least partially, even in the absence of the Canal. To rule out this statistical nuisance, we show that changes in market access attributable to the opening of the Canal are unable to predict changes in manufacturing activity that occurred between 1900 and 1910.

We interpret our findings as being driven by the changes in transport costs brought about by the presence of the Canal. An important challenge to this interpretation would be that the expectation of the Canal's opening could have led to increases in railroad construction that benefited counties that would eventually achieve greater market access thanks to the building of the Canal. However, we show that counties whose market access would eventually increase because of the Canal did not benefit from railway construction undertaken before the Canal started operating. Furthermore, all estimates are robust to the inclusion of local railroad construction *after* the Canal began commercial operations.

Our estimates contrast outcomes for different counties in terms of the extent of their gain in market access (comparing those that gained relatively *more* market access with those that gained *less* access). A natural, though different, question that follows from this had to do with the *aggregate* impact of the Canal. A naïve exercise could use the input elasticities that follow from our reduced-form exercise to provide an estimation of this effect. This approach would be unsatisfactory, however, because our findings speak about *relative* changes, which are not predictive or informative about new input *levels*. These levels are likely to be affected by general equilibrium effects. As Redding (2021) points out, differences-in-differences estimations do not distinguish between the displacement and the creation of economic activity and, therefore, cannot be used for welfare calculations.

In the context we study, a key concern is that the effects we identify are simply the result of a displacement of economic activity toward counties that benefited more from the Canal. In other words, the concern is that the Canal triggered shifts of production activities within Canada but that, in the aggregate, the benefits derived from the Canal were very limited. To address this concern, we calibrate a benchmark general equilibrium model of economic geography with frictions (Eaton and Kortum, 2002; Donaldson and Hornbeck, 2016; Hornbeck and Rotemberg, 2021) to assess a counterfactual scenario in which the Canal closes permanently in 1939. In calculating the impact of closing the Canal, we allow for the population within North America to be fully mobile. We assume that the population adjusts freely, fixing workers' welfare at its pre-closure level. We then use the model's results on counterfactual prices, wages, rents and population to compute the impact of closing the Canal on manufacturing productivity and property values.

Though not the focus of our analysis, we use the results from our general equilibrium exercise to compute a measure of the impact of closing the Canal on the agriculture, an important sector of the Canadian economy. We do so to provide a more comprehensive, back-of-the-envelope measure of the economy-wide effects of closing the Canal in 1939. We follow Donaldson and Hornbeck (2016) and take the counterfactual prediction on rents from the model to compute the total loss due to decreased land and property values in agriculture.

We find that the closure of the Panama Canal in 1939 would have led to non-negligible population and economic losses for Canada. The total Canadian population would have shrunk by 2.7%. Back-of-the-envelope calculations point to economic losses of 1.86% of GDP, reflecting lower productivity in manufacturing (0.27% of GDP) and decreased agricultural land values (1.59% of GDP), which would have followed from lower land rental rates. Although most counties would have lost population, those closer to the coasts would have experienced the largest declines. These decreases would have been more acute in the western provinces, as British Columbia and Alberta would have seen their populations fall by 7.8% and 4.2%, respectively. These declines stand in contrast to population gains of 0.07% in Manitoba and 0.2% in Saskatchewan, both central provinces that stood to gain much less – if anything – from the Canal. Our general equilibrium exercise suggests that, overall, the Panama Canal has facilitated higher levels of economic activity in Canada in the aggregate.

To shed light on the origin of the gains brought about by the Canal, we carry out a second counterfactual exercise in which we close the Canal for shipments originating in or destined for countries other than Canada. In other words, we allow the use of the Canal only for domestic trade within Canada. Along the lines of Fajbelgaum and Redding's study of the Argentinian case, this exercise sheds light on the relative importance of domestic and international market access gains in the shock we study.

Using the 1939 factual equilibrium as a base, total population and GDP both fall, by 2.4% and 1.74%, respectively. These losses reflect lower productivity in manufacturing (0.25% of GDP) and lower land values in agriculture (1.49% of GDP). Given the consequences of completely closing the Canal described in the previous paragraph, our results suggest that the bulk of the gains brought about by the Canal for Canadian counties were derived from their greater exposure to markets outside Canada. These findings are consistent with historical accounts that emphasize the importance of the Canal for trade in Canadian staples such as lumber and wheat because of the greater access it afforded to markets in the United States and Western Europe (Mackintosh, 1939).

This paper relates to several strands in the economic literature. First, there is a large body of literature on reallocation that documents imperfections – such as regulations or mark-ups – in the economy (Hsieh and Klenow, 2009; Petrin and Levinsohn, 2012; Restuccia and Rogerson, 2008). These imperfections prevent the equalization of marginal revenue products, dampen productivity and induce the misallocation of production factors. Notably, Baqaee and Farhi (2020) demonstrate that the Petrin and Levinsohn decomposition employed in this study underestimates the contribution of allocative efficiency (AE) to productivity growth. To potentially enhance our decomposition, access to firm-level data and input-output network structure would be necessary. However, since such data is unavailable for our specific context, we simply acknowledge that our findings provide suggestive evidence that AE primarily drives the productivity changes resulting from the opening of the Panama Canal. If the Petrin and Levinsohn decomposition also underestimates the contribution of AE to productivity growth in our specific setting, following Baqee and Farhi's example with modern US data, our estimation would represent a lower bound estimate of AE's role.

More recent work assesses the impact of transportation infrastructure on reallocation dynamics in the context of domestic transportation infrastructure in India, the United States and Mexico (Asturias et al., 2019; Asher and Novosad, 2020; Hornbeck and Rotemberg, 2021; Zárate, 2020). We explore a strong, one-time shock to transport costs rather than to the gradual development of a transportation network over time, which might be subject to strategic local considerations about the location of investment in infrastructure. In contrast to the existing literature, the intervention we exploit is not located within the domestic infrastructure network of the country we study. This feature alleviates reverse causality or targeting concerns, which are prevalent in this body of work. Further, we use the Canal shock to answer the question as to whether improved transportation technologies can increase productivity by enabling the reallocation of production factors.

Second, a growing body of work on international trade has emphasized the relevance of geography and trade costs in accounting for differences in economic outcomes across and within countries (Krugman, 1991; Redding and Venables, 2004). Evidence from a variety of contexts and from cross-country data indicates that market access is positively correlated with development outcomes (e.g. Liu and Meissner (2015); Cao and Chen (2022); Martinez-Galarraga et al. (2015); Missiaia (2016); Jacks and Novy (2018)). Other recent work has emphasized transport costs and infrastructure as important determinants of market integration (Banerjee et al., 2020; Donaldson, 2018; Donaldson and Hornbeck, 2016; Fajgelbaum and Redding, 2014; Hornbeck and Rotemberg, 2021; Limão and Venables, 2001; Pascali, 2017; Sotelo, 2020). Our main contribution here is to document a large shock that transformed the economic geography of the Western Hemisphere and to use it to study the impact of transport infrastructure on manufacturing outcomes in an economy marked by imperfections.

Third, we build on previous work on the impact of the Panama Canal on North American outcomes. This literature emphasizes that the Canal was a source of pecuniary externalities, social savings and structural change for the United States (Rockwell, 1971; Maurer and Yu, 2008; Maurer and Rauch, 2020) and a potential determinant of land and wage values in Canada (Umaña-Dajud, 2017). In particular, Maurer and Rauch (2020) develop a similar approach to study the effect of the Panama Canal's opening on US counties' populations, employment levels and manufacturing wages. We build on their findings by focusing on the Canadian manufacturing sector and show not only that wages and employment increased for counties that benefited from the Canal, but also that those counties increased their manufacturing revenues and productivity. Moreover, we find larger welfare gains from the Canal by using a model that considers differences in productivity between counties. Our main contribution to this literature is to assess the effects of the Canal on productivity in the context of imperfections in the economy. As we show, acknowledging them and recognizing the role that infrastructure can have in bypassing them can significantly alter our notion of the benefits of these projects and the role we assign them in the process of economic change.

Finally, this study contributes to the literature on Canada's economic history. Although contemporary observers were quick to point to the Canal as a major disruptor of economic life in Canada (Innis, 1933, 2018; Mackintosh, 1939), ours is the first project to systematically assess the impact of the improvement in transportation technology embodied by the Canal on manufacturing outcomes in that country. It also relates to work that emphasizes integration into the global economy as a force that helped to shape Canada's industrialization patterns (Alexander and Keay, 2019; Jaworski and Keay, 2021). Canadian historiography has typically overlooked this factor and emphasized scale economies, industrial policy and domestic market expansion to explain the development of this sector of the economy (Keay, 2007).

The rest of this study is structured as follows: Section 2 provides historical context; Section 3 presents a detailed discussion of the construction of the data; Section 4 describes the empirical analysis that was undertaken; Section 4.2 covers the implementation of robustness checks; Section 5 provides an assessment of the costs of closing the Canal in 1939; and Section 6 concludes.

2 Historical Background: Canada and the Panama Canal

Transportation has been a central factor in shaping Canada's development and its relationship with the rest of the world (Innis, 2018). With the fall in haulage costs that followed the introduction of the steamship and the expansion of the railroads in the nineteenth century, Canada's main staple products transitioned from cod and furs to wheat and lumber. These were mostly produced in the country's eastern and maritime provinces.

However, the effective access that some parts of the country had to key domestic and international markets remained limited even after the completion of the transcontinental Canadian Pacific Railway in 1886 (Innis, 2018). The vast distances involved, the railroad's market power and expensive transshipments remained obstacles to cheap haulage. Before the opening of the Panama Canal, a typical producer located west of Winnipeg that was eager to sell its products in New York had four potential options for shipping products to that market: (i) the Canadian transportation network, which entailed a combined journey by railroad and steamship spanning over two thousand miles over a route that was partly impassable during the winter; (ii) the United States' railroad network up to an American port that did not freeze in the winter (as did those on the St. Lawrence River); (iii) the Panama and Tehuantepec railroads; and (iv) Cape Horn, a route that was over 16,000 miles in length linking the cities of Vancouver and New York.

The costliness of these routes hindered economic agents from trading with distant markets. Lawrence (1957) points out that the shipment of lumber via the Canadian transportation network was so cumbersome that, in the case of some products, it was slightly cheaper to ship them from Vancouver to the Atlantic coast of North America via Cape Horn. Even then, these routes were too costly to be used for most products. For Huebner (1915), the very limited capacity of the Isthmian railroads and high transshipment costs cancelled out most of the savings obtained by using shorter routes than the Cape Horn passage. Given these conditions, Canada was set to tangibly benefit from the opening of a waterway in Central America.

The idea of building a canal in Central America dates back to at least the sixteenth century. Advisors to the King of Castille pointed out how beneficial such a waterway would be for the Spanish crown's profitable trade with Asia. After having surmounted what had traditionally been regarded as the impossible task of building the Suez Canal, the French developers of that route undertook the first serious effort to construct a canal through the Isthmus of Panama. Construction began in 1881, but that attempt eventually failed due to a series of conceptual flaws and challenging conditions that prevented the project from making material progress. The *Compagnie Universelle du Canal Interoceánique de Panama* filed for bankruptcy in 1889, and the works remained essentially abandoned until the next century when the United States became interested in the project.[3]

In his 1901 State of the Union address, President Theodore Roosevelt asserted that no single material work yet to be undertaken was of such consequence to the American people as an Isthmian canal. One of the strategic considerations behind this statement had to do with the expectation that such a project would eliminate the need for the United States to establish two distinct naval fleets to defend its Pacific and Atlantic shores. As part of an explicit policy to limit the influence of European powers in the Western Hemisphere, this was a powerful reason for renewed interest in the construction of a canal. After the Colombian Congress rejected a treaty that would have allowed the United States to build and manage such a canal in 1903, the Roosevelt Administration supported a revolution that would end in the secession of Panama from Colombia, after which Panama allowed for the construction of a canal under terms similar to those rejected by Bogotá. At a cost of approximately $10 billion in current terms, the Panama Canal opened for traffic on August 15, 1914, just a few weeks after the Great War erupted in Europe.

Even after it opened, however, the Panama Canal remained underutilized and effectively closed to commercial passage until 1920 (Maurer and Yu, 2010). First, the outbreak of war in Europe depressed maritime shipping markets and increased rates to prohibitive levels that would not recede until after the end of the war. Second, a series of landslides closed the waterway for several months in 1915, 1916, 1917 and 1920. Third, widespread labor strikes prevented its full operation in 1916 and 1917. As a result, the Canal did not effectively open for commercial traffic until 1920. As per the Panama Canal Act, tolls were set at levels designed only to cover operation and maintenance costs, not to maximize the revenue of the company running the waterway.

Contemporary observers pointed to the Canal as a potential factor behind the expansion of the economies of western North America during the 1920s. Mackintosh (1939) claims that: "It was not until the opening of the Panama Canal that British Columbia experienced the rapid development which comes from increasing access to world markets and

[3] Another company, the *Compagnie Nouvelle du Canal de Panama* was established in 1894 but also failed to make material progress towards completion.

a great extension of the hinterland tributary to its metropolitan centre." The share of domestic manufacturing production located in the provinces of Alberta and British Columbia increased from 7.2% in 1910 to 9.7% in 1939, a period during which domestic manufacturing production expanded substantially across the country. In a similar vein, in an article entitled "The Boom in California", The Economist (1924) asserted that: "Undoubtedly the chief factor in the expansion has been the opening of the Panama Canal, which has given the varied products of Southern California cheapened access to outside markets." Systematic tests of these hypotheses as applied to the case of Canada are presented in Section 4 of this paper.

3 Data

This study uses a combination of newly digitized census data, geographic information system (GIS) data and key parameter estimates from the recent trade and economic history literature. Our main sample consists of 217 counties in Canada that reported manufacturing activity in 1900, 1910 and 1939. The sample encompasses the universe of counties over the course of the study except those in the Yukon and the Northwest Territories. This sample of counties includes 99.6% of the total population in 1911 and 99.9% of the population over subsequent census years (1921, 1931, and 1941). We first measure how the effective opening of the Panama Canal in 1920 changed each county's exposure to other markets in North America and the rest of the world. Then, we estimate how those changes in market access led to changes in manufacturing firms' revenue, use of inputs and productivity.

To measure how much counties benefit from the Panama Canal, we compute each county's market access –a variable that we discuss in detail below– with and without the Canal in 1920. We focus our Market Access measure construction in this year because this is when the Canal effectively started operations. However, it could be the case that around the time it became clear the Canal works were going to be successful, expectations started to shift and changes in population and infrastructure happened before 1920. To deal with this concern we present several robustness checks that are described in more detail in Section 4.2. First, we measure Market Access with 1910 population figures. Second, we show railroad construction before 1920 is not correlated with Market Access gains from the Canal. Third, we show our results are robust to controlling for railroad network growth before and after the Canal.

In this section, we describe how we construct market access and how we measure manufacturing outcomes over time.

3.1 Market Access

To assess the impact of the Panama Canal on manufacturing outcomes, we focus on changes in counties' market access. The concept of market access, as developed in the economic geography literature, captures the effective exposure of agents in any given location to suppliers and consumers elsewhere. We calculate county c's market access as:

$$MA_c = \sum_{d \neq c} \tau_{cd}^{-\theta} L_d Y_d \tag{1}$$

Where τ_{cd} is the iceberg trade cost between county c and destination d, L_d is destination's population, Y_d is the GDP per capita of the country where d is located relative to Canada's GDP per capita, and θ is the elasticity of trade to trade costs. The trade cost τ takes an iceberg cost form and is computed as:

$$\tau_{cd} = 1 + \frac{t_{cd}}{\bar{P}} \tag{2}$$

here t_{cd} is the cost of moving one ton of products from county c to destination d and $\bar{P}$ is the average transportation cost per ton. The estimation of each county's market access requires, therefore, a definition of the possible set of destinations D and estimations of t_{cd}, $\bar{P}$, and θ.

Destinations

The set of destinations D to which we assume each Canadian county has access comprises all counties in Canada, all counties in the United States and selected countries in the rest of the world. For the last group, we are constrained to use a subset of 63 countries and territories for which there is population and GDP per capita data available from the Maddison project for the years from 1910 to 1920. In the case of countries that have ocean access, their population is assigned to their most historically relevant port. The population of landlocked countries is assigned to the closest international port, as measured by the distance from its borders. In total, 56 ports are used as destinations to compute transportation costs. Therefore, each Canadian county has 3,069 destinations: 216 of its peers in Canada, 2,797 in the United States and 56 ports in the rest of the world.[4] Altogether, the

[4] The transport network data for the US uses 1890 county boundaries. To use population figures for 1910 and 1920, we match counties in 1910 and 1920 to 1890, assuming that the spatial distribution of the

destinations in our sample account for approximately 86% of the global population and 93% of the global GDP in 1920.[5]

To estimate transportation costs between each origin (Canadian counties) and destination (set D), t_{cd}, we need two elements. The first is a transportation network that captures the relevant options open to Canadian producers that were shipping goods in the first decades of the twentieth century. Those options involved both the United States' and Canada's systems of railroads, canals and waterways, wagon routes and ocean liners – and, eventually, the Panama Canal. The second is an assumption about the rates charged for each mode in each country, as well as transshipment costs across different modes whenever possible. As we note below, we build such a network and estimate costs for each mode in each country using rates for wheat – a key Canadian staple with relatively simple transportation features – as inputs for computing the transportation cost estimates (t_{cd}).

Transportation Network

For Canada, our transportation network utilizes the University of Toronto's GEORIA project shapefiles, which provide georeferenced information on Canadian railroads and stations from the nineteenth century onward, including data on the year in which each line opened. We built most of the remaining components by using historical sources and then manually georeferencing them: we identified canals, waterways and harbors and drew them based upon information from traffic and other available facilities found in the *Summary of Canal Statistics* (Dominion Bureau of Statistics, 1940) and *Directory of Ports and Harbours of Canada* (Department of Marine and Fisheries, 1922) . We incorporated wagon routes by drawing straight lines between each county's geographical centroid and other county centroids – both in Canada and in the United States – in a 400km radius. A similar procedure for railroad stations and harbors in a 200km radius was used.[6] This network is available for 1910, 1920 and 1939, with railroads and the Panama Canal lines being the only time-varying components.

population is homogeneous (Eckert et al., 2020). For Canada, we use 1,941 boundaries from that year's Census of Population, which were the same as those used for the 1939 Census of Manufactures.

[5] We detail the list of countries we cover and the corresponding port to which they are assigned in the Online Appendix. In practice, the only systematically excluded area is Sub-Saharan Africa. Canadian trade with Africa, some of which is covered in our sample, represented between 0.2 and 0.8% of imports and between and 1 and 2% of exports over the 1930s.

[6] Straight-line wagon connections passing over the Great Lakes and the St. Lawrence River basin were not included.

For the United States, we rely entirely on the transportation network provided by Donaldson and Hornbeck (2016), which is based on previous work by Atack et al. (2010). Like the Canadian network, it includes railroads, canals, waterways, straight-line wagon routes and a linear ocean route that goes through Cape Horn. We supplement the domestic networks described above with the Panama Canal as an additional alternative to transport goods across the Americas. We use The Panama Canal Records, the official United States government gazette, to identify American and Canadian ports listed as origins or destinations of shipments passing through the Panama Canal from 1914 to 1939. We find 34 ports, 30 in the United States and 4 in Canada. This information is useful in constructing shipping routes between ports in the Pacific (12, 2 of which are in Canada) and the Atlantic (22, 2 of which are in Canada) passing through the Canal.[7]

Finally, we allow for the shipment of goods to destinations outside North America by creating ocean routes between the 34 North American ports and the 56 international ports around the globe that we had previously identified. We first do so by using information on actual distances between ports and key global chokepoints from the United States Navy (1911, 1917, 1920, 1931, 1943) while allowing for either direct routes between ports – whenever possible – or routes passing through global chokepoints *other than* the Panama Canal.[8] We then allow for the Canal to be used and incorporate routes that use it as part of the transportation network.

Rates

We construct estimates for rates for each mode of transportation used in the network separately for the United States and for Canada. We use wheat as the product of reference to compute our estimates because it is a staple product that has been of central importance for Canada's economy and has relatively simple transportation requirements; moreover, it is a product for which a wealth of historical data exist in both countries. We discuss our sources and compare our estimates to previous ones below. The specifics of each calculation are discussed in the Online Appendix B. All rates reported in this section are in 1910 Canadian dollars, as the American and Canadian currencies traded at par in 1910.

In the case of Canada, we rely on a variety of historical sources. For the railroads, we use historical data on rates and distances provided by the Canadian Railway Commission (1939) and compute an average rate of 0.514 cents per ton-mile. For non-oceanic waterway

[7] See the Online Appendix B.

[8] These chokepoints are the Suez Canal, Cape Horn, Cape of Good Hope, Singapore, the Strait of Gibraltar and Bishop Rock.

transportation, we use information on rates, distances, storage costs and insurance premiums provided by the Saskatchewan Grain Commission (1914) and the House of Commons of Canada (1907). We compute an average rate of 0.238 cents per ton-mile. This rate includes additional charges for storage whenever the waterways were frozen, insurance and fees that were either not applicable (winter storage) or already included in the railroad rates. For wagon transportation and in the absence of specific figures for Canada, we use data provided by the United States Department of Agriculture (1906) for states that shared a border with Canada, which is a reasonable measure of wagon transportation costs for Canada according to the Saskatchewan Grain Commission. We compute an average rate of 25.657 cents per ton-mile. We force the payment for any shipment that switches modes of a flat 50 cent rate, as per information provided by the Saskatchewan Grain Commission. All rates are in 1910 Canadian dollars.

For the United States, we use different historical sources. For rail and non-oceanic waterway transportation, we use data on average rates provided by the Interstate Commerce Commission (1913). We retrieve an average rate of 0.626 cents per ton-mile for railways and a rate of 0.260 cents per ton-mile for waterway transportation. The latter includes charges for storage and insurance following the exercise for Canada. For wagon transportation, we use data provided by the United States Department of Agriculture (1906) and retrieve an average rate of 22.639 cents per ton-mile. We require any shipment switching modes to pay a flat fee of 50 cents (Fogel, 1964).

For maritime routes, we rely on the weekly transportation rates provided by the Georgian Bay Canal Commission (1916) for shipments from Liverpool, United Kingdom, to New York City, Odessa, Karachi and Buenos Aires. We estimate an average rate of 0.052 cents per ton-mile, inclusive of insurance charges, as provided by the Saskatchewan Grain Commission (1914). According to historical records, shippers using routes through the Suez Canal and the Panama Canal paid, respectively, a flat toll fee of $1.48 and $0.95 per ton (The Panama Canal Company, 1971). All rates are in 1910 Canadian dollars.

The rates estimated above are similar for both countries, which is consistent with accounts by Innis (2018) and the House of Commons (1907) that emphasize the co- determination of transportation rates in both countries. With the exception of the rates for waterways, all of them are close to those given in Fogel (1964) for the late nineteenth century in the United States: railroads (0.630 cents per ton-mile), wagons (23.1 cents per ton-mile) and transshipments (50 cents per ton-mile). Although the rate for waterways is similar to Fogel's, our final all-inclusive estimate for non-oceanic waterway transportation (0.26

cents per ton-mile) is lower than his (0.49 cents per ton-mile) because we do not attribute the capital costs of storing wheat inventories during the winter to the transportation mode itself. Doing so might be justified for a social savings exercise but is harder to justify for the purposes of this paper. Our estimate of maritime transportation (0.052 cents per ton-mile) is within the boundaries of the literature. It is higher than Maurer's and Yu's (2008) estimate of 0.036 cents per ton-mile of variable costs and slightly lower than Donaldson's and Hornbeck's implied rate of 0.055 cents per ton-mile for a Cape Horn route.

Computation of Iceberg Trade Costs

Given the transportation network described above and the estimated rates, we implement Dijkstra's algorithm (Dijkstra, 1959) to compute the transportation costs, t_{cd}, between each Canadian county and each destination. The algorithm is not instructed to impose any penalty for using or switching between the American and Canadian networks besides the transshipment costs, which apply to both networks.

Following equation 1, these costs are divided by $\bar{P}$ in order to convert them to iceberg trade costs for use in the computation of each county's market access. We compute $\bar{P}$ following equation 3. $\bar{P}_c$, computed following equation 4, is county c's average transportation rate and μ_c is county c's share of manufacturing revenues in 1910. $distance_{cd}^{-1}$ is the inverse of the straight-line distance between county c and any possible destination d.[9] Finally, t_{cd} is the cost of transporting one ton of cargo from county c to any possible destination d before the Canal opens.

$$\bar{P} = \sum_{c} \bar{P}_c \mu_c \tag{3}$$

$$\bar{P}_c = \sum_{c \neq d} distance_{cd}^{-1} \tag{4}$$

We estimate $\bar{P}$ to be equal to 9.4. Although we follow the literature in using $\bar{P}$ in our baseline estimates, we implement a robustness check in which we use $\bar{P}_c$ to compute

[9]The weightings used for the computation of the average are the inverse of instance between county c and destination d, normalized by the sum of all weightings, so the total weightings add up to 1. For this specific computation and to minimize distortions, we use the azimuthal equidistant projection of the world centered around the geographic center of Canada.

market access.

Market Access Calculation

As equations 1 and 2 show, the computation of market access further requires an assumption about trade-to-trade-costs elasticity, θ. We use $\theta = 5$, which is Head's and Mayer's (2014) preferred estimate in the literature. Section 4.2 covers the implementation of several robustness checks, assuming a number of alternative values of θ ranging from 1 to 7.5. Finally, the computation requires population and GDP per capita statistics for each possible destination. For Canada, we use the 1911 and 1921 population statistics and the corresponding GDP per capita figures. For the United States and other countries, we use the 1910 and 1920 population and GDP per capita figures.

Covariate of Interest

We define the change in market access brought about by the opening of the Panama Canal as the difference between actual market access for each Canadian county in 1920 and market access in 1920 in the absence of the Canal:

$$\Delta \ln(MA_{1920}) = \ln(MA_{c,1920}|\mathit{Canal}) - \ln(MA_{c,1920}|\mathit{No\ Canal}) \quad (5)$$

Where $\ln(MA_{c,1920}|\mathit{Canal})$ is the natural logarithm of the market access of county c in 1920 with the transportation network that includes the Panama Canal. $\ln(MA_{c,1920}|\mathit{No\ Canal})$ is its equivalent without the Panama Canal.[10] Both of our market access measures (Canal, No Canal) use 1920 population figures and infrastructure networks. Note that the only change considered here is the possibility of shipping goods through the Canal. Figure 1 shows that the introduction of the Canal into the transportation network does *not* change the minimum-cost path between Vancouver, British Columbia, and Calgary, Alberta, but *does* change the minimum-cost path between Vancouver and New York City. Our measure indicates that *all* counties observed increases in market access in 1920 following the opening of the Canal.

Figure 2 plots a map of Canada for this variable, while Table 1 shows descriptive statistics at the national and provincial levels. These figures indicate that gains in market access due to the Canal: (i) exhibit a substantial degree of domestic variation; (ii) are generally greater in places close to the coast, the Great Lakes or the St. Lawrence River basin; and

[10] We follow the market access literature (Donaldson and Hornbeck, 2016; Hornbeck and Rotemberg, 2021; Redding and Venables, 2004) in specifying our covariate of interest in (changes in) natural logarithms.

(iii) are greater in the western part of the continent. This last pattern arises because counties in the west gained access to the relatively wealthy and populated destinations on the eastern seaboard of North America and Western Europe. Counties in the east gained access to the sparsely populated counties along the west coast of North America and the heavily populated but relatively poorer countries in Asia. Finally, a relatively sharp discontinuity arises between counties in Alberta and Saskatchewan. This follows from the original structure of the rail transportation network in the latter province, which tended to favor rail lines that would effectively cut the distance to the eastern part of North America. The by-product of this feature is that the counties located there were, in terms of the transportation network, effectively farther from the Pacific Coast than what might be expected given their location.

3.2 Manufacturing Outcomes

We collect information on manufacturing activity from the three Censuses of Manufactures conducted in 1901, 1911 and 1939. These sources include data on total revenue, capital stock, wage payments and materials for selected manufacturing establishments. We adjust the data in three ways to ensure consistency over time. First, we account for the fact that earlier censuses surveyed establishments with five or more employees, whereas the 1939 census had full coverage. Following Urquhart and Dales (2007), we use linear expansion factors derived from the (Postal) Census of Manufactures of 1906.[11]. Second, we use 1939 county boundaries and adjust data for other years at this level. To do so, we implement the geographic crosswalks proposed by Eckert et al. (2020), which assume economic activity is evenly distributed across space.[12] Finally, we estimate capital expenses using the value of capital stock provided by the different manufacturing censuses and the interest rate paid on mortgages in Ontario (Homer and Sylla, 1996).

We then digitized data from the 1911 Census of Manufactures, which was collected along with the Census of Population of the same year and included information on outcomes for 1910. It was the last census taken before the opening of the Panama Canal. To check for pre-trends in manufacturing outcomes before the opening of the Canal, we also

[11]We use the information on revenues and inputs provided in the Census of Manufactures of 1906, which is discriminated by establishment size at the province level. In practice, we multiply our 1901 and 1911 figures by a factor that captures how much each variable (revenues, employment, etc.) in 5+ employees establishments in 1906 should expand to equate the variable for *all* establishments in 1906.

[12]The implemented procedure uses the Canadian Century Research Infrastructure boundary files for all censuses with the final, consistent boundaries for 1939. This procedure is a common practice in the literature (e.g., Hornbeck (2010); Hornbeck and Rotemberg (2021); Fajgelbaum and Redding (2014)).

collected data from the 1901 Census of Manufactures. These two were the last censuses to publish data at district-industry level. Though not central to our analysis, this data limitation prevents us from studying the extent to which reallocation happened across industries. Both sources covered establishments employing five or more people. In 1901, butter and cheese factories and brick and tile works were all surveyed regardless of size. In 1911, flour mills, saw and shingle mills, lime, brick and tile works, butter and cheese factories, fish-curing plants, and electric light and power plants were all surveyed regardless of size.

Unlike the 1901 and 1911 censuses, the 1939 census: (i) surveyed all establishments regardless of size; and (ii) did not include what were described as "hand trades". To deal with (i), we use linear expansion factors from the 1906 Census of Manufactures, which surveyed establishments of all sizes, as in Urquhart (1986). To account for (ii), we remove hand-trade sectors directly whenever possible in 1901 and 1911. When they were reported in the "All other industries" category, we scale down those variables by multiplying the original values in this category by 1 minus the share of those activities in the province's aggregate values. Finally, energy and fuel expenditures were not tabulated independently at the county level in 1911 or 1901. We use the industry-level fraction of the gross value of products that relate to this item of expenditure at the national level (inferred from Urquhart) and multiply it by each of the district-industry gross product values.

We compare changes in outcomes between 1939 and 1910 for two reasons: first, even though there were yearly manufacturing censuses starting in 1916, Canada's statistical office published and compiled county-level data from only from 1932 onwards. Then, the tables included only central electrical stations from 1939 onwards. Establishment from this sector were included in 1901 and 1911. Second, the beginning of the Second World War in 1939 disrupted trade and economic activity to such an extent that the effects of that disruption are beyond the scope of this paper.

Productivity Decomposition

Although most of the metrics for manufacturing outcomes are standard (revenues and input expenditures), we also estimate three measures of productivity at the county level, following Hornbeck and Rotemberg (2021). We define total county productivity as the difference between total revenue and total cost of inputs (Basu and Fernald, 2002; Solow, 1957). Following Petrin and Levinsohn (2012), we decompose total productivity into two components: total factor revenue productivity (TFPR) and allocative efficiency (AE). The first relates to growth in revenue beyond the growth implied by increased input use, which is a classical measure of (revenue) productivity. The second relates to changes in surpluses related to the increased use of inputs in locations where they yield a higher productive use due to the presence of imperfections in the economy (such as mark-ups or distortions) that prevent the equalization of marginal revenue products. AE increases when locations that are relatively distorted due to friction increase their input use.

Following Hornbeck and Rotemberg (2021), we define log-productivity as:

$$\ln(Productivity_c) = \xi \left[\ln(P_c Q_c) - \sum_k s_{c,k} \ln(W_{c,k} X_{c,k}) \right] \tag{6}$$

Where $P_c Q_c$ is the gross value of manufacturing establishments (revenue) in county c, $s_{c,k}$ is the revenue share for input k in county c, and $W_{c,k} X_{c,k}$ are expenditures in input k in county c. The set of inputs k consists of capital expenditures, labor, and materials (including fuel and power). Finally, ξ is a scalar that converts the expression from output growth to productivity growth.[13] This measure captures changes in revenues over changes in revenue-share-weighed expenditures in inputs.

As in Petrin and Levinsohn (2012), the previous measure of productivity can be further decomposed into growth in TFPR and growth in AE:

$$TFPR_c = \xi \left[\ln(P_c Q_c) - \sum_k \alpha_{c,k} \ln(W_{c,k} X_{c,k}) \right] \tag{7}$$

$$AE_c = \xi \left[\sum_k (\alpha_{c,k} - s_{c,k}) \ln(W_{c,k} X_{c,k}) \right] \tag{8}$$

Where $\alpha_{c,k}$ corresponds to the output elasticity of input k in county c. TFPR and AE refer to different notions of productivity. The former captures whether, for the same level of inputs, agents can produce more output given the existing production technology. The latter emerges whenever we move from a perfectly competitive environment and allow for imperfections that prevent the equalization of marginal products and marginal costs across sectors and places in the economy. Because we allow for those differences between marginal products and costs to exist and for them to vary over sectors and locations, an additional unit of a given input—whether it is labor, capital or materials—yields different increases in production in different places and different sectors. The reallocation of inputs between sectors and locations can increase productivity and output because those inputs

[13] The productivity scalar used here is $\xi = 1/[1 - \frac{1}{C} \sum_c \sum_k s_{k,c}]$

can now be used in places or sectors where they produce greater economic value. Why? Because we allow for the differences between marginal products and costs to exist and vary over sectors and places. These differences are captured in equation 8 above by the term $(\alpha_{c,k} - s_{c,k})$.[14]

In practice, recovering the expressions in equations 7 and 8 require estimations of the output elasticities $\alpha_{c,k}$. We assume a Cobb-Douglas production function with constant returns to scale, which implies that these elasticities can be retrieved, under the assumption of cost minimization, from each input's cost shares. Robustness checks presented in Section 4.2 relax the constant returns to scale assumption while keeping the assumption of a Cobb-Douglas production structure. To compute $TFPR_c$ and AE_c in 1910 and 1939, we fix $\alpha_{c,k}$ and $s_{c,k}$, which also implies that input wedges $(\alpha_{c,k} - s_{c,k})$ do not change. We use the simple average of the computations for each county in 1900 and 1910 for our calculations.

One important caveat should be noted in this study. As in Petrin and Levinsohn (2012), we are using the production function approach to estimate mark-ups at the county level since we do not have access to establishment level data. Moreover, we abstract from the economy's network composition. Baqaee and Farhi (2020) argue that our productivity decomposition could benefit from incorporating input-output network data, considering information at the level where mark-ups are applicable, and employing more direct measures of mark-ups. Notably, Baqaee and Farhi (2020) demonstrate that the Petrin and Levinsohn decomposition underestimates the contribution of allocative efficiency, using contemporary US data. We acknowledge that this result may indicate that our findings in this paper represent a conservative estimate of allocative efficiency's actual contribution to productivity growth. Unfortunately, due to data limitations, we are unable to improve our productivity decomposition in accordance with the recommendations made by Baqaee and Farhi (2020).

[14] These concepts can be illustrated with a hypothetical example: Toronto is a manufacturing center where producers are highly productive, and markets work relatively well, so the differences between marginal output and marginal costs are small. Vancouver is a city where wages and the return to capital are the same as Toronto's but where wood is much cheaper because it is surrounded by forests. Production there, however, is constrained because it is distant from most major markets on the eastern seaboard of North America. Now let us assume there are two identical carpenters in both cities. Whenever their carpentry skills improved so that they could use less wood to manufacture a table, Toronto and Vancouver's TFPR increased. If the Vancouver carpenter started working more hours manufacturing tables leaving their skills constant, Vancouver's AE would increase. Note that, in a perfectly competitive economy where the difference between marginal products and marginal costs are equalized – to zero across sectors and places – the carpenter's increased labor in Vancouver yields no increase in productivity.

4 Empirical Analysis

Main Specification

Equation 9 shows the specification estimated for this study:

$$\begin{aligned} \Delta_{1939,1910} ln(Y_c) &= \beta \Delta \ln(MA_{1920}) + \gamma_p + \eta G_c + \epsilon_c \\ \Delta_{1939,1910} ln(Y_c) &= ln(Y_{c,1939}) - ln(Y_{c,1910}) \end{aligned} \tag{9}$$

Where Y_c is an outcome of interest for county c and $\Delta \ln(MA_{1920})$ is the causing variable of interest. γ_p are province-level fixed effects. G_c is a vector of county-level controls that includes: (i) a quadratic polynomial on latitude and longitude of county c; (ii) its distance to the coast, the Great Lakes or the St. Lawrence River; and (iii) (log) population in 1911. The coefficient of interest is β. This empirical specification is derived from the model that we solve for in Section 5 and that predicts a log-linear relation between the outcomes we study and market access.

We estimate equation 9 using ordinary least squares. This specification leverages within-province variation in changes in market proximity, net of geographic and baseline controls. The identification assumption is that counties within the same province, after controlling for county variables, would have followed a similar development path if the Panama Canal shock had not materialized. This specification allows us to isolate the effect of the Panama Canal over potential confounders. However, Table A.1 shows our results do not rely on the inclusion of province fixed effects or geographic controls. We report standard errors clustered at the level of 43 arbitrary grids measuring 300km by 300km to allow for a potential spatial correlation of the residuals. We also allow for smaller grid sizes (150km, 200km) that result in larger numbers of clusters as well as Conley standard errors. All our results, which are included in the Online Appendix A, are robust to these alternatives for assessing sampling variability.

4.1 Empirical Results

Panel A in Table 2 presents the results of the estimation of equation 9. These are reported, for ease of interpretation, in terms of the interquartile range of (log) market access due to the Panama Canal in 1920.[15] Moving from the 25th to the 75th percentile in terms of expo-

[15]We use the interquartile range of the residual (within) variation of our causing variable, which is the one we use for identification and which equals 0.0027. The corresponding magnitude for the original causing

sure to the Canal shock leads to an increase in revenues of 8.9%, in capital expenditures of 7.0%, in labor expenditures of 9.7%, in employment of 9.1%, in material expenditures of 8.3% and in measured productivity of 13.1%. The point estimate for effects in TFPR is small and not statistically different from zero. Increases in productivity seem to be driven by improvements in AE that reflect increased input usage in places where they yielded a higher productive use.

The specification detailed in this section relies crucially on the assumption that the change in market access is exogenous. Simultaneity can be ruled out in this case because the construction of the Canal was not determined by the economic outcomes of Canadian counties. Still, omitted variables could be confounders. One concern is that the Canadian counties that benefited the most from the building of the Panama Canal possibly would have been growing faster in any event. To assess this concern, we conduct a pre-trends analysis and conclude that our causing variable is unable to predict the changes in the outcomes of interest between 1900 and 1910. Panel B of Table 2 shows that the point estimates for $\Delta \ln(MA_{1920})$ are generally small and very noisy. Taken as a whole, Table 2 suggests that the variation that we exploit to identify our parameters of interest is not correlated with unobservable secular county-level trends.

Another concern that might arise relates to potential endogenous responses of the transportation network in anticipation of the completion of the Canal. Table 3 shows that $\Delta \ln(MA_{1920})$ has no predictive power over the (log) kilometers of railroads built in Canadian counties over different time periods. We interpret these results as evidence that the identification of parameters is not relying on variations deriving from locations that were targeted for railroad construction. These findings are consistent with our pre-trends analysis, as we would expect that counties that were growing faster would also have been likely targets for railroad construction. Our identifying variation is orthogonal to these phenomena. Indeed, robustness checks in which we explicitly control for local railroad construction before and after 1920 show that our point estimates remain essentially unchanged.

The results change very little when the population is fixed at its 1911 level. We interpret this as evidence that the results are driven by the (exogenous) reduction in transportation costs brought about by the Canal rather than by (potentially endogenous) population movements in anticipation of its completion. Further, and as robustness exercises in Section 4.2 show, our results are robust to controlling for potential *local* responses of the

variable is 0.0045.

transportation network to the Canal shock before and after 1920.

The results outlined above indicate that the Panama Canal had economically significant effects on Canadian manufacturing during the study period. They suggest several patterns. First, the Canal altered the economic geography of Canada: production inputs were used relatively more in counties that benefited more from the reduction in transportation costs. Second, increases in input usage occurred across the board and on relatively similar scales. Indeed, the 95% confidence intervals for the coefficients of interest on capital expenditures, the wage bill and material expenditures overlap. This suggests that the expansion in production brought about by the Canal was not biased toward a specific factor of production.

The gains in productivity were substantial and most likely driven by improvements in AE, not revenue productivity. This is consistent with the idea that the Canal induced a relative increase in the use of inputs in activities where they yielded a higher productive use – i.e. marginally productive firms or industries. The small point estimate and lack of statistical significance of the parameter associated to TFPR suggests, on the other hand, that the Canal brought modest improvements in the value of technical productivity, which is consistent with limited gains within the transportation sector. The bulk of the productivity gains, therefore, came about in sectors other than transportation. This result is consistent with anecdotal evidence from key industries that benefitted from the Canal. For instance, (Lawrence, 1957, ch. III) documents how during the Canal's first five years the quantity of lumber products traded through British Columbia to the US increased over 80 times. In 1920, Canada exported 4 million feet of lumber to the US. In 1925, this quantity had changed to 326 million feet. Moreover, in 1929 British Columbia exported around 60% of its output to the Eastern seaboard of the United States.

This evidence suggests that transportation infrastructure induced the reallocation of inputs and economic activity to activities where production levels were inefficiently low due to the presence of distortions and imperfections in the economy. This reallocation led to increases in AE, as highlighted by our results, because inputs shifted to activities where the difference between the value of marginal products and their marginal cost was higher, yielding an increase in the surplus generated in the county. We note that these gains in productivity are different to classical gains from trade that are typically associated with specialization and the division of labor and production – elements that are captured in our decomposition by TFPR. These findings can be rationalized in a world in which there are constant returns to scale (CRS) in manufacturing and constant elasticity of substitution

(CES) demand, so increased scale does not yield improvements in technical productivity (TFP) or its value (TFPR).

As highlighted by Baqaee and Farhi (2020), our productivity decomposition is limited by the available data. However, we believe it still provides valuable insights into the nature of reallocation that occurs with reductions in transportation costs. To offer some intuitive understanding of which areas experienced larger productivity gaps, we employ three approaches.

First, we present the distribution of wedges for all inputs in Figure A.1. The capital and labor wedges exhibit geographical correlations with regions that significantly benefited from the Canal. Second, we demonstrate the persistence of this geographical correlation by examining estimated Market Access in 1910. Figure A.2 confirms that higher Market Access in 1910 correlates with lower input wedges. Third, we establish a connection between higher sectoral concentration and elevated wedges. Places with greater concentration within industries are also associated with higher wedges. For the 1911 Manufacturing Census, we collected data on industry shares in manufacturing, considering both the value of production and employment. Using these shares, we calculated Herfindahl-Hirschman indexes, where higher values indicate greater concentration within industries. Table A.3 demonstrates a positive and statistically significant correlation between employment-based HHI and capital and labor wedges.

These three findings align with a narrative in which input usage increases in industries with lower revenue shares in highly concentrated regions, leading to enhanced productivity through allocative efficiency. However, it is crucial to note that this represents just one possible explanation, and we cannot discount other potential narratives given the limitations of the available historical data.

Taken together, our results indicate that the Panama Canal changed the economic geography of Canada during the first half of the twentieth century. Although to varying degrees, *all* the counties observed changes in market access as a consequence of the completion of the Canal. The counties that benefited the most, i.e. those that experienced greater increases in market access, enjoyed sizeable gains in economic activity. Our reduced-form evidence, though necessary to identify *comparative winners*, does not allow us to distinguish between new and displaced economic activity within Canada and, hence, is insufficient to evaluate country-wide effects. To address this concern, in Section 5 we assess the aggregate impact of the Canal by relying on the calibration of a standard general equilibrium model of economic geography. We study how population, production inputs and productivity would have changed in a counterfactual scenario in which the Canal is closed permanently in 1939.

4.2 Robustness Checks

Table 4 reports a variety of robustness checks that address concerns related to assumptions made while calculating market access and computing productivity. We present estimates that assume different values for θ, the trade elasticity in Panels B and C. Our qualitative results and conclusions remain unchanged when we assume $\theta = 1$ or $\theta = 7.5$.[16] This is because a different trade elasticity alters both elements used in the construction of our causing variable in a similar fashion. We then assume different average transportation costs while keeping all other baseline assumptions fixed for the computation of market access. Estimates in Panels D and E change very little and the conclusions remain unaltered when we set $\bar{P} = 5$ or $\bar{P} = 20$. Estimates in Panel F allow for a county-level average transportation cost for all destinations in Canada and the United States that is simply the inverse-distance-weighted average transportation cost for each county before the Canal opened. Although some estimates are slightly less precise, point estimates are similar to those in the baseline specification and do not change our conclusions.

Table 5 presents a different set of robustness checks that tackle concerns related to potential omitted variables. We control for: (i) baseline market access (Panel A); (ii) new railroad construction between 1920 and 1940 (Panel B); and (iii) new railroad construction between 1910 and 1920 (Panel C). Panel A provides evidence that our identification does not rely on better differential access to the transportation network itself in 1910, as our results remain unaltered when we control for baseline market access in 1910. This result is important since it might be the case that areas closer to the sea would have grown faster than other areas, simply because of their privileged location. Panels B and C provide evidence that our estimates are not leveraging potential endogenous responses of the Canadian transportation network to the opening of the Panama Canal. Although such responses may well have taken place, they are not omitted confounders of the parameter of interest in equation 9. Altogether, our point estimates change very little when we include all these additional controls.

Another potential concern is that there could be local shocks to economic conditions that are correlated with gains in market access. In our main specification, we exploit vari-

[16]Figure A.3 shows that our point estimates and standard errors change very little when we use different values of trade elasticity.

ation at the province level by adding province fixed effects. However, we also go a step further and leverage the panel data feature of our data to control for the potentially differential development of manufacturing activity at the local level. First, we replicate our main results estimating the following equation:

$$
\begin{aligned}
ln(Y_{ct}) = \beta_{1900}\Delta \ln(MA_{1920}) \times Year_{1900} + \beta_{1939}\Delta \ln(MA_{1920}) \times Year_{1939} + \dots \\
\dots \gamma_c + \gamma_t + \eta G_{ct} + \epsilon_{ct}
\end{aligned} \tag{10}
$$

Where $Year_t$ is a dummy variable equal to 1 for year t, γ_c, γ_t are county and year fixed effects, and G_{ct} is the same set of time-invariant covariates from equation 9 interacted with year fixed effects. In separate exercises, we add differential linear trends for each province and Province x Year fixed effects to this baseline specification. Table 6 shows the results from the three different specifications, one in each panel. Panel A includes only County and Year fixed effects. Our results do not change when we allow for Province linear trends or Province x Year fixed effects.

5 Counterfactual Analysis: Closing the Panama Canal in 1939

What were the *aggregate* effects of the Panama Canal on the Canadian economy? Our previous analysis does not necessarily enable us to answer this question, since it is possible that the observed effects are the result of the displacement of economic activity across counties. Although the mere shifting of inputs could still imply a higher level of productivity and economic activity, we want to establish whether the Canal induced a higher *aggregate* level of economic activity in Canada in a manner that allows for displacement and general equilibrium effects to occur.

To answer this question, we calibrate an extension of a benchmark model of economic geography (Eaton and Kortum, 2002). The extension developed by Donaldson and Hornbeck (2016) and Hornbeck and Rotemberg (2021) features a market structure in which there are exogenous differences between factor prices and their marginal products, input market frictions or mark-ups. This enables us to assess the effects on AE of any counterfactual shock of interest. Further, it provides a measure of changes in land rental rates that can be used to perform a back-of-the-envelope calculation of the impact of closing the

Canal that includes both the manufacturing and agriculture sectors.

We use this model to conduct two counterfactual exercises. First, we close the Panama Canal permanently in 1939. This exercise enables us to provide an answer to the question regarding *aggregate* effects. Second, we allow the use of the Canal *only* for domestic trade within Canada. This exercise allows us to study whether most of the gains derived from the Canal came from greater exposure to international or domestic markets. In both cases, we solve for counterfactual equilibrium output, input prices and quantities. We then use those to calculate the changes in productivity levels that result in a new equilibrium. We conclude by providing a back-of-the-envelope calculation that also incorporates effects on the agricultural sector.

5.1 Model Primitives[17]

Production in each county is undertaken by firms that maximize profits by optimally choosing inputs while taking their price and input frictions as givens. The production technology is characterized by a Cobb Douglas production function for variety j that utilizes capital (K), labor (L), land (T), and an intermediate good (M) as inputs. These are paid, respectively, an interest rate r_c, a wage w_c, rent q_c, and price P_c. Producers have CES preferences over the continuum of varieties that are used as intermediates with a CES of σ. P_c is, therefore, a CES Price Index. Following the standard assumption in Eaton and Kortum (2002) and the results in our reduced form exercise in Section 4, each county has an exogenous technical efficiency level for variety j which is drawn from the Fréchet distribution with CDF $F_c = 1 - e^{-A_c z^{(-\theta)}}$. The marginal cost of production is thus characterized by equation 11:

$$MC_c(j) = \frac{r_c^{\alpha_c^K} w_c^{\alpha_c^L} P_c^{\alpha_c^M}}{z_c(j)} = \frac{\Pi_k (w_c^k)^{\alpha_c^k}}{z_c(j)} \tag{11}$$

Where K inputs are capital, labor, land, and intermediate inputs. By assumption, producers face frictions in each one of the input markets in which they participate. These are taken as given and exogenous. They prevent firms from using inputs up the point where price equals marginal cost. In this context, $1 + \psi_c^k$ in equation 12 captures the factor k specific input friction that embodies firms' inability or unwillingness to expand production beyond a given level. Note that these are fixed and exogenous to the model. We make this assumption because our covariate of interest is unable to predict changes in these wedges

[17]This subsection follows Hornbeck and Rotemberg (2021) closely.

(See Table A.7).

$$p_c(j) = \frac{\Pi_k((1+\psi_c^k)w_c^k)^{\alpha_c^k}}{z_c(j)} > MC_c(j) \tag{12}$$

Workers supply labour inelastically, receive a wage w_c, and have CES preferences over the j good varieties in the same way as firms do. The indirect utility takes the form of $V = w_c/P_c$, with both w_c and P_c being endogenously determined within the model. To focus on a hypothetical long-run equilibrium, we assume that workers are perfectly mobile across counties so any difference in indirect utility can be arbitraged out. Note that this is the case for both Canada and the United States, which we see as areas where internal migration might be plausible over the long run. [18] This leads us to set $V = \bar{U}$. This implies that, although nominal wages can be different across places, they simply reflect higher nominal prices. When we conduct the counterfactual exercise, we fix worker's utility at this initial level and allow for changes in the total population in our area of study *after* the Canal shock.

We assume that capital is perfectly mobile so that interest rates are equalized across counties ($r_o = r$). Further, we assume that Canada faces a perfectly elastic supply of capital with the interest rate being exogenous and set abroad. We also assume that the land supply is completely inelastic and set at exogenous levels. The remuneration of land –rents– is, however, endogenous to the model. Trade in final or intermediate goods and between counties c and d can take place while incurring an iceberg trade cost τ_{cd}, which is defined in the same fashion as equation 2. The price in county c of a good produced in county d is $p_{cd}(j) = \tau_{cd} p_{dd}(j)$, where $p_{dd}(j)$ is the price of good-variety j in county d and $\tau_{cd} > 1$ for all $c \neq d$. Goods markets clear in general equilibrium, so demand and supply equate with each other. Production in each county equals the sum of exports to all possible destinations plus within-county sales.

5.2 Analytical Results

Solving the model yields a set of equations that are useful in understanding how changes in transportation costs might affect economic activity. A first important result, in line with Eaton and Kortum (2002), is given in equation 13. Also known as a "gravity equation", it

[18]The United States was an important source of migration to Canada over the first half of the twentieth century. By 1921, over 4% of the Canadian population and 19% of Canada's foreign-born population was born in the United States. Furthermore, between 1931 and 1940, the United States was the top country of origin of immigrants to Canada (StatisticsCanada, 2016).

states that exports from county c to county d (E_cd) are positively related with the origin's technical efficiency (A_c), the destination's income (Y_d) and its price level (P_d) (Chaney, 2018). On the other hand, it is inversely related with transportation costs (τ_{cd}), input prices (W_c^k) and distortions in the county of origin ($1+\psi_c^k$). This provides a rationalization for an important stylized fact in the international trade literature: countries with larger economic mass (income) and closer to each other (with lower trade costs) tend to trade more among themselves.

$$E_{cd} = \kappa_1 A_c \left(\Pi_k \left((1+\psi_c^k) W_c^k \right)^{\alpha_c^k} \right)^{-\theta} \tau_{cd}^{-\theta} Y_d P_d^{\theta} \tag{13}$$

$$\kappa_1 = \left(-\frac{\theta}{1-\sigma} \right) ln \left(\Gamma \left(\frac{\theta + 1 - \sigma}{\theta} \right) \right)$$

A second result is that market access – a concept that we exploit in our reduced-form work – is an inverse transformation of the CES price index. Moreover, and as Equation 14 shows, a county's market access (MA_c) is inversely related to transportation costs (τ_{cd}) and is greater whenever the locations to which a given county has access have a higher income or greater market access themselves. A third result is that, in general equilibrium, changes in a county'a market access summarize how changes in transportation costs affect each county's economic activity through changes in both goods and factor markets.[19] This relation is described in equation 14 below.[20]

$$P_c^{-\theta} = MA_c = \kappa_2 \sum_d \tau_{cd}^{-\theta} Y_d MA_d^{\frac{-(1+\theta)}{\theta}} \tag{14}$$

$$lnY_c = \kappa_1 + \kappa_{2c} + \left(\frac{\alpha_c^M + \alpha_c^L + 1}{1 + \theta_c^T} \right) ln(MA_c) \tag{15}$$

A fourth result is that the model predicts how changes in market access will affect productivity and input prices. Given the assumption that technical efficiency is exogenous, changes in county-level productivity should come from changes in AE. Equation 16 shows that this is indeed the case, with changes in county-level productivity being driven by changes in equilibrium input quantities. Equations 17 and 18 show that increases in market access yield log-linear responses in equilibrium input prices. Note that there is no

[19] For a detailed discussion of consumer market access and firm market access, see Redding and Venables (2004) and Hornbeck and Rotemberg (2021).

[20] Where $\kappa_2 = \bar{U}\rho^{\frac{1+\theta}{\theta}}$ and $\kappa_{2c} = \frac{\kappa_{1c} + ln\psi_c - \theta\alpha_c^T \ln \frac{\alpha_c^T}{T_c}}{1+\theta\alpha_c^T}$

equation for the remuneration of capital as it is assumed to be exogenous to the model.

$$\frac{d\ln PR_c}{d\ln MA_c} = \frac{P_c Q_C}{Pr_c} \sum_k \left(\alpha_c^k - s_c^k\right) \frac{d\ln X_c^k}{d\ln MA_C} \tag{16}$$

$$\frac{d\ln q_c}{d\ln MA_c} = \frac{\alpha_c^M + \alpha_c^L + 1}{1 + \theta\alpha_c^T} \tag{17}$$

$$\frac{d\ln w_c}{d\ln MA_C} = \frac{d\ln P_c}{d\ln MA_C} = -\frac{1}{\theta} \tag{18}$$

5.3 Data and Calibration

We allow for trade to take place between counties in the United States and Canada. To solve for prices, we use population data from the Canadian Census of Population of 1941 and the United States Population Census of 1940. We adjust these figures to obtain the size of the population in 1939 under the assumption that the population in each county grew at a constant rate from 1931/1930 to 1941/1940. We use our data on transportation costs, with and without the Panama Canal, to compute iceberg trade costs. We use data from the Canadian Census of Manufactures of 1939 to compute measures of the distortions and output elasticities. For this counterfactual exercise, we assume output elasticities and distortions in each input that result from output-weighted measures of the county-level information that we retrieve from the Census of Manufactures.[21] As in our empirical work, we continue to follow the literature in assuming $\theta = 5$.

We follow the procedure outlined in Donaldson and Hornbeck (2016) and Hornbeck and Rotemberg (2021) to compute the effects of our two counterfactual exercises. This entails: (i) solving for prices using the iceberg trade costs assuming that the Canal is open – that is, factual trade costs. Then: (ii) we use each county's factual population to solve for an "amenity" that captures each county's fixed endowments of land and productivity, which are not observed in the data. Next: (iii) we use the "amenity" and counterfactual trade costs to compute the counterfactual distribution of population under the assumption that the Canal closes. Then: (iv) we solve for prices again using the counterfactual distribution of population. Next: (v) we use equation 16 to compute changes in output. Then: (vi) use our assumption that revenue shares are constant to retrieve changes in input bills. Finally: (vii) use equations 17 and 18 to determine what fraction of the change in

[21]We set $\alpha_c^T = 0.1748$ (Caselli and Coleman, 2006), $\alpha_c^L = 0.255$, $\alpha_c^K = 0.0768$, and $\alpha_c^M = 0.669$. These come directly from the data.

input bills relates to changes in real input demand rather than input prices. We then have a measure of how much equilibrium quantities of labor, capital and intermediate goods inputs changed in our counterfactual scenario.

To compute the impact on manufacturing productivity of such changes as a fraction of the manufacturing sector's GDP, we use equation 19:

$$\begin{aligned} ChangeProductivity =& \frac{1}{GDP}\sum_c dProductivity_c \\ ChangeProductivity =& \frac{1}{GDP}\sum_c Productivity_c dln(Productivity_c) \end{aligned} \tag{19}$$

Given that, in the framework we are using, changes in productivity can only come from changes in AE[22], we can further simplify the expression to:

$$\begin{aligned} ChangeProductivity =& \frac{1}{GDP}\sum_c Productivity_c dln(AE_c) \\ =& \frac{1}{GDP}\sum_c P_c Q_c \sum_k \left(\alpha_c^k - s_c^k\right) d\ln(X_c^k) \\ =& \sum_c D_c \sum_k \left(\alpha_c^k - s_c^k\right) d\ln(X_c^k) \end{aligned} \tag{19}$$

Where $D_c = \frac{P_c Q_c}{GDP}$ are Domar weights and $d\ln(X_c^k)$ is the change in input k quantity as predicted by the model. Here, the K inputs are labor, capital and materials. Land is not included, as it is assumed to be in fixed supply and, hence, always used at efficient levels, which implies $d\ln(X_c^T) = 0$. We take $(\alpha_c^k - s_c^k)$ directly from Section 4 of this study.

5.4 Results

We find that closing the Panama Canal *permanently for all traffic* in 1939 would lead to substantial changes in Canada's economic geography. First, its total population would decrease by 2.7%. Second, as shown in Figure 3, those losses would be concentrated in places closer to the coasts and would be particularly severe in counties located in the western provinces. British Columbia's and Alberta's population would fall, respectively, by 7.8% and 4.2%. Counties closer to the Atlantic Ocean or the St. Lawrence River basin would see

[22] Changes in AE might be written as $d\ln(AE_c) = \frac{P_c Q_c}{Productivity_c}[\sum_k (\alpha_c^k - s_c^k) d\ln(X_c^k)]$. See equation (20) in Hornbeck and Rotemberg (2021).

decreases too, but smaller ones than those that would occur in counties located closer to the Pacific Ocean. This result suggests that the Panama Canal shock was probably stronger for the Pacific basin of the Western Hemisphere than it was for the Atlantic basin. Finally, some counties in the prairie provinces, such as Manitoba and Saskatchewan, would observe marginal changes in population. The populations of these two provinces would increase by 0.07% and 0.2%, respectively.

We also calculate substantial impacts in terms of changes in manufacturing productivity as defined in equation 19. We calculate that these losses would add up to 1.2% of manufacturing GDP in 1939 or 0.27% of Canadian GDP in 1939. These losses, as shown in equation 11, reflect both changes in aggregate inputs used in manufacturing and differences in distortions in different counties of Canada. Our results here indicate that, although there could be relevant displacements of production within Canada, the Canal paved the way for increases in the aggregate level of manufacturing activity in the country.

The model we are using also provides information on changes in land rents. This enables us to provide a back-of-the-envelope calculation of the impact of the Canal shock on the Canadian economy that includes both the manufacturing and agricultural sectors. As emphasized in Donaldson and Hornbeck (2016) and due to its immobility, land values capitalize gains stemming from increases in market access. We find that land rents fall 2.6% in our counterfactual scenario. If land values are the present discounted value of rents, we find that closing the Canal in 1939 would yield a permanent economic loss equivalent to 1.59% of Canadian GDP. This loss is additional to the estimate discussed in the previous paragraph and relates only to the agricultural sector. Thus, the total general equilibrium losses caused by closing the Panama Canal come to 0.27% + 1.59% = 1.86% of Canadian GDP in 1939.

To contrast the model's predictions, we estimate the county level effect of Market Access gains on land values. We do so by digitizing data on total land values and total area in farms from 4 different Canadian Census of Agriculture: 1901, 1911, 1921, and 1941. In an event study framework, we regress log average land values on our measure for MA gains due to the Panama Canal interacted with year dummies, leaving 1921 as reference year. Table 8 shows the results from these exercises. Column (1) only include year and county fixed effects. Column 2 includes 1911 population interacted with year dummies, our geographic controls (coordinates and distance to the coast) interacted with year dummies, and province by year fixed effects. Regardless of the specification, the message of

Table 8 is consistent with the implications of the model: moving one county from the 25th percentile of MA gains to the 75th percentile would increase land rents by 3.3%.

Table 9 compares the results obtained from our counterfactual exercise with those found in the international trade and economic geography literature, which suggest that closing the Panama Canal would be much less consequential for Canada than removing the railroad network for the United States in 1890 (Donaldson and Hornbeck, 2016; Hornbeck and Rotemberg, 2021) or for Argentina in 1914 (Fajgelbaum and Redding, 2014). Our quantitative results, however, are similar in magnitude to those from Hornbeck and Rotemberg's counterfactual exercise in which the United States railroad network is held fixed as it was in 1880 (as opposed to enjoying a decade, from 1880 to 1890, of (frenzied) railroad construction). Closing the Canal in 1939, for instance, would be much more consequential for Canada than freezing the railroad network in 1920 (with the Panama Canal open). Compared to the 1939 factual equilibrium, doing so would yield population losses of just 0.34% and economic losses (in agriculture and manufacturing) of only 0.2% of GDP.

We argue that the small, although certainly non-negligible, impacts of the Canal on Canada relate to: (i) the relatively greater importance of existing railroads for the transportation network than the Canal; and (ii) the fact that the economies of the main beneficiaries of the Canal – the western provinces of Canada – were still relatively small in 1939. Although we see substantial population losses in British Columbia and Alberta, these two provinces accounted for only 14% of the total population and 8.4% of manufacturing employment at the time. This reduces the aggregate relevance of the Canal shock for the Canadian economy.

We conclude this section by reporting the results of the exercise in which we close the Panama Canal *permanently* but *only* for traffic originating from or destined for locations outside Canada. That is, we allow the use of the Panama Canal *only* for domestic trade within Canada. We do so to understand the domestic or international origins of the effects we discuss above. We find that, compared to the initial 1939 equilibrium that features a fully open Canal, the counterfactual population falls by 2.4% and GDP by 1.74%. The losses come from lower productivity in manufacturing (0.25% of GDP) and lower agricultural land values (1.49% of GDP) due to reduced land rents. These represent roughly 93% of the losses we find in our first counterfactual exercise, which suggests that most of the gains brought about by the Panama Canal stemmed from cheaper access to international markets. Gains originating from domestic trade are much smaller.

6 Conclusion

We find that the opening of the Panama Canal had substantial impacts on Canadian economic life. First, by leveraging comparatively cheap maritime transportation, the Canal paved the way for gains in market access for all Canadian counties. Thus, the Panama Canal changed the economic geography of Canada. Second, those gains drove the growth of the manufacturing sectors of the counties that benefited more from the Canal than others. Productivity increased, thanks mostly to improvements in the allocative efficiency of the use made of production factors. The Canal allowed a greater use of inputs in places where they yielded a higher productive use – i.e. marginally productive counties. Third, these effects were not purely local, as they persist in general equilibrium. The calibration of a benchmark economic geography model indicates that closing the Canal permanently in 1939 would have led to losses of 1.86% of Canadian GDP in 1939. Most of those losses would be attributable to decreased access to overseas markets. Consequently, it turns out that there was, after all, a free ride.

References

Alexander, P. D. and Keay, I. (2019). Responding to the First Era of Globalization: Canadian Trade Policy, 1870–1913. *The Journal of Economic History*, 79(3):826–861. Publisher: Cambridge University Press.

Asher, S. and Novosad, P. (2020). Rural Roads and Local Economic Development. *American Economic Review*, 110(3):797–823.

Asturias, J., García-Santana, M., and Ramos, R. (2019). Competition and the Welfare Gains from Transportation Infrastructure: Evidence from the Golden Quadrilateral of India. *Journal of the European Economic Association*, 17(6):1881–1940.

Atack, J., Bateman, F., Haines, M., and Margo, R. A. (2010). Did Railroads Induce or Follow Economic Growth?: Urbanization and Population Growth in the American Midwest, 1850–1860. *Social Science History*, 34(2):171–197. Publisher: Cambridge University Press.

Banerjee, A., Duflo, E., and Qian, N. (2020). On the road: Access to transportation infrastructure and economic growth in China. *Journal of Development Economics*, 145:102442.

Baqaee, D. R. and Farhi, E. (2020). Productivity and misallocation in general equilibrium. *The Quarterly Journal of Economics*, 135(1):105–163.

Basu, S. and Fernald, J. G. (2002). Aggregate productivity and aggregate technology. *European Economic Review*, 46(6):963–991.

Bolt, J. and van Zanden, J. L. (2020). Maddison project database, version 2020. *URl: https://www. rug. nl/ggdc/historicaldevelopment/maddison/releases/maddison-project-database-2020 (visited on 04/23/2021)(cit. on p. 25)*.

Brancaccio, G., Kalouptsidi, M., and Papageorgiou, T. (2020). Geography, Transportation, and Endogenous Trade Costs. *Econometrica*, 88(2):657–691. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA15455.

Cao, Y. and Chen, S. (2022). Rebel on the Canal: Disrupted Trade Access and Social Conflict in China, 1650–1911. *American Economic Review*, 112(5):1555–1590.

Carlos, A. M. and Lewis, F. (1992). The Profitability of Early Canadian Railroads: Evidence from the Grand Trunk and Great Western Railway Companies. In *Strategic Factors in Nineteenth Century American Economic History: A Volume to Honor Robert W. Fogel*, pages 401–426. University of Chicago Press.

Chaney, T. (2018). The Gravity Equation in International Trade: An Explanation. *Journal of Political Economy*, 126(1).

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

Donaldson, D. and Hornbeck, R. (2016). Railroads and American Economic Growth: A "Market Access" Approach. *The Quarterly Journal of Economics*, 131(2):799–858.

Eaton, J. and Kortum, S. (2002). Technology, Geography, and Trade. *Econometrica*, 70(5):1741–1779. Publisher: [Wiley, Econometric Society].

Eckert, F., Gvirtz, A., Liang, J., and Peters, M. (2020). A method to construct geographical crosswalks with an application to us counties since 1790. Technical report, National Bureau of Economic Research.

Fajgelbaum, P. and Redding, S. J. (2014). Trade, Structural Transformation and Development: Evidence from Argentina 1869-1914. Working Paper 20217, National Bureau of Economic Research. Series: Working Paper Series.

Fogel, R. (1964). *Railroads and American Economic Growth: Essays in Economic History*. Johns Hopkins Press, Baltimore.

Head, K. and Mayer, T. (2014). Chapter 3 - Gravity Equations: Workhorse,Toolkit, and Cookbook. In Gopinath, G., Helpman, E., and Rogoff, K., editors, *Handbook of International Economics*, volume 4 of *Handbook of International Economics*, pages 131–195. Elsevier.

Homer, S. and Sylla, R. E. (1996). *A History of Interest Rates*. Rutgers University Press.

Hornbeck, R. (2010). Barbed wire: Property rights and agricultural development. *The Quarterly Journal of Economics*, 125(2):767–810. Publisher: MIT Press.

Hornbeck, R. and Rotemberg, M. (2021). Railroads, Reallocation, and the Rise of American Manufacturing. Working Paper 26594, National Bureau of Economic Research. Series: Working Paper Series.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Huebner, G. G. (1915). Economic Aspects of the Panama Canal. *The American Economic Review*, 5(4):816–829. Publisher: American Economic Association.

Innis, H. (2018). *Essays in Canadian Economic History*. University of Toronto Press. Publication Title: Essays in Canadian Economic History.

Innis, H. A. (1933). *Canada and the Panama Canal*.

Jacks, D. S. and Novy, D. (2018). Market Potential and Global Growth over the Long Twentieth Century. *Journal of International Economics*, 114:221–237.

Jaworski, T. and Keay, I. (2021). Globalization and the Spread of Industrialization in Canada, 1871–1891.

Keay, I. (2007). The Engine or the Caboose? Resource Industries and Twentieth-Century Canadian Economic Performance. *The Journal of Economic History*, 67(1):1–32. Publisher: Cambridge University Press.

Krugman, P. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy*, 99(3):483–499. Publisher: The University of Chicago Press.

Lawrence, J. C. (1957). *Markets and capital: a history of the lumber industry of British Columbia (1778-1952)*. PhD thesis, University of British Columbia.

Limão, N. and Venables, A. J. (2001). Infrastructure, Geographical Disadvantage, Transport Costs, and Trade. *The World Bank Economic Review*, 15(3):451–479.

Liu, D. and Meissner, C. M. (2015). Market potential and the rise of US productivity leadership. *Journal of International Economics*, 96(1):72–87.

Mackintosh, W. A. (1939). *The economic background of dominion-provincial relations : a study prepared for the Royal Commission on Dominion-Provincial Relations*. J.O. Patenaude, Ottawa.

Martinez-Galarraga, J., Tirado, D. A., and González-Val, R. (2015). Market potential and regional economic growth in Spain (1860–1930). *European Review of Economic History*, 19(4):335–358.

Maurer, N. and Yu, C. (2008). What T. R. Took: The Economic Impact of the Panama Canal, 1903–1937. *The Journal of Economic History*, 68(3):686–721.

Maurer, N. and Yu, C. (2010). The big ditch. In *The Big Ditch*. Princeton University Press.

Maurer, S. and Rauch, F. (2020). Economic Geography Aspects of the Panama Canal.

McCullough, D. (2001). *The Path Between the Seas: The Creation of the Panama Canal, 1870-1914*. Simon and Schuster. Google-Books-ID: uVaeIXMxJpQC.

Missiaia, A. (2016). Where do we go from here? Market access and regional development in Italy (1871–1911). *European Review of Economic History*, 20(2):215–241.

of Marine and Fisheries, D. (1922). *Ports and Harbours Directory of Canada*.

of Statistics, D. B. (1940). *Summary of Canal Statistics*.

Pascali, L. (2017). The Wind of Change: Maritime Technology, Trade, and Economic Development. *American Economic Review*, 107(9):2821–2854.

Petrin, A. and Levinsohn, J. (2012). Measuring aggregate productivity growth using plant-level data. *The RAND Journal of Economics*, 43(4):705–725. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1756-2171.12005.

Redding, S. and Venables, A. J. (2004). Economic geography and international inequality. *Journal of International Economics*, 62(1):53–82.

Restuccia, D. and Rogerson, R. (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic dynamics*, 11(4):707–720. Publisher: Elsevier.

Rockwell, A. E. (1971). The Lumber Trade and the Panama Canal, 1921-1940. *The Economic History Review*, 24(3):445–462. Publisher: [Economic History Society, Wiley].

Solow, R. M. (1957). Technical Change and the Aggregate Production Function. *The Review of Economics and Statistics*, 39(3):312–320. Publisher: The MIT Press.

Sotelo, S. (2020). Domestic Trade Frictions and Agriculture. *Journal of Political Economy*, 128(7):2690–2738. Publisher: The University of Chicago Press.

StatisticsCanada (2016). 150 years of immigration in canada.

Umaña-Dajud, C. (2017). Domestic transport costs, Canada, and the Panama Canal.

Urquhart, M. C. and Dales, J. H. (2007). 2. New Estimates of Gross National Product, Canada, 1870-1926: Some Implications for Canadian Development. In *2. New Estimates of Gross National Product, Canada, 1870-1926: Some Implications for Canadian Development*, pages 9–94. University of Chicago Press.

Zárate, R. D. (2020). Spatial Misallocation, Informality, and Transit Improvements: Evidence from Mexico City.

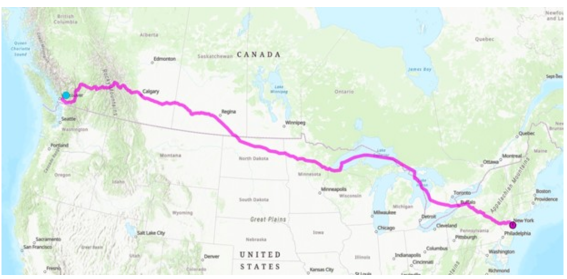
Figures and Tables

Figure 1: Minimum Cost Path Illustrations With and Without the Panama Canal

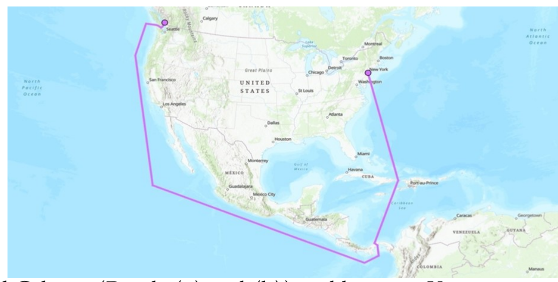
(a) Vancouver, British Columbia, to Calgary, Alberta - Without Canal

(b) Vancouver, British Columbia, to Calgary, Alberta - With Canal

(c) Vancouver, British Columbia, to New York City, NY - Without Canal

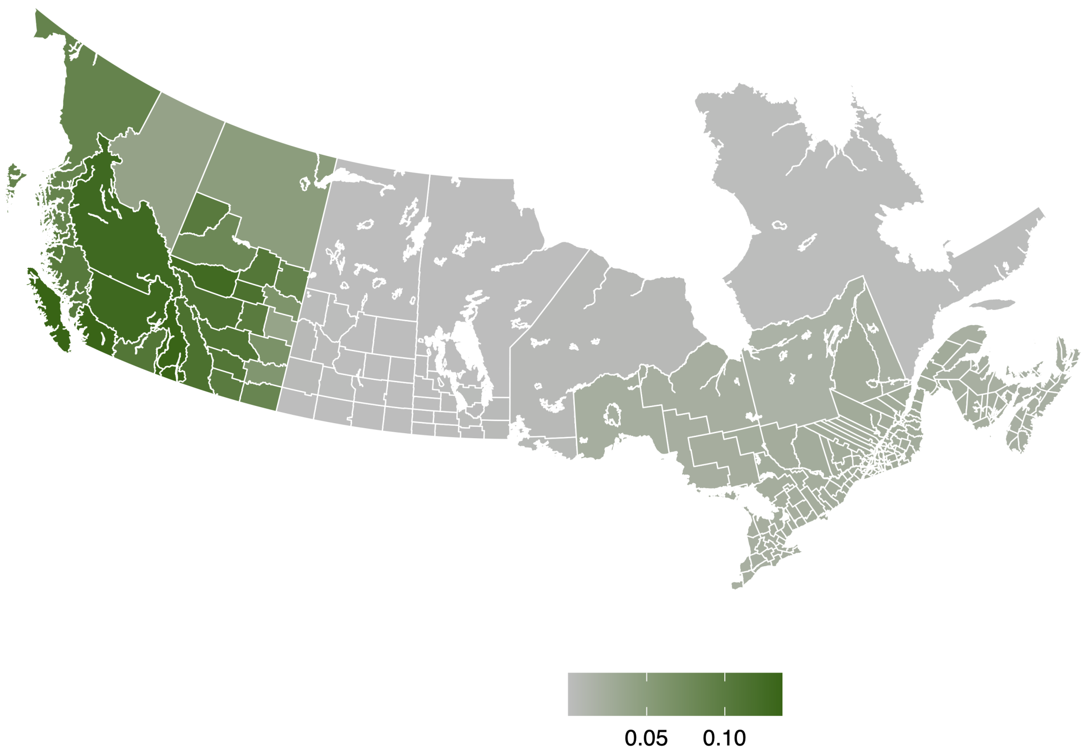


(d) Vancouver, British Columbia, to New York City, NY - With Canal



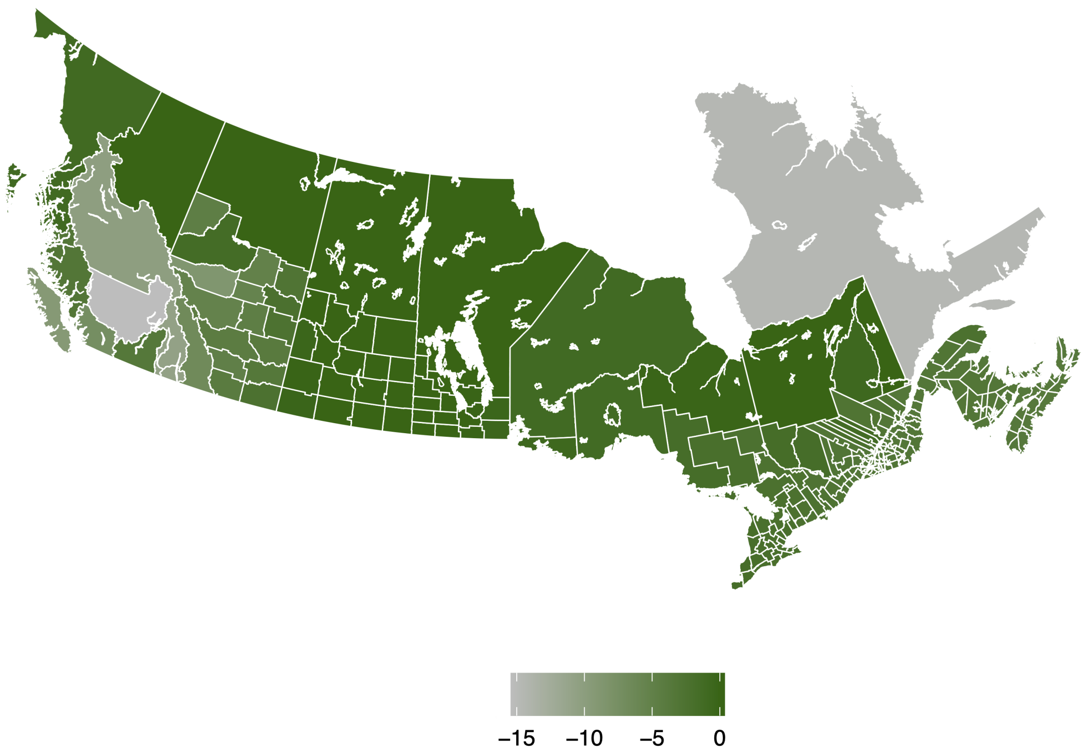
Note: The figure shows the minimum cost path between Vancouver and Calgary (Panels (a) and (b)) and between Vancouver and New York City (Panels (c) and (d)). When the Panama Canal is open, the path changes only between Vancouver and New York City.

Figure 2: Changes in Log Market Access Due to the Opening of the Panama Canal in 1920



Note: The map shows the difference in log market access in 1920 between a scenario where the Panama Canal is open and another where it is closed for all Canadian counties.

Figure 3: Percentage-Point Changes in Population in a Counterfactual Scenario Where the Panama Canal Closes in 1939



The figure shows the change in total population between the actual 1939 data and a counterfactual scenario where the Panama Canal is closed. Details about the counterfactual scenario and the calculations are given in Section 5.

Table 1: Changes in Log Market Access Induced by the Panama Canal

Province	Counties	Average	Std. Dev.	Min	Max
Alberta	17	0.089	0.026	0.039	0.129
British Columbia	10	0.112	0.03	0.04	0.137
Manitoba	16	0.001	0.001	0.000	0.003
New Brunswick	15	0.021	0.001	0.019	0.025
Nova Scotia	18	0.020	0.000	0.019	0.02
Ontario	54	0.022	0.003	0.004	0.024
Quebec	66	0.024	0.003	0.000	0.026
Prince Edward	3	0.019	0.000	0.019	0.02
Saskatchewan	18	0.001	0.001	0.000	0.004
Canada	217	0.028	0.029	0.000	0.137

Note: The table gives descriptive statistics for the difference in market access in 1920 when the Canal is open and closed for all Canadian provinces.

Table 2: Main Results

Manuf. Outcomes: y_c	(1) Value of Products	(2) Capital Exp.	(3) Wage Bill	(4) Employment	(5) Materials	(6) Productivity	(7) TFPR	(8) AE
				Panel A: Δy_c 1939-1910				
$\Delta Ln(MA_{1920})$	0.089**	0.070*	0.097**	0.091***	0.083**	0.131**	0.013	0.118*
	(0.035)	(0.037)	(0.037)	(0.027)	(0.037)	(0.053)	(0.027)	(0.068)
Observations	217	217	217	217	217	217	217	217
r2	0.236	0.248	0.300	0.340	0.243	0.168	0.077	0.233
				Panel B: Pre-Trends Δy_c 1910-1900				
$\Delta Ln(MA_{1920})$	-0.000	-0.003	-0.008	-0.005	0.015	0.005	-0.038	0.042
	(0.029)	(0.028)	(0.028)	(0.027)	(0.031)	(0.050)	(0.023)	(0.042)
Observations	217	217	217	217	217	217	217	217
r2	0.445	0.502	0.520	0.529	0.439	0.137	0.291	0.391

Note: Coefficients are standardized for moving from the 25th to the 75th percentile in terms of gains in market access in 1920, with and without the Canal. All specifications include province fixed effects and control for a quadratic polynomial in latitude and longitude, for a quadratic polynomial in distance to the coast, the Great Lakes or the St. Lawrence River, and for log population (in 1910 for regressions in Panel A, in 1900 for regressions in Panel B). Standard errors clustered at 300km x 300km cells of an arbitrary grid are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Market Access and Railroad Construction

	(1)	(2)	(3)	(4)	(5)	(6)
	Ln(New Km of Railroad built Between...)					
	1900-1915	1900-1920	1905-1915	1905-1920	1910-1915	1910-1920
$\Delta Ln(MA_{1920})$	0.103	0.116	0.141	0.154	0.178*	0.177*
	(0.093)	(0.094)	(0.103)	(0.104)	(0.095)	(0.097)
Observations	217	217	217	217	217	217
Mean Dep. Var.	9.493	9.843	8.688	9.174	6.533	7.225
r2	0.290	0.276	0.362	0.330	0.404	0.330

Note: Coefficients are standardized for moving from the 25th to the 75th percentile in terms of gains in market access in 1920, with and without the Canal. All specifications include province fixed effects and control for a quadratic polynomial in latitude and longitude, for a quadratic polynomial in distance to the coast, the Great Lakes or the St. Lawrence River, and for log population in 1900. Standard errors clustered at 300km x 300km cells of an arbitrary grid are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Robustness to Different MA Measures

Dependent Var: Δy_c	(1) Value of Products	(2) Capital Exp.	(3) Wage Bill	(4) Employment	(5) Materials	(6) Productivity	(7) TFPR	(8) AE
				Panel A: Fixed Population 1910, $\theta = 5$, $\bar{P} = 9.4$				
$\Delta Ln(MA_{1920})$	0.082**	0.064*	0.090***	0.084***	0.077**	0.120**	0.011	0.109*
	(0.031)	(0.034)	(0.033)	(0.024)	(0.033)	(0.048)	(0.025)	(0.062)
Observations	217	217	217	217	217	217	217	217
r2	0.238	0.248	0.302	0.342	0.245	0.169	0.077	0.235
				Panel B: $\theta = -1$, $\bar{P} = 9.4$				
$\Delta Ln(MA_{1920}^{\theta=-1})$	0.121**	0.096	0.136**	0.132***	0.107*	0.198**	0.043	0.155
	(0.055)	(0.058)	(0.057)	(0.042)	(0.058)	(0.082)	(0.047)	(0.100)
Observations	217	217	217	217	217	217	217	217
r2	0.227	0.243	0.290	0.332	0.233	0.166	0.079	0.219
				Panel C: $\theta = -7.5$, $\bar{P} = 9.4$				
$\Delta Ln(MA_{1920}^{\theta=-7.5})$	0.091**	0.068*	0.100**	0.092***	0.085*	0.134**	0.021	0.114
	(0.041)	(0.038)	(0.038)	(0.030)	(0.044)	(0.065)	(0.035)	(0.069)
Observations	217	217	217	217	217	217	217	217
r2	0.238	0.245	0.300	0.340	0.243	0.168	0.078	0.227
				Panel D: $\theta = -5$, $\bar{P} = 5$				
$\Delta Ln(MA_{1920}^{\bar{P}=5})$	0.096**	0.074*	0.106**	0.098***	0.090**	0.142**	0.015	0.127*
	(0.042)	(0.040)	(0.040)	(0.031)	(0.044)	(0.066)	(0.035)	(0.074)
Observations	217	217	217	217	217	217	217	217
r2	0.235	0.246	0.298	0.337	0.242	0.167	0.077	0.230
				Panel E: $\theta = -5$, $\bar{P} = 20$				
$\Delta Ln(MA_{1920}^{\bar{p}=20})$	0.077**	0.063*	0.085**	0.081***	0.071**	0.118**	0.014	0.105*
	(0.030)	(0.033)	(0.033)	(0.024)	(0.032)	(0.046)	(0.023)	(0.060)
Observations	217	217	217	217	217	217	217	217
r2	0.234	0.248	0.298	0.339	0.241	0.169	0.077	0.232
				Panel F: $\theta = -5$, $\bar{P}_c$ calculated for each county				
$\Delta Ln(MA_{1920}^{\bar{P}_c})$	0.200*	0.143	0.219**	0.210**	0.185*	0.310*	0.052	0.257
	(0.104)	(0.106)	(0.103)	(0.078)	(0.110)	(0.158)	(0.083)	(0.178)
Observations	217	217	217	217	217	217	217	217
r2	0.221	0.237	0.280	0.320	0.230	0.157	0.078	0.212

Note: Coefficients are standardized for moving from the 25th to the 75th percentile in terms of gains in market access in 1920, with and without the Canal. Dependent variables for specifications in each column are the change in manufacturing outcomes detailed in the first row between 1910 and 1939. Regressions in each panel differ in terms of the measurement of change in market access that they use. Panel A's measure holds population fixed at 1910 levels. Measures in Panels B to E use different levels of the key parameters described in Section 6. All specifications include province fixed effects and control for a quadratic polynomial in latitude and longitude, for a quadratic polynomial in distance to the coast, the Great Lakes or the St. Lawrence River, and for log population in 1910. Standard errors clustered at 300km x 300km cells of an arbitrary grid are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Robustness Checks: Alternative Hypothesis

Dependent Var: Δy_c	(1) Value of Products	(2) Capital Exp.	(3) Wage Bill	(4) Employment	(5) Materials	(6) Productivity	(7) TFPR	(8) AE
Panel A: Controlling for Baseline MA								
$\Delta Ln(MA_{1920})$	0.084**	0.066*	0.092**	0.086***	0.080**	0.122**	0.009	0.113*
	(0.034)	(0.036)	(0.035)	(0.025)	(0.036)	(0.053)	(0.029)	(0.067)
$Ln(MA_{1910})$	0.150*	0.109	0.170**	0.152**	0.116	0.298**	0.126	0.172*
	(0.089)	(0.085)	(0.072)	(0.061)	(0.089)	(0.135)	(0.114)	(0.092)
Observations	217	217	217	217	217	217	217	217
r2	0.251	0.254	0.320	0.361	0.251	0.191	0.085	0.245
Panel B: Controlling for Extension of Railroad Lines, 1920-1940								
$\Delta Ln(MA_{1920})$	0.091**	0.072*	0.100***	0.093***	0.085**	0.136**	0.013	0.123*
	(0.035)	(0.036)	(0.037)	(0.027)	(0.037)	(0.054)	(0.029)	(0.068)
New Railroad Km,	-0.047	-0.037	-0.061	-0.041	-0.037	-0.107	-0.003	-0.104
1920-1940	(0.092)	(0.089)	(0.079)	(0.068)	(0.098)	(0.179)	(0.087)	(0.152)
Observations	217	217	217	217	217	217	217	217
r2	0.238	0.249	0.302	0.341	0.244	0.171	0.077	0.237
Panel C: Controlling for Extension of Railroad Lines, 1910-1920								
$\Delta Ln(MA_{1920})$	0.089**	0.070*	0.097**	0.090***	0.082**	0.135**	0.020	0.115*
	(0.034)	(0.036)	(0.036)	(0.027)	(0.036)	(0.053)	(0.027)	(0.066)
New Railroad Km,	-0.005	0.003	0.011	0.016	0.025	-0.085	-0.158	0.072
1910-1920	(0.076)	(0.097)	(0.079)	(0.064)	(0.075)	(0.135)	(0.120)	(0.091)
Observations	217	217	217	217	217	217	217	217
r2	0.236	0.248	0.300	0.340	0.244	0.170	0.089	0.235
Panel D: Controlling for Pre-Trend in TFPR								
$\Delta Ln(MA_{1920})$	0.090**	0.071*	0.099***	0.091***	0.088**	0.118*	-0.003	0.121*
	(0.034)	(0.036)	(0.036)	(0.027)	(0.034)	(0.059)	(0.024)	(0.067)
$\Delta TFPR_{1910-1900}$	0.036	0.040	0.055	0.015	0.138	-0.378**	-0.448***	0.070
	(0.095)	(0.100)	(0.083)	(0.071)	(0.090)	(0.168)	(0.155)	(0.160)
Observations	217	217	217	217	217	217	217	217
r2	0.237	0.249	0.302	0.340	0.255	0.205	0.174	0.235
Panel E: Controlling for Distance to Each Coast Separately								
$\Delta Ln(MA_{1920})$	0.089**	0.078**	0.103**	0.094***	0.080**	0.140**	0.011	0.128*
	(0.037)	(0.037)	(0.040)	(0.030)	(0.039)	(0.065)	(0.025)	(0.074)
Observations	217	217	217	217	217	217	217	217
r2	0.245	0.253	0.297	0.341	0.251	0.191	0.107	0.241

Note: Coefficients are standardized for moving from the 25th to the 75th percentile of the independent variables. Dependent variables for specifications in each column are the change in manufacturing outcomes between 1910 and 1939 detailed in the first row. All specifications for Panels A to D include province fixed effects and control for a quadratic polynomial in latitude and longitude, for a quadratic polynomial in distance to the coast, the Great Lakes or the St. Lawrence River, and for log population in 1910. Specifications in Panel E replace the polynomial on distance to the nearest coast to separate polynomials for distance to the Atlantic and Pacific coasts. Standard errors clustered at 300km x 300km cells of an arbitrary grid are shown in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Main Results Panel Data

Dependent Variables: y_c	(1) Value of Products	(2) Capital Exp.	(3) Wage Bill	(4) Employment	(5) Materials	(6) Productivity	(7) TFPR	(8) AE
Panel A: County and Year FE.								
$\Delta Ln(MA_{1920}) \times \ldots$								
... year=1900	-0.001	0.003	0.007	0.004	-0.016	-0.006	0.037	-0.043
	(0.034)	(0.032)	(0.032)	(0.030)	(0.036)	(0.046)	(0.036)	(0.041)
... year=1939	0.088**	0.069*	0.097**	0.090***	0.083**	0.131**	0.013	0.118*
	(0.037)	(0.038)	(0.040)	(0.031)	(0.040)	(0.057)	(0.026)	(0.071)
N	651	651	651	651	651	651	651	651
r2	0.877	0.870	0.883	0.887	0.872	0.981	0.674	0.985
Panel B: County and Year FE + Province Time Trends								
$\Delta Ln(MA_{1920}) \times \ldots$								
... year=1900	-0.001	0.002	0.007	0.004	-0.016	-0.006	0.037	-0.043
	(0.033)	(0.032)	(0.031)	(0.029)	(0.035)	(0.044)	(0.032)	(0.041)
... year=1939	0.088**	0.070*	0.097***	0.090***	0.083**	0.131**	0.013	0.118*
	(0.036)	(0.036)	(0.036)	(0.027)	(0.039)	(0.052)	(0.034)	(0.071)
N	651	651	651	651	651	651	651	651
r2	0.887	0.883	0.896	0.903	0.880	0.982	0.686	0.986
Panel C: County and Year FE + Province x Year FE								
$\Delta Ln(MA_{1920}) \times \ldots$								
... year=1900	-0.000	0.003	0.008	0.004	-0.016	-0.006	0.037*	-0.043
	(0.031)	(0.030)	(0.030)	(0.029)	(0.033)	(0.043)	(0.021)	(0.039)
... year=1939	0.089**	0.070*	0.097***	0.091***	0.083**	0.131**	0.013	0.118*
	(0.037)	(0.037)	(0.037)	(0.028)	(0.041)	(0.054)	(0.026)	(0.068)
N	651	651	651	651	651	651	651	651
r2	0.891	0.888	0.900	0.905	0.886	0.982	0.702	0.987

Note: Coefficients are standardized for moving from the 25th to the 75th percentile in terms of gains in market access in 1920, with and without the Canal. All specifications include county and year fixed effects and control for a quadratic polynomial in latitude and longitude, for a quadratic polynomial in distance to the coast, the Great Lakes or the St. Lawrence River, and for log population; all are interacted with year fixed effects. Specifications in Panel B include province linear trends. Specifications in Panel C include province x year FE. Standard errors clustered at the county level are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Heterogeneity

Dep. Var. Δy_c 1939 - 1910	(1) Value of Products	(2) Capital Exp.	(3) Wage Bill	(4) Employment	(5) Materials	(6) Productivity	(7) TFPR	(8) AE
Panel A: Above and Below Average 1910 Employment/Population Share								
$\Delta Ln(MA_{1920})$	0.059**	0.052	0.073**	0.079***	0.048*	0.061**	0.018	0.044
	(0.026)	(0.035)	(0.033)	(0.026)	(0.026)	(0.027)	(0.022)	(0.032)
$\Delta Ln(MA\_1920) \times$ ... $\mathbb{1}$(Empl./Pop.>Mean)	0.078	0.048	0.065	0.034	0.092	0.180*	-0.010	0.190
	(0.061)	(0.067)	(0.067)	(0.049)	(0.066)	(0.097)	(0.064)	(0.128)
$\mathbb{1}$(Empl./Pop.>Mean)	-0.606***	-0.426*	-0.532***	-0.496***	-0.642***	-0.736**	-0.145	-0.591**
	(0.197)	(0.214)	(0.183)	(0.155)	(0.206)	(0.313)	(0.241)	(0.280)
Counties	217	217	217	217	217	217	217	217
r2	0.277	0.265	0.333	0.377	0.286	0.200	0.079	0.274
Panel B: Above and Below Average Total Population 1910								
$\Delta Ln(MA_{1920})$	0.075*	0.058	0.083**	0.079**	0.070	0.126**	0.014	0.112
	(0.039)	(0.041)	(0.041)	(0.030)	(0.043)	(0.062)	(0.030)	(0.078)
$\Delta Ln(MA_{1920}) \times$ ... $\mathbb{1}$(Pop.>Mean)	0.130	0.117	0.133	0.115	0.125	0.064	-0.003	0.067
	(0.147)	(0.144)	(0.131)	(0.107)	(0.163)	(0.179)	(0.061)	(0.203)
$\mathbb{1}$(Pop.>Mean)	0.066	0.081	0.146	0.157	0.079	-0.126	-0.153	0.027
	(0.285)	(0.268)	(0.250)	(0.209)	(0.296)	(0.448)	(0.344)	(0.367)
Counties	217	217	217	217	217	217	217	217
r2	0.243	0.253	0.308	0.349	0.249	0.169	0.078	0.234
Panel C: Above and Below Average Wood Products Revenue/Value of Products 1910								
$\Delta Ln(MA_{1920})$	0.066*	0.045	0.075**	0.084***	0.055	0.077*	0.035	0.042
	(0.035)	(0.036)	(0.035)	(0.027)	(0.037)	(0.042)	(0.028)	(0.041)
$\Delta Ln(MA_{1920}) \times$ ... $\mathbb{1}$(Wood/Val. Prod.>Mean)	0.072	0.084	0.074	0.025	0.089	0.164	-0.073	0.237*
	(0.074)	(0.059)	(0.075)	(0.063)	(0.079)	(0.130)	(0.069)	(0.137)
$\mathbb{1}$(Wood/Val. Prod.>Mean)	-0.075	-0.243	-0.139	-0.147	-0.078	0.020	0.175	-0.155
	(0.206)	(0.212)	(0.174)	(0.146)	(0.209)	(0.302)	(0.195)	(0.216)
Counties	217	217	217	217	217	217	217	217
r2	0.242	0.259	0.307	0.344	0.251	0.179	0.083	0.269
Panel D: Above and Below Average Food Products Revenue/Total Revenue 1910								
$\Delta Ln(MA_{1920})$	0.078	0.054	0.073	0.067	0.074	0.110	0.051	0.059
	(0.082)	(0.053)	(0.057)	(0.051)	(0.089)	(0.118)	(0.068)	(0.069)
$\Delta Ln(MA_{1920}) \times$ ... $\mathbb{1}$(Food/Val. Prod.>Mean)	0.024	0.029	0.042	0.038	0.027	0.036	-0.064	0.100
	(0.106)	(0.070)	(0.079)	(0.069)	(0.114)	(0.143)	(0.083)	(0.094)
$\mathbb{1}$(Food/Val. Prod.>Mean)	0.172	0.092	0.087	0.026	0.254	0.052	-0.141	0.193
	(0.206)	(0.189)	(0.176)	(0.145)	(0.206)	(0.333)	(0.214)	(0.213)
Counties	217	217	217	217	217	217	217	217
r2	0.241	0.250	0.303	0.342	0.253	0.169	0.082	0.243

Note: Coefficients are standardized for moving from the 25th to the 75th percentile in terms of gains in market access in 1920, with and without the Canal. All specifications include province fixed effects and control for a quadratic polynomial in latitude and longitude, for a quadratic polynomial in distance to the coast, the Great Lakes or the St. Lawrence River, and for log population in 1910. Standard errors clustered at 300km x 300km cells of an arbitrary grid are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Land Values

	(1)	(2)
Dep. Variable:	Log Average Land Values	
year=1901 × $\Delta Ln(MA_{1920})$	0.004	0.003
	(0.016)	(0.009)
year=1911 × $\Delta Ln(MA_{1920})$	0.012	0.012
	(0.016)	(0.007)
year=1941 × $\Delta Ln(MA_{1920})$	0.032**	0.033***
	(0.015)	(0.011)
Observations	868	868
Dep. Var. Mean	3.095	
Pop. x Year FE		x
Coord. Poly.		x
Prov. x Year FE		x
r2	0.832	0.903

Note: Coefficients are standardized for moving from the 25th to the 75th percentile in terms of gains in market access in 1920, with and without the Canal. 1921 is the omitted category. All specifications include county and year fixed effects. Column 2 controls for population in 1911 interacted with year dummies, a quadratic polynomial in latitude and longitude, and a quadratic polynomial in distance to the coast, the Great Lakes or the St. Lawrence River, interacted with year dummies, and province x year fixed effects. Standard errors clustered at 300km x 300km cells of an arbitrary grid are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 9: The Panama Canal Shock in Perspective

Country (Sector)	Scenario	GDP Change	Population Loss (With Fixed Real Wages)	Source
Canada	Close the Panama Canal in 1939	-1.86%	2.7%	Galiani, Jaramillo, and Uribe-Castro (2022)
US	Close the Panama Canal in 1940	-0.2%		Maurer and Rausch (2022)
US	Social savings from the Canal	0.12%		Maurer and Yu (2008)
US (Agriculture)	Remove RR in 1890	-3.2%	58%	Donaldson and Hornbeck (2016)
US (Manufacturing)	Remove RR in 1890	-28%	68%	Hornbeck and Rotemberg (2022)
	Fix RR network in 1880	-2.7%		
Argentina	Remove RR in 1914	-11.8%	8.6%	Fajgelbaum and Redding (2022)
	Pre-steamship freight rates in 1914	-15.6%	12.1%	
Canada	Remove international borders in 1900	32%		Liu and Meissner (2015)
		Social Savings Rates		
Country	Project	Rate		Source
United States	Panama Canal Average: 1921-1937	10.9%		Maurer and Yu (2008)
Canada	Grand Western Railroad	6.1%		Carlos and Lewis (1992)
	Grand Trunk Railroad	2.8%		

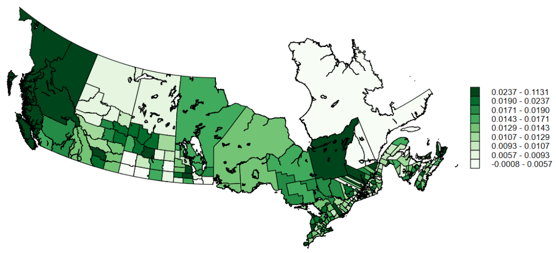
Note: The table compares the results of different counterfactual estimates of the effect of changes in trade costs on GDP and population. Results from different sources may come from models with different fundamentals. Our model is closest to that of Hornbeck and Rotemberg (2021). For the Maurer and Yu (2008) figure, we use their social savings estimate for 1926 (Table 9) and nominal GDP from the NBER Macrohistory Database.

Online Appendixes

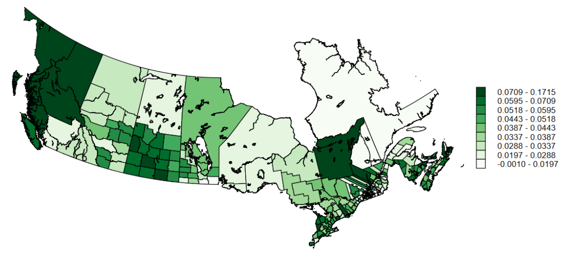
Appendix A Supporting Figures and Tables

Figure A.1: Geographic Distribution of Wedges

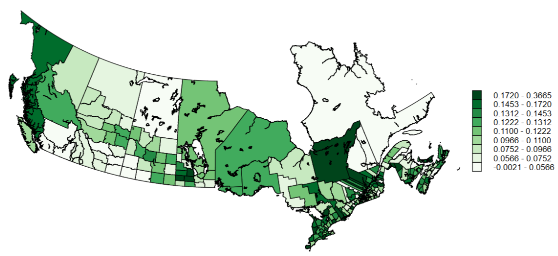
(a) Capital



(b) Labor



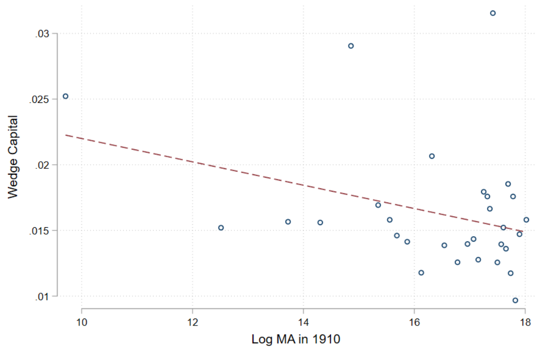
(c) Materials



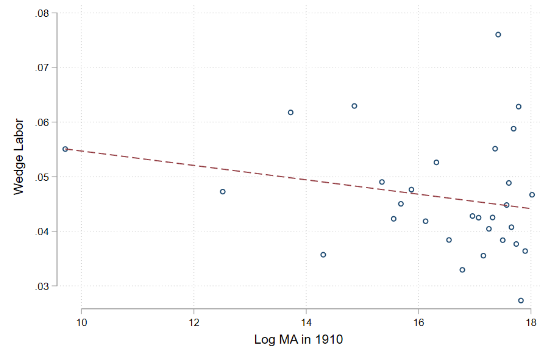
Note: The figures show the geographic distribution of input wedges. For each input and each county, we estimate the difference between the output elasticity ($\alpha_{k,c}$) and the revenue share ($s_{k,c}$). The output elasticity is defined as the cost share of input k in county c. See Section 3 for more details.

Figure A.2: Correlation Between Wedges and Market Access in 1910

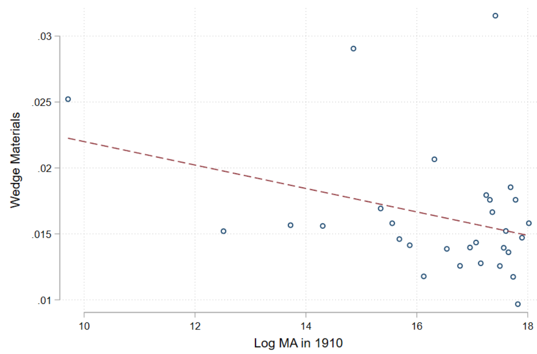
(a) Capital



(b) Labor



(c) Materials



Note: The figures show the binscatter plot for input wedges over the levels of estimated (log) Market Access in 1910. For each input and each county, we estimate wedges as the difference between the output elasticity ($\alpha_{k,c}$) and the revenue share ($s_{k,c}$). The output elasticity is defined as the cost share of input k in county c. See Section 3 for more details.

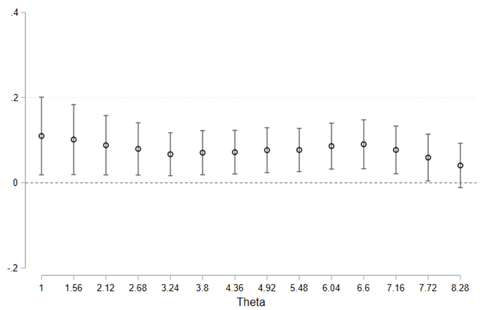
Table A.1: Robustness Checks: $\Delta Ln(MA_{1920})$ coefficients for different specifications

Dependent Var: Δy_c	(1) Value of Products	(2) Capital Exp.	(3) Wage Bill	(4) Employment	(5) Materials	(6) Productivity	(7) TFPR	(8) AE
Main Spec.	0.084**	0.066*	0.092**	0.086***	0.080**	0.122**	0.009	0.113*
	(0.034)	(0.036)	(0.035)	(0.025)	(0.036)	(0.053)	(0.029)	(0.067)
No Prov. FE, No Geo. Controls	0.088*	0.070	0.098*	0.091**	0.083*	0.132*	0.013	0.119
	(0.045)	(0.048)	(0.050)	(0.041)	(0.046)	(0.073)	(0.028)	(0.089)
Geo. Controls	0.088**	0.069*	0.097**	0.090***	0.083**	0.131**	0.013	0.118
	(0.036)	(0.040)	(0.042)	(0.032)	(0.038)	(0.057)	(0.027)	(0.072)
Prov. FE	0.089**	0.070	0.098**	0.091**	0.084*	0.133*	0.013	0.119
	(0.043)	(0.045)	(0.048)	(0.036)	(0.045)	(0.072)	(0.028)	(0.088)

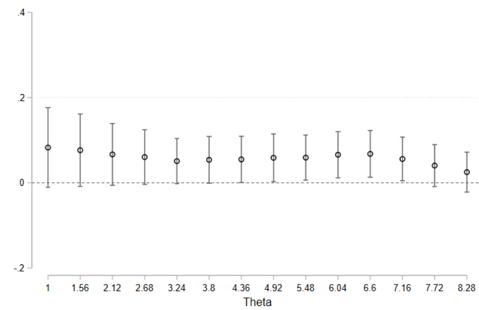
Note: Table shows $\Delta Ln(MA_{1920})$ coefficients for different dependent variables (columns) and specifications (rows). Coefficients are standardized for moving from the 25th to the 75th percentile of the independent variables. Dependent variables for specifications in each column are the change in manufacturing outcomes between 1910 and 1939 detailed in the first row. All specifications control for log total population. Main specification includes 1) province fixed effects and 2) controls for a quadratic polynomial in latitude and longitude, for a quadratic polynomial in distance to the coast, the Great Lakes or the St. Lawrence River. Other specifications are different combinations of 1) and 2). Standard errors clustered at 300km x 300km cells of an arbitrary grid are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure A.3: Robustness to Different Levels of θ for MA Calculations

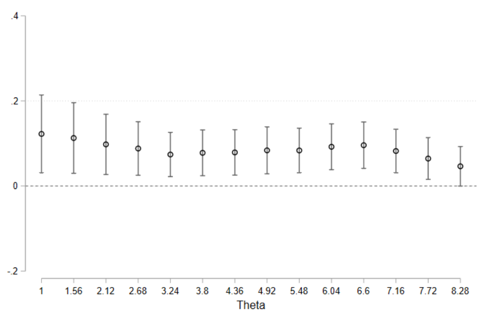
(a) Value of Products



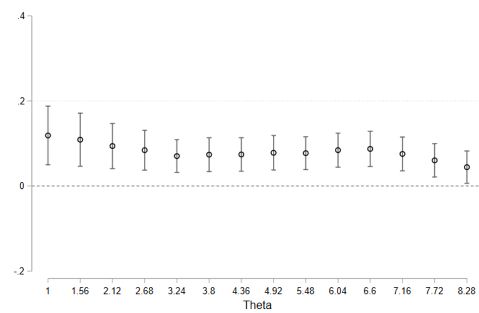
(b) Capital Expenditures



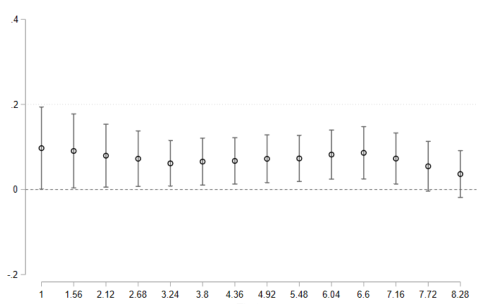
(c) Wage Bill



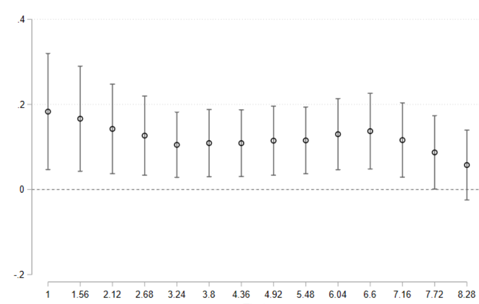
(d) Employment



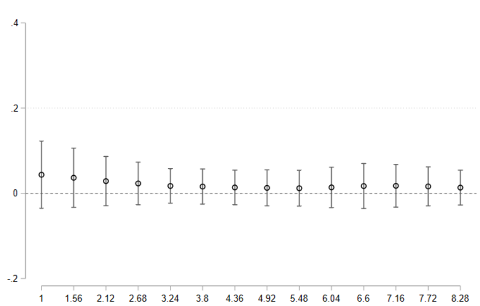
(e) Materials



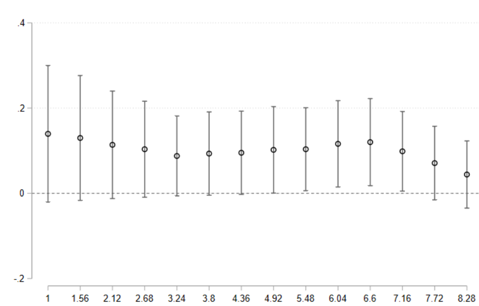
(f) Productivity



(g) TFPR



(h) Reallocative Eff.



Note: Each subfigure plots the β from Equation 9 for different estimates of Market Access gains. We estimate $\Delta ln(MA_c)$ for different levels of θ, the elasticity of trade to trade costs. Each coefficient on each subfigure comes from a separate linear regression estimation.

Table A.2: Main Results: Different Estimates of Standard Errors

Dep Variable: Δy_c	(1) Value of Products	(2) Capital Exp.	(3) Wage Bill	(4) Employment	(5) Materials	(6) Productivity	(7) TFPR	(8) AE
$\Delta Ln(MA_{1920})$	0.089	0.070	0.097	0.091	0.083	0.131	0.013	0.118
Clustering Std. Errors Using Arbitrary Grid of Size:								
300km	(0.035)	(0.037)	(0.037)	(0.027)	(0.037)	(0.053)	(0.027)	(0.068)
150km	(0.039)	(0.037)	(0.038)	(0.029)	(0.043)	(0.057)	(0.026)	(0.070)
200km	(0.036)	(0.036)	(0.035)	(0.028)	(0.040)	(0.054)	(0.026)	(0.068)
Conley Adjusted Std. Errors With Distance Cutoff:								
150km	(0.035)	(0.031)	(0.032)	(0.025)	(0.039)	(0.054)	(0.024)	(0.065)
300km	(0.029)	(0.031)	(0.031)	(0.023)	(0.032)	(0.051)	(0.025)	(0.062)
600km	(0.034)	(0.031)	(0.031)	(0.025)	(0.035)	(0.055)	(0.020)	(0.055)

Note: All specifications include province fixed effects and control for a quadratic polynomial in latitude and longitude, for a quadratic polynomial in distance to the coast, the Great Lakes or the St. Lawrence River, and for log population in 1910.

Table A.3: Sectoral Concentration and Input Wedges

Dep. Var.:	(1) Capital Wedge	(2)	(3) Labor Wedge	(4)	(5) Materials Wedge	(6)
HHI (Revenue)	0.01167		0.00770		0.01195	
	(0.00963)		(0.01343)		(0.02453)	
HHI (Employment)		0.01758**		0.03541**		0.01703
		(0.00847)		(0.01370)		(0.02161)
Observations	217	217	217	217	217	217
Mean Dep. Variable	0.01636	0.01636	0.04632	0.04632	0.11714	0.11714
r2	0.08496	0.10210	0.09793	0.12882	0.09581	0.09670

Note: The dependent variables are the input wedges for capital (columns 1 and 2), labor (columns 3 and 4), and materials (columns 5 and 6). Wedges are defined as the difference between cost share and revenue share, as explained in Section 3. All specifications control for population in 1910. HHI indexes are calculated using industry level data for 1910. Robust standard errors are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: Change in MA induced by the Panama Canal as a Function of 1910 MA

Dep. Var.:	(1)	(2) $\Delta Ln(MA_{1920})$	(3)	(4)
$Ln(MA_{1910})$	0.00069	0.00069	0.00085	0.00077
	(0.00367)	(0.00110)	(0.00113)	(0.00109)
Observations	217	217	217	217
Mean Dep. Var.	0.02165	0.02165	0.02165	0.02165
Province FE		X	X	X
Population 1900			X	
Geo Controls				X
r2	0.00056	0.88963	0.89298	0.91863

Note: The dependent variable is the change in log market access in 1920 induced by the Panama Canal. Coefficients are standardized in terms of standard deviations of the level of log market access in 1910. Geographic controls include a quadratic polynomial in latitude and longitude and a quadratic polynomial in distance to the coast, the Great Lakes or the St. Lawrence River. Standard errors clustered at 300km x 300km cells of an arbitrary grid are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: Robustness Check: Non-Constant Returns to Scale

	(1)	(2)	(3)	(4)	(5)	(6)
	Decreasing Returns to Scale (0.95)			Increasing Returns to Scale (1.05)		
Dep Variable: Δy_c	Productivity	TFPR	Allocative Eff.	Productivity	TFPR	AE
$\Delta Ln(MA_{1920})$	0.364**	-0.039	0.396*	0.364**	0.116*	0.248
	(0.149)	(0.010)	(0.228)	(0.149)	(0.068)	(0.157)
Observations	217	217	217	217	217	217
r2	0.157	0.102	0.230	0.157	0.078	0.207

Note: All specifications include province fixed effects and control for a quadratic polynomial in latitude and longitude, for a quadratic polynomial in distance to the coast, the Great Lakes or the St. Lawrence River, and for log population in 1910.

Table A.6: Manufacturing Subsectors, 1910

Subsector	Pctg. of Manufacturing Employment			Pctg. of Value of Production		
	Average	Std. Dev.	Max	Average	Std. Dev.	Max
Chemical products	0.10	0.92	10.85	0.09	0.67	7.44
Clothing	1.20	3.17	22.19	0.83	2.62	24.33
Electric light and power	0.23	0.64	4.06	0.29	1.01	11.18
Electrical apparatus and supplies	0.05	0.55	7.83	0.03	0.32	3.88
Food and beverage	14.34	18.27	88.97	25.24	21.05	91.33
Iron and steel products	2.23	5.24	45.50	2.18	5.95	52.58
Leather and leather products	0.77	4.03	34.67	1.23	6.79	67.74
Miscellaneous industries	0.18	0.80	7.80	0.15	0.70	6.86
Non-ferrous metal products	0.21	1.85	19.24	0.53	5.22	54.51
Non-metallic mineral products	2.88	6.89	35.09	1.23	3.16	18.99
Paper products	1.00	5.53	43.37	1.10	6.13	46.61
Petroleum and coal	0.04	0.43	4.44	0.05	0.53	5.49
Printing and publishing	1.52	3.45	20.73	0.87	1.94	12.94
Rubber and products	0.03	0.42	6.20	0.01	0.11	1.57
Textiles other than clothing)	0.65	4.04	54.43	0.42	2.91	40.14
Tobacco and tobacco products	0.31	1.71	16.49	0.24	1.29	13.57
Transportation equipment	2.24	6.66	62.11	1.37	3.97	26.71
Wood products	31.48	26.51	94.30	24.03	23.69	97.43
All Other Industries	40.54	26.56	100.00	40.10	26.62	100.00

Note: The data comes from the 1911 Census of Manufactures. The table shows the county-level average share of manufacturing employment and value for different subsectors. For the average county, 40% of employment and manufacturing revenues were classified as "All Other Industries."

Table A.7: Effect of MA Gains on Input Wedges

	(1)	(2)	(3)
Dependent Variable:	$\Delta_{1939-1910}$ Wedge for...		
	K	L	M
$\Delta Ln(MA_{1920})$	-0.001	-0.000	-0.002
	(0.000)	(0.001)	(0.001)
Observations	217	217	217
Mean Dependent Variable	0.011	0.034	0.097
r2	0.083	0.101	0.123

Note: Coefficients are standardized for moving from the 25th to the 75th percentile in terms of gains in market access in 1920, with and without the Canal. Dependent variables are the change between 1910 and 1939 of the wedges for each of the production input, respectively: capital, labor and materials. Wedges are defined as the difference between the output elasticity α and the revenue share of each input. All specifications include county and year fixed effects and control for a quadratic polynomial in latitude and longitude, for a quadratic polynomial in distance to the coast, the Great Lakes or the St. Lawrence River, and for log population. Standard errors clustered at the county level are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix B Data Appendix

Market Access

We compiled information on wheat trade costs by mode of transportation, including rail, wagon, rivers and canals, and international oceanic routes. Additionally, all trips pay a 50-cent fee whenever modes are switched. These follow from Fogel (1964) and the Saskatchewan Grain Commission (1914). Table A.8 summarizes information on routes and rates.

Population

For each Canadian county, we rely on the Canadian Census of Population for 1911 and 1921. For US counties, we use the Population Census for 1910 and 1920 and adjust by GDP per capita relative to Canada. For destinations in the rest of the world, we rely on Maddison Project GDP and population data (Bolt and van Zanden, 2020). We use countries with nonmissing GDP per capita records for 1920 and countries with data for some point between 1910 and 1930 if data for 1920 is missing. As in the case of the US data, for each country we adjust by its GDP per capita relative to Canada in the year for which population data is available in order to adjust for the GDP-per-capita weighted population.

Adjustments to 1911 Canadian Manufacturing Outcomes

We use the Census of Manufactures for 1901, 1911, and 1931. There are three key issues when using the 1911 census. First, the results for 1911 and 1939 are not comparable due to differences in coverage. The 1911 census covered establishments with five or more employees, except in the case of flour mills, saw and shingle mills, lime, brick and tile works, butter and cheese factories, fish-curing plants, and electric light and power plants, which were all surveyed regardless of size. The 1939 census surveyed all establishments, regardless of size. To address this problem, we follow Urquhart and Dales (2007), at least in spirit, by using linear expansion factors (blow-up factors) for the industries that were not fully covered in 1911. These factors are available for each outcome (revenue, capital, employment, wages) at the province level. They are calculated from the 1906 Postal Census, which published information on these outcomes for all establishments and those employing five or more people. This is simply a linear expansion that follows from those two values.

Table A.8: Transportation Network for Trade Cost Calculation.

Mode	Information	Canada	United States
Rail	Routes	GEORIA Project, University of Toronto	Donaldson and Hornbeck (2016)
	Rate	0.514 cents per ton-mile. Method: Simple average of per-ton-mile rates for 17 routes between Fort William, Ontario to 17 destinations in Canada. Standard deviation = 0.04. Source: Railways Commission of Canada (1939)	0.626 cents per ton-mile Method: National average for grains, taken directly from the Interstate Commerce Commission (1913). Source: Interstate Commerce Commission (1913).
Wagon routes	Routes	Straight lines from county centroids to other county centroids, railroad stations, and harbours within 200km of the county centroid.	Donaldson and Hornbeck (2016)
	Rate	25.657 cents per ton-mile Method: Simple average of per-ton-mile rates for wheat of US States that shared a border with Canada. Cross-checked with Saskatchewan Grain Commission (1914). Source: United States Department of Agriculture (1906), Saskatchewan Grain Commission (1914)	22.639 cents per ton-mile Method: National average for wheat. Sources: United States Department of Agriculture (1906)
Waterways, rivers, and canals	Routes	Manually drawn from historical sources including Department of Marine and Fisheries of Canada (1922) and Canada Dominion Bureau of Statistics (1940)	Donaldson and Hornbeck (2016)
	Rate	0.238 cents per ton-mile. Method: Sum of (i) per-ton-mile rate of the Fort William, Ontario – Port Arthur, Ontario (Great Lakes) route, (ii) insurance costs assuming a cargo value of 86 cents per bushel (Great Lakes), and (iii) storage costs for up to six months. Source: Saskatchewan Grain Commission (1914), House of Commons of Canada (1908), and Ward (1994).	0.260 cents per ton-mile. Method: Sum of (i) per-ton-mile rate of the Chicago – New York all-water route, ii) insurance costs assuming a cargo value of 86 cents per bushel (Great Lakes), and (iii) storage costs for up to six months. Source: Interstate Commerce Commission (1913), Saskatchewan Grain Commission (1914), House of Commons of Canada (1908), and Ward (1994).
International oceanic routes	Routes	Donaldson and Hornbeck (2016) and United States Navy (1911, 1917, 1920, 1931, 1943)	
	Rate	0.052 cents per ton-mile + toll (if applicable) If using the Suez Canal: $1.48 toll. If using the Panama Canal: $0.95 toll. Method: Sum of per-ton-mile rate and insurance costs assuming a cargo value of 90 cents per bushel (New York). The per-ton-mile results from the simple regression of (a) historical freight costs between Liverpool and the ports of New York City, Odessa, Karachi, and Buenos Aires retrieved from the Georgian Bay Commission (1916) on (b) distance. Sources: Georgian Bay Commission (1916), Saskatchewan Grain Commission (1914), United States Navy (1911, 1917, 1920, 1931, 1943), and The Panama Canal Company (1971).	

The second issue is that the 1939 census does not include "hand-trades" whereas the 1901 and 1911 censuses did. To address this problem, we explicitly removed industries that were classified as hand-trades from the 1911 census data. Given that many times these industries were classified under "All Other Industries" in a given county, we blew down this particular industry by the proportion that hand-trades represented of the to-

Table A.10: Ports used to build shipping routes passing through the Panama Canal

Ports	
Key West, FL	Providence, RI
Tampa, FL	Boston, MA
Mobile, AL	Portland, ME
New Orleans, LA	Halifax, NS
Port Arthur, TX	Montreal, ON
Pensacola, FL	San Diego, CA
Galveston, TX	Los Angeles, CA
Jacksonville, FL	San Francisco, CA
Charleston, SC	Astoria, OR
Savannah, GA	South Bend, WA
Wilmington, NC	Portland, OR
Alexandria, VA	Port Townsend, WA
Norfolk, VA	Seattle, WA
Newport News, VA	Tacoma, WA
Baltimore, MD	Everett, WA
Philadelphia, PA	Victoria, BC
New York, NY	Vancouver, BC

Source: The Panama Canal Records

tal national values for the outcomes of interest (capital, employment, wages, materials, revenue).

The third and final adjustment is related to fuel and electricity expenditures, which were recorded but not tabulated in 1911. We take the average share of revenues (value of production) that these expenditures represented for 19 Standard Industrial Classification (SIC) industry groups at the national level. To match each industry to an industry group, we use a 1948 SIC classification, which is comparable, according to Urquhart and Dales (2007), to industry classifications in 1911 and 1901. We then calculate total fuel and electricity expenses as the product of total revenues and Urquhart shares for each county/industry. For the "All Other Industries" category, we use national averages for all industries. Finally, we add these expenditures at the county level and include them in the materials expenses to make them comparable to the figures for 1901 and 1939.