Working Paper

The Local Impact of Containerization

Leah Brooks
Trachtenberg School of Public Policy and Public Administration
George Washington University

Nicolas Gendron-Carrier
Department of Economics
University of Toronto

Gisela Rua
Division of Research and Statistics
Board of Governors of the Federal Reserve System

Wednesday 24th February, 2016
Preliminary. Please do not cite without authors' permission.

Thanks to Elliot Anenberg, Nathaniel Baum-Snow, Paul Carrillo, Jonathan Dingel, Marco Gonzalez-Navarro, Jessie Handbury, Peter Morrow, Justin Pierce, Jordan Rappaport, Daniel Trefler, and Matthew Turner for many helpful comments and discussions. We are also very appreciative to research assistants James Calello, Adrian Hamins-Puertolas, Arthi Rabbane, Alex Severn, and Daniel Walker who helped with data entry. We are grateful to Matthew Turner and Gilles Duranton for sharing the 1956 County Business Patterns data. Finally, we thank Nora Brooks, retired reference librarian and Leah's mother, whose combination of digital and pre-digital skills uncovered the 1948 port-level trade data.

The Local Impact of Containerization

Abstract

This paper exploits the advent of containerization, a technological shock that dramatically reduced international shipping costs, to examine how access to international markets affects the spatial organization of economic activity. We contend with the non-random adoption of containerization by employing a novel instrument: being a very deep port in 1953, before containerization. The key idea behind our empirical strategy is that ports have to be deep enough in order to take advantage of containerization because container megaships are much larger than previously used steamships. Analogous to a cost shifter, port depth should affect the supply of ports, but have no effect on the demand for ports. We construct a historical dataset describing the evolution of employment, population, and port facilities for a panel of counties in the United States to estimate the causal impact of containerization on local economic activity. Consistent with the predictions of a standard new economic geography model, we find substantial population and employment growth in U.S. counties near containerized ports.

Leah Brooks
Trachtenberg School of Public Policy
George Washington University
805 21st St. NW, Room 601F
Washington, DC 20006
lfbrooks@gwu.edu

Nicolas Gendron-Carrier
Department of Economics
University of Toronto
150 St. George Street
Toronto, ON, Canada, M5S 3G7
nicolas.gendron.carrier@mail.utoronto.ca

Gisela Rua
Division of Research and Statistics
Board of Governors of the Federal Reserve System
20th and C Streets, Mail Stop 82
Washington, DC 20551
gisela.rua@frb.gov

1 Introduction

Despite a vast and prominent theoretical literature that emphasizes the role of transportation costs in explaining the spatial distribution of economic activity (e.g. Krugman, 1991; Fujita et al., 1999; Allen and Arkolakis, 2014; Redding, 2015), we know very little empirically about the causal effects of trade on local economic growth. In this paper, we use the advent of containerization, a technological shock that dramatically reduced international shipping costs, to examine how access to international markets affects the distribution of economic activity within a country.

Containerization is premised on a simple insight: Packaging goods for waterborne trade into a standardized container makes them cheaper to move. Containerization simplifies packing, transit, pricing, and the transfer from ship to train to truck; it also limits previously frequent and lucrative pilferage. Since the advent of containerization in 1956, international trade has grown tremendously.[1] Bernhofen et al. (2016) estimate that containerization caused international trade to grow by more than 1000 percent in the 15 years since 1966. Containerized trade now dominates ocean shipping, and containers account for well over 75 percent of U.S. domestic rail traffic (Rodrigue, 2015).

Containerization's impact on economic activity is theoretically ambiguous. The new economic geography literature predicts that firms will locate in regions where market potential is high because these regions are the most profitable. A decline in international trade costs can promote local economic growth by allowing firms in the region to reach foreign consumers, thereby attracting new firms to the region and increasing local employment. Improved access to international markets may, however, make a region poorer if increased foreign competition causes firms to depart the region entirely for a lower in-

---

[1] There is some dispute about the magnitude of the cost decline caused by containerization. While Bridgman (2014) and Hummels (2007) argue that direct declines in shipping rates were small, Levinson (2008) argues that declines were large. We address this issue in Section 2.

put cost area, for example China (Autor et al., 2013). In addition, it may increase land and congestion costs which cause firms to relocate. Our empirical work tests whether, in response to this sharp decline in international transportation costs, the agglomerative forces keeping firms and workers together are stronger than those repelling them.

Critical to our investigation is the construction of a historical dataset describing the evolution of population and port facilities for a panel of locations. We focus on the United States, and our unit of analysis is the county. For the period of analysis, 1950 to 2010, we observe population, employment by sector, and factor prices for each county in each year. We combine these data with port level data on location, depth, and size, as well as measures of pre-containerization international trade volumes. We also observe the year of first containerization of any port in every county.

We use these data in both reduced form and general equilibrium approaches. In the reduced form work, we address the non-random selection of counties into proximity to a containerized port with a novel instrument—specifically, proximity to a very deep port in 1953 is an instrument for proximity to a containerized port.

The first requirement for a good instrument is that it is correlated with the endogenous variable. Container ships are substantially larger than their predecessors, and displace more water. They therefore require deep ports. While a port can be arbitrarily deep in the absence of cost concerns and environmental regulations, initially deeper ports are cheaper to convert to container ports because they require less drilling and dredging. This instrument is analogous to the cost shifter instruments used in the industrial organization literature. Empirically, we find a very strong relationship between the instrument and the endogenous variable.

The second requirement is that proximity to a very deep port in 1953 affects county economic activity only through its impact on containerization. Although ports varied in depth before containerization, being a very deep port—beyond 25 or 30 feet—posed no particular competitive advantage. Most ships did not displace enough water to require more depth. Crucially, being a very deep port matters only after the invention and diffusion of containerized shipping. Thus, we parametarize our instrument as a county's proximity to a very deep port, where the depth cut-off is beyond what was generally considered a useful depth in the pre-containerization era.

Our causal estimates of the impact of containerization on local economic growth rely on the quasi-random variation in initial depths. The estimates compare counties that are treated with a container terminal—because they had nearby ports that were very deep before the invention of containerization—to otherwise similar counties.

We find that proximity to a containerized port causes significant population growth. Our most complete instrumental variables estimates indicate that being within 50 km of a container port causes a 30 percent increase in population growth. We also find statistically significant population growth for counties somewhat farther from containerized ports: Population growth in counties 200 to 250 km from containerized ports experience changes equal to 35 percent of the mean growth over the period.

As a robustness check, we repeat the analysis for a panel of world cities. We find that, from 1950 to 2010, cities within 100 km of a containerized port grew about 25 percentage points more than other cities. This is 16 percent of the mean city population growth over the period. Effects are strongest for cities within 100 kilometers of a containerized port, and are marginally statistically significant at distances greater than 200 kilometers.

These reduced form methods help us understand the shift in the distribution of population. However, these methods do not allow us to assess whether population increases near ports come at the expense of other locations. To tackle this general equilibrium proposition, we turn to a market access analysis, as in Donaldson and Hornbeck (forthcoming). We are currently assembling historical maps for this purpose, and anticipate having the results in a future draft.

Our paper is closely related to previous work that examines the role of market access in explaining spatial variation in economic activity (e.g. Davis and Weinstein, 2002; Hanson, 2005; Redding and Sturm, 2008). These authors consider a variety of changes in market access, ranging from the bombing of Japanese cities during the Second World War to the division and reunification of Germany, to test new economic geography predictions. Our paper contributes to this literature by considering directly the effects of a large decline in international transportation costs on local economic growth.[2]

There is also an active academic literature investigating the effect of transportation infrastructure on the growth of cities and regions (e.g. Baum-Snow, 2007; Michaels, 2008; Duranton and Turner, 2012; Donaldson, forthcoming; Gonzalez-Navarro and Turner, 2015). These studies examine how investments in highways, railways, and subways, have shaped the spatial distribution of economic activity.[3] Most existing studies consider infrastructure that reduce domestic transportation costs. Our results contribute to this literature by showing that investments in transportation infrastructure that reduce international transportation costs, such as the construction of new container terminals, can also improve the economic condition of target areas. Methodologically, an important contribution of our paper to this literature is the introduction of a new instrumental variable strategy to contend with the non-random assignment of transport improvements. Specifically, we show how a cost shifter instrument can be used to obtain a source of quasi-random variation in the observed infrastructure.

Finally, our work draws on the large literature concerned with the fundamental determinants of economic growth, pioneered by the work of Barro (1991). A consistent finding in this literature is that landlocked countries are much poorer than other countries. Our

---

[2] Our paper is also related to a growing literature in international trade that looks at the impact of trade on local labour markets (e.g. Topalova, 2010; Autor et al., 2013; Kondo, 2013; Dix-Carneiro and Kovak, 2015). These studies suggest that trade can have substantial localized effects.

[3] See Redding and Turner (2015) for a recent survey of the literature.

results lend credence to the hypothesis that good access to international markets matters for economic growth.[4]

The remainder of this paper is organized as follows. The following section provides background on containerization, and Section 3 discusses the data. We present empirical methods in Section 4, results in Section 5, and robustness checks in Section 6. We conclude with Section 7.

2 Containerization

Containerization brought about an unprecedented change in transportation: It made moving goods across the world dramatically easier, cheaper, and faster. Before the advent of containerization, shipping was expensive and slow. Vessels spent weeks at ports while cargo was handled, piece by piece, by gangs of dockworkers. Containerization was first used in sea and land transportation in 1956, when Malcolm McLean moved 58 truck trailers on a ship from Newark to Houston. This first application of containerization started a revolution that caused a tremendous growth in international trade. For example, Bernhofen et al. (2016) estimate that containerization caused international trade to grow by more than 1000 percent over the 15 years following 1966.

Containerized shipping began in the United States in the mid-1950s, and its diffusion across the country was extremely rapid. Figure 1a reports the total number of container ports per year in the United States. As we can see, the bulk of containerization adoption occurred in the 1960s and 1970s, a smaller number of adoptions happened in the 1980s, and an even smaller number after that. Adoption of containerization was also exceptionally rapid across the world. As shown in Figure 1b and in Rua (2014), the bulk of international adoption occurred in the 1970s, and by 1983, the majority of coun-

---

[4] Romer and Frankel (1999) Feyrer (2009), and Pascali (2014) provide country level evidence of the effects of trade on economic growth.

tries had already adopted containerization. As of 2013, containerized trade accounted for over half of global non-commodity trade (United Nations Conference on Trade and Development, 2013).[5]

Containerization's success relies on two key innovations. The first is the mechanization of container movements—specialized container cranes lift containers in and out of ships, around the port, and onto rail cars and trucks. This simple innovation eased transportation both around and beyond the port, radically changing the entire process of on- and off-loading. With containerization, ships can now spend just a few days at port, while in the breakbulk/pre-containerization era, they spent weeks. And this quick turnaround makes large ships more profitable. Moreover, since containerization greatly reduces the risk of loss and damage, it allows all different kinds of goods, with different destinations, to be shipped together.[6]

The second key innovation is the development of common standards for container size, stacking techniques, and grip mechanisms. These international standards allow a container to be used across modes of transportation—ships, trucks, rail—and across countries. While the U.S. standard for containers was adopted in the early 1960s, the international standard, of the International Organization for Standardization, was promulgated in the late 1960s. These national and international agreements on standards were a successful resolution for a potentially severe collective action problem. (In Figure 1, we note the dates of the U.S. and international standards adoption, as well as when international diffusion plateaus in the early 1980s.)

One of the main benefits of containerization that derive from these two innovations is its impact on ocean shipping costs—both directly and through improvements in the

---

[5] While containers are appropriate for carrying many goods, as diverse as toys and frozen meat, some goods are not yet containerizable. Both "non-dry cargo" and "dry-bulk commodities" such as oil, fertilizers, ore, and grain cannot be shipped inside "the box."

[6] Losses to pilferage have plummeted. Wilson (1982) estimates loses to pilferage at roughly 25% in the breakbulk era, and near zero in the container era.

quality of shipping services. Although a strand of the trade literature argues that containerization did little to lower direct ocean shipping costs, as Hummels (2007) explains, direct shipping costs do not fully capture containerization's impact on the quality of shipping services, particularly delivery times and the risk of pilferage and damage. To the extent that these gains in the quality of shipping services do not show up in traditional measures of shipping costs, they fail to capture containerization's full impact on transportation costs. In fact, in what concerns delivery times alone, Hummels and Schaur (2013) estimate that each day in transit is worth between 0.6 to 2.1 percent of the value of the good, lending credence to the argument that unmeasured benefits of containerized shipping are non-negligible.

For the purposes of this paper, and consistent with the industry definition, we call a port "containerized" when it has special infrastructure and equipment to handle containers. Specifically, the port has invested in equipment to handle shipping containers which enables their movement in and out of ship and onto a train or a truck. Container ports also require extensive marshalling yard where containers in transit can wait to be moved (Rua, 2014).

The conversion from a traditional port to a containerized port is exceedingly expensive. For example, Kendall (1986) writes that "In the period between 1968 and 1973, shipowners, terminal operators, and port agencies in the United States alone invested seven and a half billion dollars in ships, containers, and port facilities." In addition, a survey published by the *The Journal of Commerce and Commercial* in 1977 estimated that acquiring container-handling cranes cost \$1.75 million per 30-ton capacity and securing waterfront space for terminals and marshalling yards cost \$250,000-\$300,000 per acre (Morison, 1977).[7]

There are several reasons why ports decide to containerize, among them, pre-existing

---

[7] See Talley (2002) for containerization's impact on dockworkers.

trade volumes, trade linkages, the industry-mix near the port, and expectations of future gains from increased trade volumes. An additional reason is harbor depth, since it provide an important cost-advantage to adopting containerization. The economies of scale present in containerized shipping imply that ships that carry more volume are more profitable. Since these larger container ships ride deeper in the water, and therefore require deeper harbors than breakbulk ships, ports that are deeper have an advantage with containerization. We use this depth advantage in the reduced-form work to address the non-random adoption of containerization.

3 Data

To study containerization's impact on county growth, we construct a panel with data on population and port facilities for all U.S. counties from 1950 to 2010. This section summarizes the data. We present full details in the data appendix.

Our sample frame is the county level Decennial Census, for the years 1910 to 2010.[8] We assemble a time-invariant panel of counties by aggregating 1950 counties to their 2010 counterparts (most county changes 1950 to 2010 are splits) and by dropping a very few counties with land area changes greater than 35 percent. For the period of analysis, 1950 to 2010, we observe population, employment, share of manufacturing employment, age distribution, income distribution, and education by county in each year.[9] We omit Alaska from our analysis because its administrative districts in 1950 do not correspond to its modern counties. This yields 2,702 counties with complete data, compared to the 2010 total of 3,007 counties.

---

[8] For the 2010 sample, we use the Decennial Census for population figures and the American Community Survey (years 2008–2012) for other demographic covariates.

[9] The share of manufacturing employment comes either from the Decennial Census or from the 1956 County Business Patterns, which we received courtesy of Matt Turner and Gilles Duranton. We hope to expand our use of these data in future drafts, using the 1956 version as our pre-containerization period.

To this sample frame, we add port attribute data. Our universe of ports is all ports that existed in either 1953 or 2015, as defined by the 1953 and 2015 *World Port Index*. For each port, we observe its location (latitude and longitude), size (in 4 discrete categories), and depth (in 8 discrete categories). We use depth of the wharf in 1953 as our measure of pre-containerization port depth.[10] The year of first containerization comes from the *Containerisation International Yearbook*, volumes 1968 and 1970 to 2010. We also observe 1948 and 1955 international trade in dollars by port from the Census Bureau's Foreign Trade Statistics.

We associate each county with a vector of ports and port characteristics. First, we calculate the distance from each county center to each port.[11] Second, for each distance bin $d \in D$, we calculate the number of 1953 ports, the year of first containerization across all ports, the maximum 1953 depth across all ports, and the value of 1948 and 1955 international trade across all ports.[12]

Our dataset covers all continental U.S. land area, and therefore allows for a complete description of containerization's impact on the spatial organization of economic activity. As a robustness check, we also constructed a world sample of international cities. This additional sample has the benefit of describing a larger share of cities affected by changes in international trade. However, since cities must be of sufficient size to enter the sample, it offers a truncated view of containerization's impact on population.

Our sample frame for international cities is the 2014 Revision of World Urbaniza-

---

[10] Results are generally robust to using anchorage and channel depth, which the *World Port Index* also reports.

[11] The county center is the geographic center of the (grouped) county.

[12] Appendix Figure 1 is a graphical depiction of this classification for counties in Southern California. Red triangles mark the geographic center of the county, called a centroid. Containerized ports are denoted with the pink anchor, and the grey circles show rings of 50 kilometers around each port (the Los Angeles and Long Beach ports are next door to one another). Only one county centroid is within the 50 kilometer ring. Thus, at the distance interval of 0 to 50 kilometers, only Orange County is treated with having a containerized port in this range. The bottom panel shows that at a distance of between 50 and 100 km from a port, both San Diego (the southernmost) and Los Angeles are treated counties. Ventura County, just to the north of Los Angeles, remains untreated.

tion Prospects. This dataset contains all 1,692 urban agglomerations with population exceeding 300,000 at any time between 1950 and 2014. By construction, this sample over-represents fast growing cities that were small in 1950 but grew rapidly in the second half of the twentieth century. To mitigate this sampling issue, we restrict the sample to cities with population over 50,000 in 1950, which yields a world panel of 1,051 cities. Our results are robust to different sample selection criteria.[13]

Table 1 presents summary statistics for our main sample of U.S. counties. We report summary statistics for six distance-to-containerized-port bins (columns 1 to 6), and by the categories of ever and never containerized (columns 7 and 8). A county may appear in more than one distance-to-containerized-port bin, but the number of observations in the "ever" and "never" columns sums up to the total sample size.

We observe U.S. population every ten years from 1910 to 2010. From 1910 to 1950—the pre-containerization years—log population in counties near future container ports is larger, and it increases at a faster rate, than in counties farther from future container ports. In addition, the average population among ever-containerized U.S. counties is larger than among counties never near a containerized port. These differences between counties generate a possible bias in the OLS estimation that we address in the empirical section.

The second half of the table uses the Census Bureau's division of the U.S. into four regions. Almost half of the counties within 300 km of container ports are in the Southern region, slightly under one-third are in the Midwest, about one-sixth are in the Northeast, and nearly 1 percent are in the West (note that Western counties tend to be geographically larger). On average, counties near future container ports tend to have substantially more of their employed population in the manufacturing sector than counties never near

---

[13] At this point, for the world analysis we only use 1953 ports that were classified as larger than "very small" in 2015. We are currently entering data to be able to use the sample of all 1953 ports, regardless of size.

future container ports—43 versus 27 percent (columns 7 and 8).

Finally, U.S. counties near container ports have been exposed to containerization for 44 years, on average (column 7). Appendix Table 1 reports total 1948 and 1955 international trade by distance bin to port and shows that counties near future container ports have, on average, more pre-containerization international waterborne trade.

4 Empirical Methods

In this section, we explain our empirical strategy for estimating the causal effect of being near a container port on county population growth. We begin by presenting a naive regression of containerization's impact on population to clearly explain the potential endogeneity issues. We then motivate and explain our instrumental variable strategy. We conclude by discussing how we parameterize a county's proximity to a containerized port.

4.1 First Difference Specification

Our goal is to understand how county population responds to the advent of containerization. Empirically, we measure this in two ways: whether a county is near a port that ever containerizes, and the number of years the county has been near a containerized port. We focus our discussion in this section on the former measure, which is easier to interpret, but we discuss both measures in the estimation results section.

We estimate an equation of the form:

$$\Delta y_{i,t} = \beta_0 + \beta_1 \Delta c_{i,t} + \beta_2 x_{i,1950} + \Delta \epsilon_{i,t} \ , \tag{1}$$

where $i \in I$ are counties, and $t \in T$ are years. Our dependent variable, population, is

$y_{i,t}$. The operator Δ denotes long-run differences, so that $\Delta y_{i,t} = y_{i,t} - y_{i,1950}$. Since there were no container ports in 1950 ($c_{i,1950} = 0 \; \forall i$), our main explanatory variable of interest is whether a county is near a container port at time t, $c_{i,t}$. We also control for baseline covariates in $x_{i,1950}$. Standard errors are clustered at the county level, which is equivalent to having robust standard errors in this two-period case.

To establish the causal effects of containerization on the growth of counties, however, we must contend with the non-random assignment of container ports to counties. For example, if economically healthier counties, which are more likely to take advantage of increased trade, are also more likely to be proximate to container ports, OLS estimates are biased upward. In contrast, if large counties attract proximate container ports, to take advantage of the larger markets those counties offer, our estimates of containerization's impact on county growth would be biased downward. This is because large counties, on average, grow more slowly than smaller counties.

The first difference strategy addresses some of these concerns. The first difference nets out time-invariant county characteristics that may make container terminals more likely to locate near particular counties. For example, equation 1 controls for changes in population due to a county's geographic location, its long-run industry mix, or its climate. This method also accounts for changes in population that impact all counties equally from 1950 to 2010, for example an economic downturn—worldwide or in the U.S.—that might have impacted the likelihood of containerization adoption.

In addition, in a first difference approach, in contrast to a panel fixed effects approach, we can also control for initial conditions, $x_{i,1950}$. Including initial conditions in the first difference model allows for powerful controls, such as differential trends in population growth by initial period covariates. Therefore, we can directly address the concern of differential growth rates by initial county size. We further control for being within 300 km of a 1953 port, the number of 1953 ports within 300 km, and the initial population in

1920 to 1950, the 1956 manufacturing share of employment, and the total value of 1955 international trade.

Nevertheless, these estimates do not allow a distinction between population reallocation or net growth (Redding and Turner, 2015). A positive estimate for β_1 could result from a mix of domestic migration, international migration, or natural population increase. In future work, we hope to be able to dissect some of these differences in the U.S. data.

This empirical strategy would yield a causal estimate for the effect of containerization on population growth if containerization was exogenous, conditional on time-invariant factors at the county level and on the initial covariates that we include. However, suppose that counties are more likely to be near a container port if they made better industrial choices in the 1940s and early 1950s. This is something we may fail to capture, even after netting out county-specific time-invariant factors. This type of endogeneity would yield a positive bias in the OLS estimates. Conversely, if there is more containerization adoption near counties with less successful industrial choices in the 1940s and 1950s, the OLS estimates would be biased downwards.

4.2 Instrumental Variables

To deal with selection bias in the adoption of containerization, we use being near a very deep port in 1953, z_i, as an instrument for containerization $\Delta c_{i,t}$:

$$\Delta c_{i,t} = \alpha_0 + \alpha_1 z_i + \alpha_2 x_{i,1950} + \eta_i \ , \tag{2}$$

There are two requirements for a successful instrumental variable strategy. The first is a strong relationship between containerization and initial depth. The second requirement is that, conditional on covariates, being near a very deep 1953 port is uncorrelated with unobserved determinants of population growth between 1950 and period t. In other words, proximity to a very deep 1953 port affects county i's population growth only through its impact on containerization:

$$\text{Cov}(z_i, \Delta\epsilon_{i,t}) = 0 \tag{3}$$

Conditional on these assumptions, β_1 yields a causal estimate of proximity to a container port on population growth.

We explore the two instrumental variable requirements in turn. First, we anticipate that county proximity to a very deep port pre-containerization should be strongly correlated to county proximity to a container port. Even the first container ships were substantially larger than their predecessors, and larger ships sit deeper in the water and require greater depth to navigate and dock.

Although harbor depth is malleable, it is malleable only at great cost. Given enough money and sufficiently lax environmental regulation, a harbor can arguably be made arbitrarily deep. However, ports which are initially deep have a competitive advantage when technology changes to favor very deep ports. This inability of all ports to adjust equally is confirmed by Broeze, who notes that while "ship designers [keep] turning out larger and larger vessels," and "the engineering limits of port construction and channel deepening have by no means been reached[, t]his, however, may not be said of the capacity of all port authorities to carry the cost of such ventures" Broeze (2002, pp. 175–177). Converting a breakbulk port into a container port is substantially cheaper when the harbor is already deep.

This intuition is borne out in practice by containerization adoption patterns. Figure 2 shows how the likelihood of a county being within 300 km of a containerized port varies over time with proximity (within 300 km) to ports of a given depth. We see a strong relationship between proximity to deep 1953 ports and later proximity to containerized ports. Counties within 300 km of a port deeper than 35 feet are virtually always within 300 km of a container port by the end of the sample period. Only roughly 25 percent of counties within 300 km of ports with depths between 25 and 35 feet are not near a container port by the end of the sample period. For counties near less deep ports, however, containerization is decidedly not a certainty. Indeed, counties that are near only shallow ports—those less than 20 feet deep—are never near a container port.

In Table 2, we show that these differences are statistically meaningful. The specification in the table is not precisely what we will use in our estimation, which relies on a vector of port proximity measures and a vector of depth proximity measures as instruments, but it illustrates clearly the intuition behind our method. We defer discussion of the precise instrument specifications to the end of this section, and results in the following section.

The first column of Table 2 shows that, relative to counties near ports of 1953 maximal depth less than 10 feet, counties near ports over 40-feet deep in 1953 are almost certain to be near containerized ports in 2010. The coefficient declines almost monotonically with depth. Consistent with Figure 2, counties near ports that are less than 20 feet deep in 1953 are very unlikely to be near a container port in 2010; both coefficients are near zero and insignificant. Using the dichotomous specification, counties near ports greater than 30 feet deep in 1953 are 7 percent more likely to be near a container port in 2010. This specification has a robust F statistic of 41.

In the final two columns of the table, the dependent variable is a county's years of exposure to containerization. Counties near ports of depth 40 feet or more experience about 20 years more exposure since first containerization. This coefficient falls monotonically with depth, consistent with our argument about how depth should impact the likelihood of containerization. Using the dichotomous specification, proximity to a 1953 port that is 30 or more feet deep increases a county's exposure to containerization by almost four additional years, or about ten percent of the mean from Table 1.

Given this evidence, which is consistent with a strong relationship between the dependent variable and the instrument, we now turn to the second condition for instrument validity—that proximity to a very deep 1953 port affects county i's population growth only through its impact on containerization. A key concern with the instrument is that port depth may explain county success even before containerization. This is surely true. However, the minimum depth for pre-containerization success was substantially shallower than the minimum depth most useful for containerization. We account for this concern by limiting the depth variation in the instrument to be binary: whether the port is very deep in 1953.

Before containerization, port depth conveyed some advantage, but it was not particularly useful for a port to be very deep. Given the limited draft of breakbulk ships, greater depth was only useful up to a certain point. This is clear even from how data on port depth was collected. The 1953 *World Port Index*'s deepest category is "40 feet and above," while the deepest category in the 2015 *World Port Index* is "76 feet and over."

Our claim that depths beyond 30 feet were not particularly advantageous to port success is supported by a number of contemporary commentators. As late as 1952, F. W. Morgan argues in *Ports and Harbours* that beyond a certain level, depth is not a particularly useful feature of a port:

> The importance for a few ports of maintaining a ruling depth sufficient to admit the largest liners [a draft of 40 feet] emphasizes unduly their importance to the port world. A super-liner which comes into a port every few weeks will, it is true, amplify that port's tonnage figures by half a million tons or so annually. . . . The greater part of world trade by sea and the greater part of the traffic of many ports is concerned with ships of more modest size.

> It would certainly be possible to devise a classification of ports by the draught of ship which can be berthed in them. Halifax and Wellington would appear in the first class, and their ability to berth the largest ships is a great asset in wartime. It tells, however, only a little about their normal significance as ports. (p. 15, Morgan (1952))

Earlier writers also confirm this view. A 1938 monograph argues that "For the ports with which we are dealing, the 30-foot channel at low-water will be taken as the minimum standard in relation to the needs of modern ships" (Sargent, 1938).[14] However, he notes that the cost of making a channel deeper is no small endeavor: "It is a question how far the rest of the world, Europe in particular, is prepared, except in special circumstances, to face the very heavy cost of providing for the needs of the ocean mammoth" (Sargent, 1938, p. 21).

Our instrument is therefore analogous to a cost shifter instrument used in the industrial organization literature. Port depth should affect the supply of ports after the advent of containerization, but have no effect on the demand for ports.

In Section 5, we empirically allay concerns that the instrument is correlated with pre-containerization changes at the county level. To do so, we examine the correlation between pre-containerization factors and the identifying variation in the instrument.

4.3 Parameterization of Distance to Container Port

Until now, we have treated proximity to a containerized port as a uniform category. In practice, our specification allows for different impacts of proximity to a container port on population growth, by distance to a container port. Therefore, we measure change in

---

[14] He goes on to write that in the U.S., a 35-foot draught is becoming standard (p. 21).

access to container ports by:

$$\Delta c_{i,t} \equiv \sum_{d \in D} \beta_{1,d} \mathbb{1}\{\text{Container port between } d_1 \text{ and } d_2 \text{ km}\}_{i,t} , \tag{4}$$

where $d \in D$ is a set of distance bins for county i, and $\{d_1, d_2\}$ are the lower and upper bounds of each bin. We use kilometer bins of $\{0-50, 50-100, 100-150, 150-200, 200-250, 250-300\}$.

This flexible parameterization allows for potentially non-linear effects of distance to the container port on population growth. Our goal is to let the data tell us whether counties need to be very near container ports to experience gains from trade, or whether the specifics of this technology allow for more dispersed growth.

We also propose a set of instrumental variables that parallels this specification:

$$\Delta c_{i,t} = \alpha_0 + \sum_{d \in D} \alpha_{1,d} \mathbb{1}\{\text{Very deep port in 1953 between } d_1 \text{ and } d_2 \text{ km}\}_i + \alpha_2 x_{i,1950} + \eta_i , \tag{5}$$

where $\mathbb{1}\{\text{Very deep port in 1953 between } d_1 \text{ and } d_2 \text{ km}\}_i$ is a dummy variable equal to 1 if the maximum depth of any 1953 port in the bin d_1 to d_2 is greater than 30 feet.

5 Results

We now turn to estimates of the impact of proximity to a containerized port on county population growth. We start with the OLS results, followed by the first-stage instrument results. We then evaluate the instrument's validity before presenting the full two-stage least squares results.

The lefthand panel of Table 3 reports the OLS coefficients for the estimation of equation 1. The first column controls only for the presence of a port in 1953 within 300 kilometers. Results from this specification show that counties nearest to containerized ports have the largest absolute increases in population growth. From 1950 to 2010, their population growth was 31 percentage points higher than in counties never near a container port. This is 30 percent of the average change in population for all counties. Column 2 adds a vector of controls for initial port intensity: the number of 1953 ports in each of the six distance bins. This inclusion increases the coefficient for counties within 50 kilometers of container ports, but leaves the other coefficients relatively unchanged, suggesting that containerization adoption is not strongly related to the presence of many nearby ports in 1953.

To address the issue that counties of different sizes may grow at different rates—and that counties near container ports are larger, as we know from Table 1—the third column adds controls for log population in 1920, 1930, 1940 and 1950. In other words, we allow differential growth rates by initial county size. Here initial size is not just in the first pre-treatment year, but an additional thirty years preceding the treatment. The addition of these controls decreases the estimates of container port proximity on population growth by somewhat less than 50 percent for counties very close (less than 100 km) to container ports, and has a smaller effect on counties farther from container ports.

Finally, as we saw in Table 1, counties near future container ports had, on average, much higher rates of manufacturing employment in 1956. The final column includes this variable as a control. It also attempts to control for pre-containerization port prominence by including, for all ports in each distance bin, total 1955 international trade in millions of dollars. The addition of these covariates has little effect on the coefficients. In this final specification, counties within 50 km of a port that containerizes experience a statistically significant 24 percentage point increase in population growth, which is about one-quarter of the average change in the dependent variable. Counties within 50 to 100 km of a container port experience a statistically significant additional 15 percentage point growth, or a 16 percent increase relative to the mean. Counties 100 to 150 km from a container port experience 13 percent greater growth, and counties 150 to 200 km from a container port experience additional relative growth of 12 percent.

As we remain concerned that additional time varying factors may cause both proximity to containerization and population growth, we turn to our instrumental variable strategy. Table 4 reports estimates from the first stage, using the maximal set of controls from Table 3. We have six endogenous variables—proximity to container ports at the six distance bins—and six instruments, which are the depth of the deepest port in 1953 in each of the six distance bins. We expect that the relationship in this table should be strongest on the diagonal: proximity to a container port at distance d_1 to d_2 should be most correlated with the 1953 depth of the deepest port in that same distance interval.

This is in fact the pattern we see. In panel A, where the dependent variable is proximity to an ever-containerized port, counties in proximity to very deep ports are between 33 and 53 percent more likely to be near a container port at the same distance. The coefficient is smallest at the 250 to 300 distance. All coefficients on the diagonal are strongly significant, and the F statistics for these regressions are never lower than 44; all but five are about 100.

The off-diagonal coefficients in this table are generally negative. This suggests that there is some geographic competition in the location of container ports. Intuitively, we would expect that increasing the number of suitable locations to build a containerized port would decrease the likelihood of containerization of every location. The negative coefficients off-diagonal are consistent with this.

The bottom panel of the table uses years of proximity to the first container port by distance bin as the dependent variable. In this specification, counties located within 0 to 50 km of a very deep 1953 port experience an additional 26 years of proximity to a container port. The instrument explains roughly half of the 41 average years of proximity in this distance bin. Coefficients on the other bins are slightly larger, in the 28 to 32 year range, save the final coefficient for the 250 to 300 km bin, which is 19 years. All coefficients are significant at the 1 percent level or above, and the F statistics for each estimation are never lower than 92.

To further test the validity of the instruments, we evaluate whether they are correlated with county-level characteristics that might plausibly be in the error term. While we cannot do this for all potential confounders, we can observe whether the identifying variation—the residual from a regression of the instrument on the full set of covariates—is correlated with specific pre-treatment covariates. Figure 3 uses the instrument $\mathbb{1}\{\text{Very deep port in 1953 between 0 and 50 km}\}_i$.

Our regression specification controls for log of population density in 1920, 1930, 1940 and 1950. Were the identifying variation in the instrument to be related to the log of 1910 population density, this would suggest that the pre-treatment controls were not adequately capturing the historical pattern of population growth. We do not find this to be the case. Figure 3a shows the identifying variation from the instrument on the y axis, and log of 1910 population density on the horizontal axis. We find no significant relationship ($t = 0.55$) between these two variables.

Similarly, recall that the regression controls for the 1955 value of international trade at a set of distances from each county. If this covariate did not sufficiently control for the impact of pre-containerization port strength on population, we would expect that the identifying variation would be related to the 1948 value of international trade at a set of distances from each county. Figure 3b shows that this is not the case. Again, the relationship between the identifying variation and the variable of concern is insignificant ($t = -0.95$).

In fact, we do twelve estimations: a regression of the identifying variation from each of the six distance bins with 1910 population, and with the dollar value of 1948 international trade at ports in that distance interval. In these 12 regressions, we find one significant relationship. This one significant coefficient is almost what we would expect only by random chance.

Having allayed concerns about instrument strength and validity, we turn to the right-hand panel of Table 3, which shows results from the instrumental variable estimation. The pattern of coefficient change as we add covariates is very similar between the IV and OLS estimates. One notable difference is that the IV results show significant effects at the 250 to 300 km distance range, while the OLS results do not. Counties at this distance from a port have statistically significantly more population growth. In the final column, counties closest to ports (0 to 50 km) have 40 percentage points (and 40 percent) more population growth. Counties 100 to 150 km of a port see no statistically significant difference in growth and then counties 150 to 200 km see an additional 17 percentage points of population growth, or 27 percent of the average change over the period. Counties located from 200 to 250 kilometers of a port see an additional 28 percentage points of population growth, or a 49 percent increase relative to the mean.

This U-shaped pattern of the coefficients with respect to distance from treatment is somewhat surprising. From a standard new economic geography model (e.g. Redding, 2015), we would expect the effect of containerization to be largest for counties closer to containerized ports. However, the non-linear geographic effect of treatment may be explained by the intermodal nature of containers. In the United States, more than anywhere else in the world, containerization impacted the entire transit network. As an intermodal system, firms and consumers can reap the benefits of better access to international markets without necessarily being very close to a container port. As such, the counties that gained the most from the new shipping technology are those near the eight percent of U.S. ports that containerize, as well as counties farther from container ports with more to gain from the new shipping technology in terms of access to international markets. That the treatment effect follows a non-linear pattern could also reflect spillovers across space if people migrate to grab some of the rent associated with large infrastructure investments concentrated near container terminals. To the extent that people are more likely to migrate to nearby destinations, we would expect treatment effects to be non-linear.

In general, the IV results here are larger than the OLS ones. Why is this the case? Suppose that already populous counties grow at a slower rate than less populous counties, and that these larger counties are more likely to be near container ports. When we correct for this endogeneity with the instrument, in principle giving more weight to smaller counties—where proximity to container ports is more likely to be driven by the supply constraint posed by depth, the coefficient should increase.

The specification in Table 3 measures proximity to containerized ports with a dummy variable. While this has the benefit of being easy to interpret, it reports an average across counties near a container port for very few years and counties near a container port for many years. If counties near early-adopting container ports have different population growth trajectories than counties near later-adopting container ports, the average results could be quite misleading. For example, if there are many later-adopting container ports, the average effect could be quite small, even while some counties experience large effects. Alternatively, if proximity to containerization requires a certain number of years to achieve population growth, or stops after a number of years, the average could again be misleading.

To address this concern, Table 5 reports results from re-estimating the specifications in Table 3, using years of proximity to a containerized port by distance bin as the dependent variable.

Concentrating on column 8, our most complete instrumental variables estimate, being a county 0 to 50 km from a container port is associated with a (0.007*41) 29 percentage point increase in population growth, which for this distance category is also a roughly 30 percent increase relative to the mean. Counties 150 to 200 km experience (0.003*39/0.63) 19 percent more population growth, and counties 200 to 250 km experience an additional (0.005*40/0.57) 35 percent growth, relative to the mean in each distance bin.

Comparing these results to those using the proximity measured dichotomously, magnitudes are generally somewhat smaller. This suggests that counties that have been exposed to container ports longer have smaller (per-year) impacts. If larger counties are more likely to be near ports that adopt container technology early, and larger counties grow more slowly, we would expect smaller coefficients. It is also possible that, over time, there are decreasing marginal returns from proximity to a container port, so that beyond a certain point, additional years do not contribute to additional population growth.

6 Robustness Checks

6.1 Sub-samples

We conclude our analysis by investigating potential concerns with our identification. One concern is our use of all counties for counterfactual population growth. Perhaps port counties—specifically, counties near 1953 ports—grow differently than all counties, conceivably because they are all more affected by industries that specialize in waterborne trade. Under this assumption, other port counties could be the better counterfactual. Nonetheless, one could equally argue that the set of all counties provides a better counterfactual, particularly in countries with few very large port counties for which the other port counties are a poor counterfactual. To test this contention, we re-estimate the results using only counties near 1953 ports. The first column in Table 6 limits the sample to these 1,296 counties only. Estimates in this specification are generally smaller and less significant than those in Column 8 of Table 3. This suggests that U.S. counties near 1953 ports grow faster, on average, than counties far from 1953 ports. Despite the decline in magnitude, the pattern of larger growth changes when the county is either quite close to a container port (0 to 50 km) or not particularly close (200 to 250 km) holds.[15]

To explore how much our results are influenced by the correlation between county population and proximity to container ports, we split the sample into counties that in 1950 were above and below median population. As the sample size declines, we lose the ability to obtain precise estimates at many distance intervals. However, a comparison of the coefficients for counties closest to ports (0 to 50 km, comparing columns 2 and 3) shows that counties that had below median population in 1950 grew much more quickly after the advent of containerization. In counties with 1950 population above the median, growth is primarily associated with being not particularly proximate (200 to 250 km) to a container port. It may be that these more populous counties cannot, or are too expensive, to house a container port—which requires substantial land area—in the direct vicinity, but still benefit from the trade deriving from the port.

The final panel of the table repeats these estimates using years of proximity to a container port as the endogenous variable. We find a very similar geographic pattern to the specification using the proximity to an ever-containerized port, and the coefficients yield changes of very similar magnitudes. This suggests that the ever-containerized results are not distorted by unusual patterns of population growth in response to additional years of containerization.

6.2 Sample of World Cities

In this subsection, we use a panel of world cities to explore the robustness of our results to an alternative sampling frame. Table 7 reports summary statistics for the worldwide

---

[15] In Appendix Table 3 we show that the instrument remains strong.

panel, in a format similar to that of Table 1 for the U.S. sample. As we can see in columns 7 and 8 in the top portion of the table, in 1950, cities near future containerized ports are roughly 30 percent larger than cities never near a containerized port. Not surprisingly, given that the U.S. sample covers the entire country, the average population among ever-containerized world cities is almost 2 log points (or 200,000 people) larger than among ever-containerized U.S. counties. Cities closer to future containerized ports are also larger than those farther from containerized ports: 1950 log population declines from column 1 to column 6, with only small exceptions. Comparing 1950 log population to 2010 log population, cities farther from container ports (column 6) have, on average, smaller population increases than cities closer to container ports (column 1).

The second section of the table shows that about fifty percent of cities in the sample are in Asia, roughly one-fifth are in Europe, a slightly smaller fraction are in North America, and the remainder of cities are split between South America, Australia, and Africa. Also, the average container port has existed for 35 years (column 7); unsurprisingly, the average world city has had shorter exposure to containerization than the average U.S. county (44 years).

Figure 5 shows that the containerization adoption patterns in the world sample follow the same strong relationship between proximity to deep 1953 ports and later proximity to containerized ports as in the U.S. sample. Cities are more likely to be near a container port earlier when they are near a deep 1953 port. By 1975, virtually all cities within 300 km of a port that was greater than 40-feet deep in 1953 are near a container port (blue line). Adoption is substantially slower for international cities near ports that are less than 20-feet deep, though roughly eighty percent of these cities are near a containerized port by the end of the sample period.

Table 8 shows that these differences are statistically meaningful. As in the U.S. sample, cities near ports of 1953 maximal depth 40 feet or above are more than fifty percent more likely to be near a container port in 2010, relative to cities near ports of 1953 maximal depth less than 10 feet (column 1). This likelihood is never less than fifty percent for cities near ports greater than 30 feet deep in 1953, and it is always significant at the five percent level. Proximity to a very deep port (above 30 feet) makes a city 10 percent more likely to be within 300 kilometers of a container port (column 2). The F statistic for this estimation is 10, indicating a reasonably strong relationship between the instrument and the endogenous variable. Consistent with Figure 5, we see the strong pattern of cities near deeper ports having had longer exposure to containerization (columns 3 and 4). Relative to cities near ports less-than-10 feet deep, cities near ports 40 or more feet deep have had containerization for 18 more years. In addition, cities near ports that are 30 or more feet deep experience an additional 9 and a half years of exposure to containerization. The F statistic for this specification is a very robust 55.

Table 9 presents OLS and IV results for the world sample. Here, we measure city proximity to container ports in three bins: 0 to 100 km, 100 to 200 km, and 200 to 300 km. Column 1 controls only for whether the city is near (within 300 km) of a port in 1953, which allows for differential population growth trends for cities initially near and far from ports. Column 2 additionally controls for the number of ports within 300 km of the city in 1953, to measure port intensity, and column 3 adds country fixed effects, which allow for different trends in population growth by country. To address the concern that population growth is a function of initial size, Column 4 adds a control for 1950 log population; this specification allows comparisons of population growth between cities of similar initial sizes. Focusing on the OLS estimates in Column 4, cities located 0 to 100 km from a container port experience a statistically significant 14.54 percentage points greater population growth, which is about 9 percent of the average city growth over the period. Estimates for larger distances are smaller and not statistically significant. Comparing with the U.S. sample, these estimates are somewhat smaller in magnitude than the OLS estimates in Table 3.

Similar to what we did for the U.S. sample, we address endogeneity concerns by instrumenting the three containerization proximity measures with three measures of proximity to a port of depth greater than 30 feet, using the same distance bins. We present the full first-stage estimates in Table 10. In the world sample, the instrument also works as we hypothesized, that is, we see strong and significant results along the diagonals of the two panels of the table–in the first panel, the dependent variable is proximity to a containerized port by distance bin, and in the second panel, the dependent variable is a city's years since first containerization by distance bin. In all columns, the F statistic is quite high, and is never less than 62.

Moving to the two-stage least squares results in Table 9, in our most complete specification (column 8), cities within 100 km of a container port show a precisely estimated increase in population growth of about 25 percentage points, or about 16 percent of the mean. For cities within 100 to 200 km and within 200 to 300 km of a container port, we estimate a population growth increase of 23 percentage points. As with the U.S. sample, the IV results here are larger than the OLS ones.

We also estimated the impact of each additional year since first containerization—measured by years of proximity to a container port—for the world sample. The final IV column (column 8) suggests that cities within 100 km and 100 to 200 km of a container port grow (0.0067*32.25) 21.6 percentage points and (0.0063*30.5) 19.2 percentage points faster. These imply growth 14 percent faster than the average city. Similar to the U.S. sample, these estimates are somewhat smaller than those using the dichotomous measure; we hypothesize that the same factors that drove the U.S. sample estimates using years of proximity to a container port to be smaller are also at play here.

Table 12 addresses potential identification concerns. As with the U.S. sample, we limit the world sample to only port cities. Each column pair in this table shows results for the full sample of cities followed by the sample of port cities only. In all specifications, the results in the paired columns are extremely similar, and are not qualitatively differentiable from one another.

An additional concern with the world sample is that the rule for entry into the sample—a population greater than 300,000 at any point from 1950 to 2014—biases the sample toward faster growing cities, and could lead us to overestimate containerization's impact on urban growth. We address this concern by limiting the sample to cities that have 300,000 people or more in all years of the sample. This is, in effect, a sample without any selection biases due to growth rates. It is a more limited sample of cities, but one without interpretation concerns.

Regardless of whether we use all cities in this range (303 cities, column 3) or only port cities (213 cities, column 4), we still find that cities near ports have faster population growth after the advent of containerization. However, for the larger cities, we see that growth is primarily associated with being only somewhat close to containerized ports. For these larger cities, we estimate that being 100 to 200 km from a port causes roughly 41 percentage points more growth, or a 40 percent increase relative to the mean. As with the U.S. sample, we speculate that this change in distance pattern may be due to initially larger cities having less room to accommodate the requirements of a container port—which are substantial in terms of land area—in their immediate vicinity.

The final two columns of the table, which repeat the specification using a city's years of proximity to a container port by distance bin, show similar results for both port cities and the full sample.[16]

Overall, the world sample yields results that are similar to those using our main sample of all U.S. counties: proximity to a container port, at certain distances, is associ-

---

[16] Appendix Table 4 shows the first-stage results using the sample of port cities only; the instrument remains strong.

ated with population growth above the average. However, the estimates for the world sample are somewhat smaller in magnitude than for the U.S. sample, and differ in their geographic pattern. There are multiple possible reasons for larger results in the U.S. First, the U.S. was the first country in the world to adopt container shipping; the average U.S. county proximate to a container port has experienced ten more years since first containerization than the average world city. In addition, U.S. container trade was initially primarily domestic, which may not be true of other countries, and which may yield additional population growth impacts further inland. Finally, population in U.S. counties is substantially smaller than the average population in the world cities data. If smaller cities grow more rapidly, this could account for part of the difference.

The geographic pattern of city proximity to container ports and urban growth also differs in the world and U.S. samples. This comparison is difficult to make with precision, however, given the difference distance bands we have used in the two analyses. We defer analysis of this difference for a future draft.

7 Conclusion

We construct a historical dataset describing the evolution of employment, population, and port facilities for a panel of counties in the United States to estimate the causal impact of containerization on local economic activity. This technology not only transformed global trade, but had substantial local consequences. Consistent with the predictions of a standard new economic geography model, we find substantial increases in population for counties closest and farthest to container ports, and still sizeable increases for counties at a middle distance from container ports.

In future work, we would like to push the reduced form methods to probe the drivers of population change, and containerization's impact on other local variables. However, the reduced form methods do not allow us to assess whether population increases near ports comes at the expense of other locations. To tackle this general equilibrium proposition, we plan to turn to a market access analysis, as in Donaldson and Hornbeck (forthcoming). We are currently assembling historical maps for this purpose, and anticipate having the results in a future draft.

References

Allen, Treb and Arkolakis, Costas, 2014. "Trade and the Topography of the Spatial Economy." *Quarterly Journal of Economics* 129(3): 1085–1140.

Autor, David H., Dorn, David, and Hanson, Gordon H., 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review* 103(6): 2121–68.

Barro, Robert J, 1991. "Economic Growth in a Cross Section of Countries." *The Quarterly Journal of Economics* 106(2): 407–43.

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

Bernhofen, Daniel, El-Sahli, Zouheir, and Kneller, Richard, 2016. "Estimating the Effects of the Container Revolution on World Trade." *Journal of International Economics* 98: 36–50.

Bridgman, Benjamin, 2014. "Why Containerization Did Not Reduce Ocean Trade Shipping Costs." Working Paper.

Broeze, Frank, 2002. *The Globalisation of the Oceans: Containerisation From the 1950s to the Present*, vol. 23 of *Research in Maritime History*. St. John's, Newfoundland: International Maritime Economic History Association.

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

Dix-Carneiro, Rafael and Kovak, Brian, 2015. "Trade Reform and Regional Dynamics: Evidence From 25 Years of Brazilian Matched Employer-Employee Data." Working paper.

Donaldson, Dave, forthcoming. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review* .

Donaldson, Dave and Hornbeck, Richard, forthcoming. "Railroads and American Economic Growth: A Market Access Approach." *Quarterly Journal of Economics* .

Duranton, Gilles, Morrow, Peter M., and Turner, Matthew A., 2014. "Roads and Trade: Evidence from the US." *Review of Economic Studies* 81: 681–724.

Duranton, Gilles and Turner, Matthew, 2012. "Urban Growth and Transportation." *Review of Economic Studies* 79: 1407–1440.

Feyrer, James, 2009. "Trade and Income – Exploiting Time Series in Geography." NBER Working Papers 14910, National Bureau of Economic Research, Inc.

Fujita, Masahisa, Krugman, Paul, and Venables, Anthony J., 1999. *The Spatial Economy: "Cities, Regions and International Trade"*. The MIT Press.

Gonzalez-Navarro, Marco and Turner, Matthew, 2015. "Subways and Urban Growth: Evidence from Earth." Working paper.

Hanson, Gordon, 2005. "Market Potential, Increasing Returns, and Geographic Concentration." *Journal of International Economics* 67: 1–24.

Hummels, David, 2007. "Transportation Costs and International Trade in the Second Era of Globalization." *Journal of Economic Perspectives* 21(3): 131–154.

Hummels, David and Schaur, Georg, 2013. "Time as a Trade Barrier." *American Economic Review* 103(7): 2935–2959.

Kendall, L.C., 1986. *The Business of Shipping*. Cornell Maritime Press, fifth edn.

Kondo, Illenin O., 2013. "Trade Reforms, Foreign Competition, and Labor Market Adjustments in the U.S." International Finance Discussion Papers 1095, Board of Governors of the Federal Reserve System (U.S.).

Krugman, Paul, 1991. "Increasing Returns and Economic Geography." *Journal of Political Economy* 99(3): 483–99.

Levinson, Mark, 2008. *The Box: How the Shipping Container Made the World Smaller and the World Economy Bigger*. Princeton, NJ: Princeton University Press.

Michaels, Guy, 2008. "The effect of trade on the demand for skill: Evidence from the Interstate Highway System." *Review of Economics and Statistics* 09: 683–701.

Morgan, F. W., 1952. *Ports and Harbours*. London, UK: Hutchison's University Library.

Morison, Robert F., 1977. "Containerization: Good and Bad News." *The Journal of Commerce and Commercial* pages 1A, 15A.

National Geospatial Intelligence Agency, 1953. "World Port Index, 1953." H.o. pub. no. 950, first edition.

National Geospatial Intelligence Agency, 2015. "World Port Index, 2015." Twenty-fourth edition.

Pascali, Luigi, 2014. "The Wind of Change: Maritime Technology, Trade and Economic Development." Working Paper 165, CAGE.

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

Redding, Stephen J., 2015. "Goods Trade, Factor Mobility and Welfare." NBER Working Papers 18008, National Bureau of Economic Research, Inc.

Redding, Stephen J. and Turner, Matthew A., 2015. "Transportation Costs and the Spatial Organization of Economic Activity." In Gilles Duranton and William Strange, (Eds.) "Handbook of Regional and Urban Economics," Elsevier B. V.

Rodrigue, Jean-Paul, 2015.

Romer, David H. and Frankel, Jeffrey A., 1999. "Does Trade Cause Growth?" *American Economic Review* 89(3): 379–399.

Rua, Gisela, 2014. "Diffusion of Containerization." Finance and Economics Discussion Series 88, Board of Governors of the Federal Reserve System (U.S.).

Sargent, A. J., 1938. *Seaports and Hinterlands*. London, UK: Adam and Charles Black.

Talley, Wayne K., 2002. "Dockworker Earnings, Containerisation, and Shipping Deregulation." *Journal of Transport Economics and Policy* 36(3): 447–467.

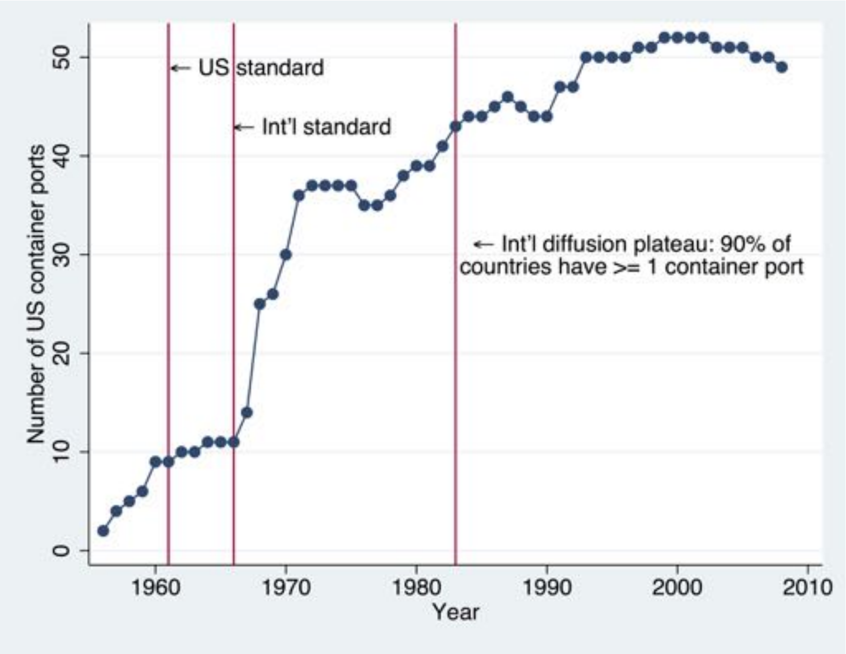
Topalova, Petia, 2010. "Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India." *American Economic Journal: Applied Economics* 2(4): 1–41.

United Nations Conference on Trade and Development, 2013. *Review of Maritime Transport*. New York, NY: United Nations.

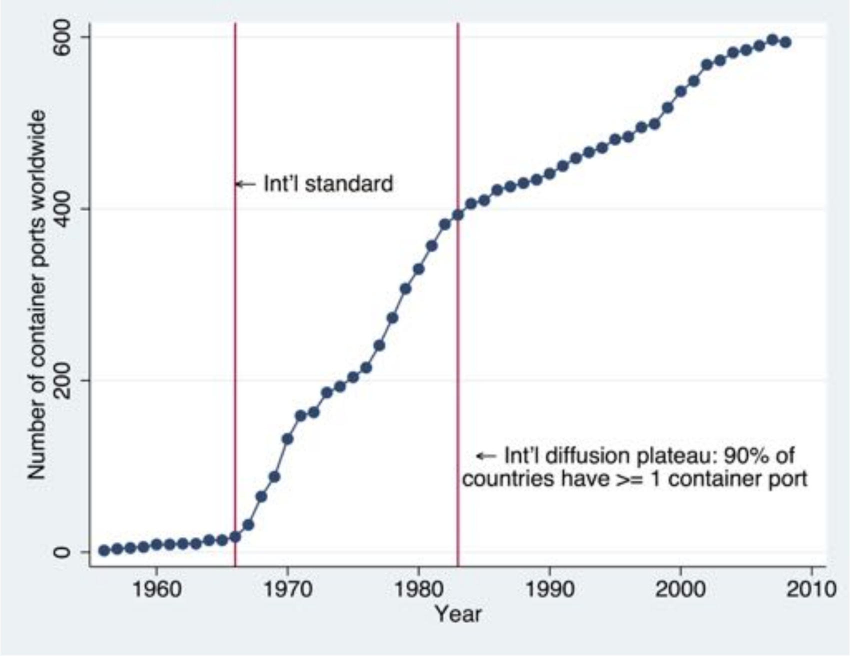
Wilson, R. P., 1982. "The Containerization of Ports." UC Santa Cruz Working Paper.

Figure 1: Adoption of Containerization: 1956–2008

(a) United States

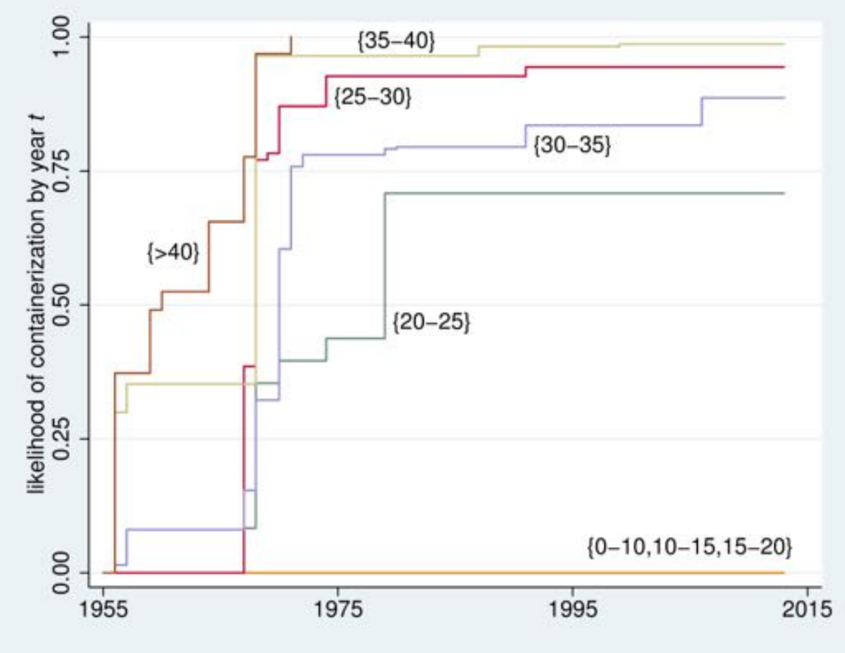


(b) Worldwide



Source: *Containerisation International Yearbook*, volumes 1968 and 1970–2010.

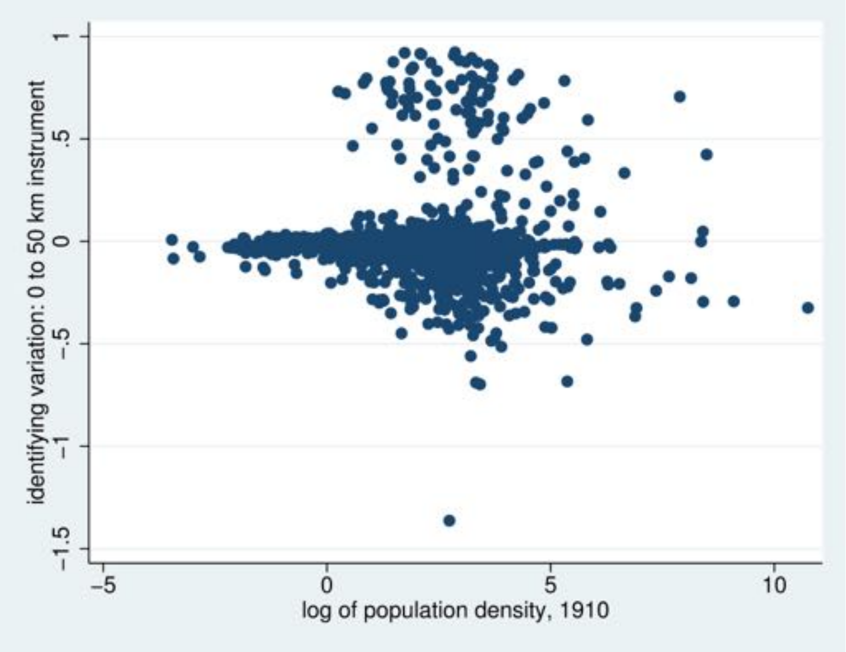
Figure 2: Likelihood of Having a Containerized Port by 1953 Port Depth



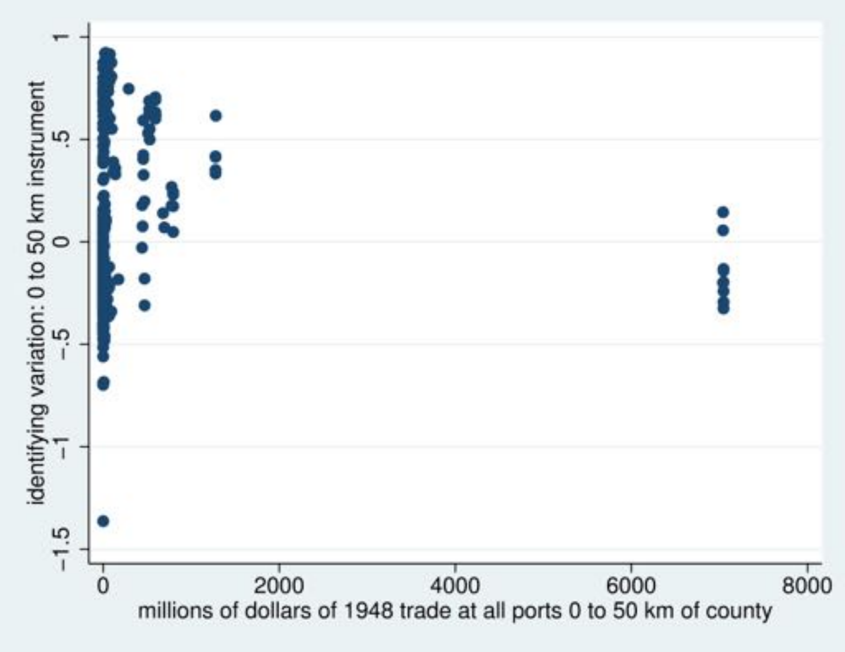
Notes: In this picture, a U.S. county has a port if there is a port within 300 km. We call this county "containerized" if $t >$ year of first containerization of any port within 300 km. We measure depth as the depth of the deepest port within 300 km. On average, deeper ports are more likely to ever containerize, and more likely to containerize early.

Figure 3: Instrument Variation vs. Pre-Treatment Covariates

(a) Versus Log of Population Density, 1910

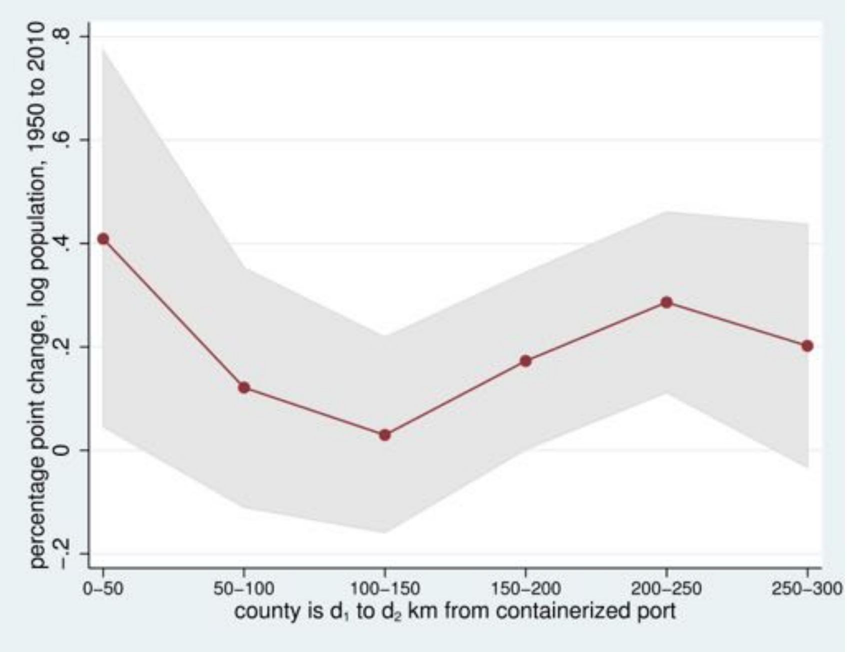


(b) Versus Millions of Dollars of 1948 International Trade at Port within 50 to 100 km



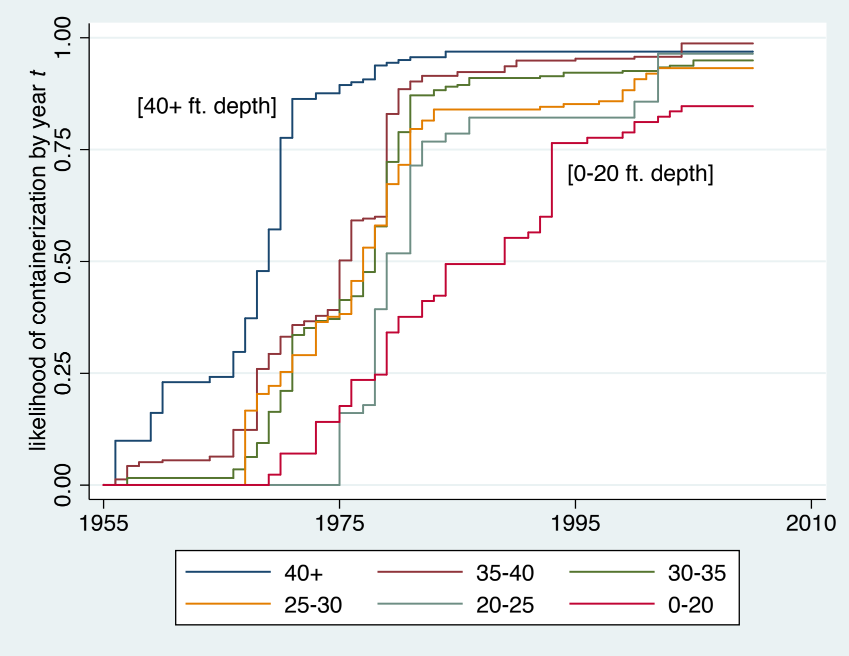
Notes: "Identifying variation" is the residual from a regression of the instrument (county is within 0 to 50 kilometers of a "very deep" port) on the full set of covariates from equation 1 (as in Table 3, Columns 4 and 8).

Figure 4: Containerization's Relationship with Population Strongest at Closer Distances



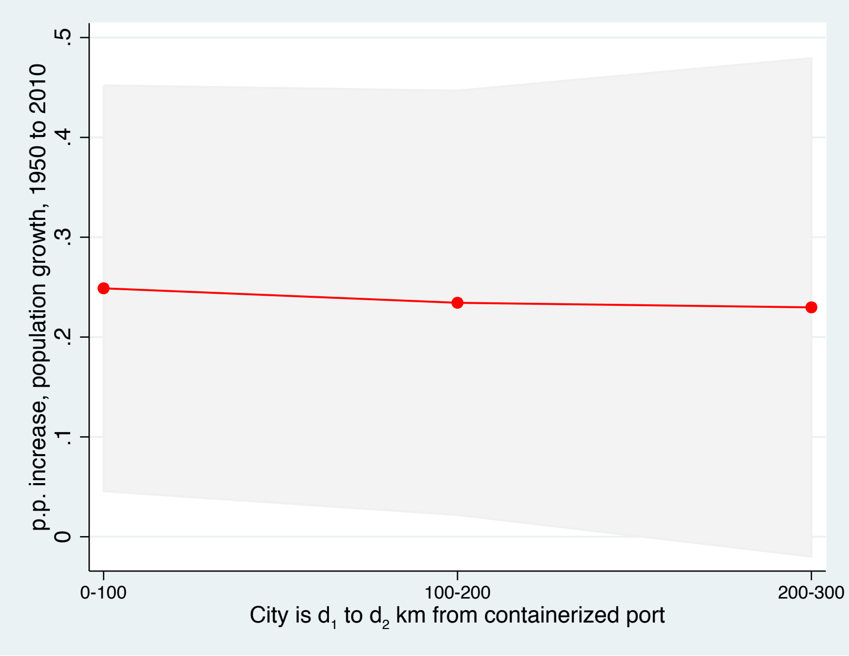
Notes: This picture presents results from Column 8 of Table 3; each dot corresponds to an estimated coefficient for each distance bin. The gray band is the 95% confidence interval.

Figure 5: Likelihood of Having a Containerized Port by 1953 Port Depth (World Sample)



Notes: In this picture, an international city has a port if there is a port within 300 km. We call this city "containerized" if $t >$ year of first containerization of any port within 300 km. We measure depth as the depth of the deepest port within 300 km. On average, deeper ports are more likely to ever containerize, and more likely to containerize early.

Figure 6: Containerization's Relationship with Population Strongest at Closer Distances (World Sample)



Notes: This picture presents results from Column 8 of Table 9; each dot corresponds to an estimated coefficient for each distance bin. The gray band is the 95% confidence interval.

Table 1: County Characteristics by Distance to Containerized Port

	Distance to Containerized Port							
	0 to 50	50 to 100	100 to 150	150 to 200	200 to 250	250 to 300	Ever Cont.	Never Cont.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Population								
1910	10.85	10.23	10.2	10.14	10.11	10.11	10.15	9.5
	[1.45]	[1.09]	[1.06]	[0.95]	[0.86]	[0.91]	[0.95]	[0.89]
1920	11.05	10.34	10.29	10.21	10.16	10.16	10.21	9.56
	[1.49]	[1.15]	[1.1]	[1]	[0.92]	[0.96]	[1]	[0.92]
1930	11.29	10.48	10.4	10.29	10.22	10.21	10.28	9.64
	[1.52]	[1.23]	[1.16]	[1.06]	[0.98]	[1.02]	[1.06]	[0.84]
1940	11.41	10.59	10.48	10.37	10.29	10.28	10.36	9.68
	[1.5]	[1.24]	[1.17]	[1.07]	[0.98]	[1.03]	[1.07]	[0.86]
1950	11.67	10.76	10.61	10.46	10.36	10.34	10.44	9.66
	[1.52]	[1.33]	[1.23]	[1.14]	[1.06]	[1.1]	[1.15]	[0.92]
2010	12.63	11.66	11.33	11.09	10.93	10.85	11	9.89
	[1.3]	[1.45]	[1.36]	[1.32]	[1.28]	[1.28]	[1.34]	[1.29]
Region								
Northeast	0.25	0.18	0.2	0.18	0.16	0.16	0.16	0.01
Midwest	0.18	0.2	0.22	0.26	0.27	0.28	0.31	0.41
South	0.41	0.45	0.45	0.45	0.49	0.49	0.44	0.44
West	0.17	0.18	0.13	0.11	0.09	0.07	0.09	0.15
Share Manuf.	0.45	0.42	0.43	0.43	0.43	0.43	0.43	0.27
Emp., 1956	[0.17]	[0.18]	[0.18]	[0.18]	[0.18]	[0.18]	[0.18]	[0.22]
Years Since	41.24	39.69	40.13	39.37	39.81	40.34	44.36	.
First Cont.	[9.19]	[10.1]	[9.59]	[9.85]	[9.19]	[9.11]	[7.6]	.
Observations	120	256	392	522	614	692	1209	1493

Notes: The unit of observation in this table is the county. We report means and standard deviations (brackets) for each variables. Source: See data appendix.

Table 2: Containerization More Likely in Counties Near Deeper Ports

	Dependent Variable is			
	Ever Containerized		Years Since First Cont.	
	(1)	(2)	(3)	(4)
Port Depth in feet is				
40 and over	1.028***		46.435***	
	(0.064)		(3.119)	
35-40	0.999***		45.434***	
	(0.063)		(3.08)	
30-35	0.888***		35.199***	
	(0.062)		(3.067)	
25-30	0.959***		40.709***	
	(0.063)		(3.072)	
20-25	0.705***		27.323***	
	(0.065)		(3.214)	
15-20	-0.012		0.183	
	(0.107)		(5.253)	
10-15	0.002		0.103	
	(0.084)		(4.127)	
$1\{\text{Depth} \geq 30 \text{ Feet}\}$		0.069***		3.793***
		(0.011)		(0.532)
Mean, Dependent Variable	0.447	0.447	19.85	19.85
R-squared	0.909	0.884	0.894	0.864
Observations	2702	2702	2702	2702
F for Excluded Instrument(s)	111	41	116	51
Increase in R^2 due to instrument	0.026	0.002	0.032	0.003

Notes: We report standard errors below coefficients in parentheses. + indicates significance at the 10% level, * at the 5% level, ** at the 1% level, and *** at the 0.1 % level. Port depth is the depth of the deepest port within 300 km. All specifications control for a dummy for ever being within 300 km of a 1953 port, the number of 1953 ports in each of six distance bins to 300 km, log population 1920 to 1950, 1956 manufacturing share, and the total value of waterborne international trade in each of six distance bins to 300 km. This table reports results using wharf depth; we show similar results for channel depth and anchorage depth in Appendix Table 2.

Source: See data appendix.

Table 3: Change in Log Population, 1950 to 2010, by Distance to Containerized Port

	OLS Estimates				IV Estimates			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ County is d_1 to d_2 km from a containerized port								
0 to 50	0.311***	0.409***	0.241***	0.240***	0.347**	0.694***	0.349*	0.409*
	(0.068)	(0.078)	(0.069)	(0.07)	(0.127)	(0.192)	(0.171)	(0.177)
50 to 100	0.351***	0.295***	0.153**	0.152**	0.428***	0.354**	0.086	0.121
	(0.049)	(0.056)	(0.048)	(0.049)	(0.093)	(0.129)	(0.113)	(0.114)
100 to 150	0.138**	0.137**	0.101*	0.095*	0.154+	0.147	0.039	0.03
	(0.043)	(0.048)	(0.041)	(0.041)	(0.083)	(0.111)	(0.097)	(0.098)
150 to 200	0.081*	0.083*	0.078*	0.076*	0.094	0.122	0.163+	0.173*
	(0.039)	(0.042)	(0.036)	(0.036)	(0.08)	(0.098)	(0.085)	(0.088)
200 to 250	0.049	0.064	0.069+	0.057	0.150+	0.192+	0.287***	0.286**
	(0.038)	(0.041)	(0.035)	(0.035)	(0.081)	(0.098)	(0.086)	(0.088)
250 to 300	-0.007	0.007	0.026	0.016	0.039	0.125	0.211+	0.202
	(0.039)	(0.04)	(0.035)	(0.035)	(0.11)	(0.135)	(0.117)	(0.126)
Covariates								
Ever 1953 Port	x	x	x	x	x	x	x	x
Distance Bins to 1953 Port		x	x	x		x	x	x
Log Population, 1920-1950			x	x			x	x
1955 Int'l Trade & 1956 Manuf				x				x

Notes: Standard errors are below coefficients in parentheses. + indicates significance at the 10% level, * at the 5% level, ** at the 1% level, and *** at the 0.1 % level. The dependent variable mean is 0.376. See text for details on exact covariates. All regressions have 2,702 observations. Sources: See data appendix for details.

Table 4: First Stage: Containerization More Likely When Ports are Deep

	Dependent Variable: County is d_1 to d_2 of containerized port					
	0 to 50	50 to 100	100 to 150	150 to 200	200 to 250	250 to 300
	(1)	(2)	(3)	(4)	(5)	(6)
A. Dependent Variable is Ever Containerized						
Depth is ≥ 30 feet, county is d_1 to d_2 km of any port, $t > 1956$						
0 to 50	0.427***	0.029	-0.023	-0.034	-0.103**	-0.008
	(0.019)	(0.026)	(0.031)	(0.035)	(0.037)	(0.039)
50 to 100	0.030*	0.521***	0.021	-0.028	0.039	-0.018
	(0.015)	(0.021)	(0.025)	(0.028)	(0.029)	(0.031)
100 to 150	0.018	-0.043*	0.518***	-0.046+	-0.004	0.025
	(0.013)	(0.018)	(0.022)	(0.024)	(0.026)	(0.027)
150 to 200	-0.025*	0.041**	-0.026	0.533***	-0.072**	-0.107***
	(0.012)	(0.016)	(0.019)	(0.021)	(0.022)	(0.024)
200 to 250	-0.002	-0.021	-0.019	-0.039*	0.480***	-0.100***
	(0.011)	(0.015)	(0.018)	(0.02)	(0.021)	(0.022)
250 to 300	-0.033**	-0.043**	-0.059***	-0.070***	-0.094***	0.330***
	(0.01)	(0.014)	(0.017)	(0.019)	(0.02)	(0.021)
Joint F test	92.3	122.3	107.9	112.4	96.8	44.4
Mean, Dep. Var.	0.044	0.095	0.145	0.193	0.227	0.256
B. Dependent Variable is Years Since First Containerization						
Depth is ≥ 30 feet, county is d_1 to d_2 km of any port, $t > 1956$						
0 to 50	25.591***	1.742	-1.373	-2.038	-6.203**	-0.5
	(1.141)	(1.577)	(1.888)	(2.12)	(2.223)	(2.341)
50 to 100	1.808*	31.261***	1.267	-1.688	2.356	-1.081
	(0.892)	(1.233)	(1.477)	(1.658)	(1.738)	(1.83)
100 to 150	1.082	-2.558*	31.083***	-2.757+	-0.238	1.5
	(0.786)	(1.086)	(1.3)	(1.46)	(1.531)	(1.612)
150 to 200	-1.476*	2.485**	-1.588	32.006***	-4.322**	-6.399***
	(0.691)	(0.954)	(1.143)	(1.283)	(1.345)	(1.416)
200 to 250	-0.148	-1.236	-1.113	-2.343*	28.817***	-6.028***
	(0.637)	(0.88)	(1.054)	(1.183)	(1.24)	(1.306)
250 to 300	-1.983**	-2.551**	-3.542***	-4.199***	-5.643***	19.802***
	(0.604)	(0.835)	(1)	(1.122)	(1.177)	(1.239)
Joint F test	92.3	122.3	107.9	112.4	96.8	44.4
Mean, Dep. Var.	2.7	5.7	8.7	11.6	13.6	15.4

Notes: Standard errors are below coefficients in parentheses. + indicates significance at the 10% level, * at the 5% level, ** at the 1% level, and *** at the 0.1 % level. All regression have 2,702 observations, and controls are as noted in Table 2. Sources: See data appendix for details.

Table 5: Change in Log Population, 1950 to 2010, by Distance to Containerized Port and Years Since Containerization

	OLS Estimates				IV Estimates			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ County is d_1 to d_2 km from a containerized port, Years Since First Cont.								
0 to 50	0.005***	0.007***	0.004***	0.004***	0.006**	0.012***	0.006*	0.007*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)
50 to 100	0.006***	0.005***	0.003**	0.003**	0.007***	0.006**	0.001	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
100 to 150	0.002**	0.002**	0.002*	0.002*	0.003+	0.002	0.001	0
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
150 to 200	0.001*	0.001*	0.001*	0.001*	0.002	0.002	0.003+	0.003*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
200 to 250	0.001	0.001	0.001+	0.001	0.003+	0.003+	0.005***	0.005**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
250 to 300	0	0	0	0	0.001	0.002	0.004+	0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Covariates								
Ever 1953 Port	x	x	x	x	x	x	x	x
Distance Bins to 1953 Port		x	x	x		x	x	x
Log Population, 1920-1950			x	x			x	x
1955 Int'l Trade & 1956 Manuf				x				x

Notes: Standard errors are below coefficients in parentheses. + indicates significance at the 10% level, * at the 5% level, ** at the 1% level, and *** at the 0.1 % level. The dependent variable mean is 0.376. See text for details on exact covariates. All regressions have 2,702 observations. Sources: See data appendix for details.

Table 6: Containerization Impact on Population Robust to Sample and Covariate Changes

	Independent Variable is					
	Ever Containerized			Years Since First Containerization		
	Port Counties Only (1)	1950 Pop $\geq$ Median (2)	1950 Pop $<$ Median (3)	Port Counties Only (4)	1950 Pop $\geq$ Median (5)	1950 Pop $<$ Median (6)
Δ County is d_1 to d_2 km from a containerized port						
0 to 50	0.283+	0.193	0.908	0.005+	0.003	0.015
	(0.155)	(0.194)	(0.637)	(0.003)	(0.003)	(0.011)
50 to 100	0.032	0.16	0.156	0.001	0.003	0.003
	(0.099)	(0.145)	(0.233)	(0.002)	(0.002)	(0.004)
100 to 150	0	-0.012	0.003	0	0	0
	(0.083)	(0.119)	(0.204)	(0.001)	(0.002)	(0.003)
150 to 200	0.077	0.091	0.259	0.001	0.002	0.004
	(0.074)	(0.119)	(0.166)	(0.001)	(0.002)	(0.003)
200 to 250	0.224**	0.345***	0.122	0.004**	0.006***	0.002
	(0.074)	(0.103)	(0.176)	(0.001)	(0.002)	(0.003)
250 to 300	0.155	0.354*	0.066	0.003	0.006*	0.001
	(0.105)	(0.161)	(0.245)	(0.002)	(0.003)	(0.004)
Observations	1296	1351	1351	1296	1351	1351
Dep. Var. Mean	0.55	0.54	0.21	0.55	0.54	0.21

Notes: Standard errors are below coefficients in parentheses. + indicates significance at the 10% level, * at the 5% level, ** at the 1% level, and *** at the 0.1 % level. Controls are as noted in Table 2. Sources: See data appendix for details.

Table 7: City Characteristics by Distance to Containerized Port (World Sample)

	Distance to Containerized Port							
	0 to 50	50 to 100	100 to 150	150 to 200	200 to 250	250 to 300	Ever Cont.	Never Cont.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Population								
1950	12.674	12.493	12.399	12.365	12.331	12.321	12.322	11.985
	[1.147]	[1.087]	[1.105]	[1.081]	[1.003]	[1.028]	[1.058]	[0.811]
2010	14.112	13.799	13.722	13.666	13.653	13.666	13.810	13.603
	[1.067]	[1.026]	[0.999]	[0.962]	[0.903]	[0.917]	[0.978]	[0.804]
Continent								
Africa	0.101	0.055	0.056	0.100	0.055	0.067	0.097	0.048
Asia	0.348	0.346	0.372	0.363	0.376	0.412	0.378	0.585
Australia	0.033	0.011	0.017	0.012	0.000	0.011	0.014	0.000
Europe	0.272	0.313	0.325	0.327	0.345	0.326	0.241	0.186
North America	0.156	0.203	0.175	0.143	0.188	0.139	0.184	0.115
South America	0.091	0.071	0.056	0.056	0.035	0.045	0.087	0.067
Years Since	32.438	31.852	30.774	30.378	31.176	30.899	35.171	.
First Cont.	[11.158]	[11.790]	[12.406]	[11.845]	[12.125]	[11.513]	[9.667]	.
Observations	276	182	234	251	255	267	632	419

Notes: The unit of observation in this table is the city. We report means and standard deviations (brackets) for each variables. Source: See data appendix.

Table 8: Containerization More Likely in Cities Near Deeper Ports (World Sample)

	Ever Cont.		Years Since Cont.	
	(1)	(2)	(3)	(4)
40+ Depth	0.524^{**}		18.022^{**}	
	(0.205)		(8.117)	
35-40 ft. Depth	0.533^{***}		16.414^{**}	
	(0.205)		(8.077)	
30-35 ft. Depth	0.515^{**}		15.754^{*}	
	(0.206)		(8.098)	
25-30 ft. Depth	0.443^{**}		9.878	
	(0.208)		(8.176)	
20-25 ft. Depth	0.486^{**}		2.389	
	(0.211)		(8.392)	
15-20 ft. Depth	0.154		-2.885	
	(0.252)		(8.662)	
10-15 ft. Depth	0.413^{*}		0.898	
	(0.214)		(8.260)	
30+ ft. Depth		0.105^{***}		9.425^{***}
		(0.033)		(1.269)
Mean, Dependent Variable	0.60	0.60	21.15	21.15
R-squared	0.89	0.89	0.88	0.87
F Stat Excluded Instrument(s)	3.14	10.00	18.17	55.12
Observations	1051	1051	1051	1051

Notes: Sample of world cities over 50K in 1950. Dependent variable (1)-(2): Adoption of containerization within 300km. Dependent variable (3)-(4): Number of years since the adoption of containerization within 300km. All specifications include country fixed effects, log population in 1950, a dummy variable for having a port within 300km in 1953, and the number of ports within 300km in 1953. Port depth is the depth of the deepest port within 300km. Robust standard errors in parentheses. Stars denote significance levels: * 0.10 ** 0.05 *** 0.01. Sources: See data appendix.

Table 9: Change in Log Population, 1950 to 2010, by Distance to Containerized Port (World Sample)

	OLS Estimates				IV Estimates			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ever Cont. (0-100km)	-0.0613	0.0362	0.0817^{*}	0.1454^{***}	-0.8494^{***}	-0.5278^{***}	0.1695	0.2488^{**}
	(0.0656)	(0.0617)	(0.0491)	(0.0488)	(0.1473)	(0.1370)	(0.1078)	(0.1037)
Ever Cont. (100-200km)	-0.1910^{***}	-0.0252	0.0436	0.0183	-0.5693^{***}	-0.2839^{**}	0.2938^{**}	0.2343^{**}
	(0.0671)	(0.0637)	(0.0528)	(0.0511)	(0.1481)	(0.1414)	(0.1155)	(0.1084)
Ever Cont. (200-300km)	-0.1553^{**}	-0.0389	0.0929^{*}	0.0790	-1.0758^{***}	-0.6995^{***}	0.2958^{**}	0.2297^{*}
	(0.0661)	(0.0609)	(0.0564)	(0.0529)	(0.1558)	(0.1434)	(0.1355)	(0.1274)
Number of ports in 1953	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Country dummies	No	No	Yes	Yes	No	No	Yes	Yes
Log population in 1950	No	No	No	Yes	No	No	No	Yes

Notes: Sample of world cities over 50K in 1950. Dependent variable: Change in log population between 1950 and 2010. All specifications include a dummy variable for having a port within 300km in 1953. The mean of the dependent variable is 1.54. All regressions have 1051 observations. Robust standard errors in parentheses. Stars denote significance levels: * 0.10 ** 0.05 *** 0.01. Sources: See data appendix.

Table 10: First Stage: Containerization More Likely When Ports are Deep (World Sample)

	Ever Cont.			Years since Cont.		
	(1) 0 to 100	(2) 100 to 200	(3) 200 to 300	(4) 0 to 100	(5) 100 to 200	(6) 200 to 300
30+ ft. Depth (0-100km)	0.552^{***}	-0.095^{**}	-0.045	21.236^{***}	-3.558^{***}	-2.925^{**}
	(0.036)	(0.038)	(0.040)	(1.413)	(1.328)	(1.397)
30+ ft. Depth (100-200km)	-0.086^{**}	0.499^{***}	-0.066	-2.895^{**}	18.119^{***}	-2.820^{**}
	(0.037)	(0.036)	(0.040)	(1.338)	(1.425)	(1.366)
30+ ft. Depth (200-300km)	-0.106^{***}	-0.096^{**}	0.484^{***}	-2.817^{*}	-2.457^{*}	19.454^{***}
	(0.041)	(0.042)	(0.040)	(1.456)	(1.463)	(1.498)
Mean, Dependent Variable	0.35	0.35	0.37	11.71	11.08	12.02
R-squared	0.65	0.63	0.63	0.68	0.66	0.67
F Stat Excluded Instrument(s)	99.72	73.97	62.39	97.77	64.31	84.79

Notes: Sample of world cities over 50K in 1950. Dependent variable (1)-(3): Adoption of containerization between d1 and d2 km. Dependent variable (4)-(6): Number of years since the adoption of containerization between d1 and d2 km. All specifications control for country fixed effects, log population in 1950, a dummy variable for having a port within 300km in 1953, and the number of ports within 300km in 1953. All regressions have 1051 observations. Port depth is the depth of the deepest port within 300km. Robust standard errors in parentheses. Stars denote significance levels: * 0.10 ** 0.05 *** 0.01. Sources: See data appendix.

Table 11: Change in Log Population, 1950 to 2010, by Distance to Containerized Port and Years Since Containerization (World Sample)

	OLS Estimates				IV Estimates			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years since Cont.	-0.0071***	-0.0031*	0.0019	0.0040***	-0.0171***	-0.0131***	0.0047*	0.0067**
(0-100km)	(0.0017)	(0.0018)	(0.0015)	(0.0015)	(0.0029)	(0.0031)	(0.0028)	(0.0027)
Years since Cont.	-0.0098***	-0.0039**	0.0019	0.0012	-0.0129***	-0.0087**	0.0080***	0.0063**
(100-200km)	(0.0018)	(0.0019)	(0.0016)	(0.0015)	(0.0031)	(0.0034)	(0.0031)	(0.0029)
Years since Cont.	-0.0099***	-0.0053***	0.0017	0.0010	-0.0204***	-0.0162***	0.0067**	0.0050*
(200-300km)	(0.0017)	(0.0018)	(0.0016)	(0.0015)	(0.0031)	(0.0034)	(0.0032)	(0.0030)
Number of Ports 1953	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Country dummies	No	No	Yes	Yes	No	No	Yes	Yes
Log Population 1950	No	No	No	Yes	No	No	No	Yes

Notes: Sample of world cities over 50K in 1950. Dependent variable: Change in log population between 1950 and 2010. All specifications include a dummy variable for having a port within 300km in 1953. The mean of the dependent variable is 1.54. All regressions have 1051 observations. Robust standard errors in parentheses. Stars denote significance levels: * 0.10 ** 0.05 *** 0.01. Sources: See data appendix.

Table 12: Containerization Impact on Population Robust to Sample and Covariate Changes (World Sample)

	Above 50K in 1950		Above 300K in 1950		Above 50K in 1950		Above 300K in 1950	
	(1) All	(2) Port cities	(3) All	(4) Port cities	(5) All	(6) Port cities	(7) All	(8) Port cities
Ever Cont. (0-100km)	0.2488**	0.2636**	0.2096	0.2012				
	(0.1037)	(0.1099)	(0.1665)	(0.1655)				
Ever Cont. (100-200km)	0.2343**	0.2192*	0.4099**	0.4623**				
	(0.1084)	(0.1186)	(0.1980)	(0.2224)				
Ever Cont. (200-300km)	0.2297*	0.2147	0.3712	0.3522				
	(0.1274)	(0.1415)	(0.2777)	(0.3075)				
Years since Cont. (0-100km)					0.0067**	0.0071**	0.0049	0.0051
					(0.0027)	(0.0028)	(0.0040)	(0.0041)
Years since Cont. (100-200km)					0.0063**	0.0057*	0.0112**	0.0118**
					(0.0029)	(0.0031)	(0.0050)	(0.0054)
Years since Cont. (200-300km)					0.0050*	0.0049	0.0058	0.0054
					(0.0030)	(0.0033)	(0.0054)	(0.0061)
Mean of Dep Variable	1.54	1.47	1.02	0.93	1.54	1.47	1.02	0.93
R-squared	0.68	0.71	0.78	0.79	0.68	0.72	0.79	0.80
Observations	1051	636	303	213	1051	636	303	213

Notes: Columns (1)-(2), (5)-(6): Sample of world cities over 50K in 1950. Columns (3)-(4), (7)-(8): Sample of world cities over 300K in 1950. Even numbered columns: Restrict sample to cities with a port within 300K in 1953. Dependent variable: Change in log population between 1950 and 2010. All specifications control for country fixed effects, log population in 1950, a dummy variable for having a port within 300km in 1953, and the number of ports within 300km in 1953. Robust standard errors in parentheses. Stars denote significance levels: * 0.10 ** 0.05 *** 0.01. Sources: See data appendix.

A Data Appendix

A.1 Data Sources

We use data from a variety of sources. This appendix provides source information.

1. County Business Patterns
   These data include total employment, total number of establishments (with some variation in this definition over time), and total payroll.
   - 1956: Courtesy of Gilles Duranton and Matthew Turner. See Duranton et al. (2014) for source details. We collected a small number of additional counties that were missing from the Duranton and Turner data.
   - 1967 to 1985: U.S. National Archives, identifier 313576.
   - 1986 to 2011: U.S. Census Bureau. Downloaded from `https://www.census.gov/econ/cbp/download/`
2. Decennial Census: Population and employment data by county
   - 1950: ICPSR 02896, Historical, Demographic, Economic and Social Data: The United States, 1790-2002, Dataset 38: 1950 Census I (County and State)
   - 1960: ICPSR 02896, Historical, Demographic, Economic and Social Data: The United States, 1790-2002, Dataset 38: 1960 Census I (County and State)
   - 1970: ICPSR 8107, Census of Population and Housing, 1970: Summary Statistic File 4C – Population [Fourth Count]
   - 1980: ICPSR 8071, Census of Population and Housing, 1980: Summary Tape File 3A
   - 1990: ICPSR 9782, Census of Population and Housing, 1990: Summary Tape File 3A
   - 2000: ICPSR 13342, Census of Population and Housing, 2000: Summary File 3
   - 2010: U.S. Census Bureau, 2010 Decennial Census Summary File 1, Downloaded from `http://www2.census.gov/census_2010/04-Summary_File_1/`
3. Port Universe and Depth
   - We use these documents to establish the population of ports in any given year.
   - 1953: National Geospatial Intelligence Agency (1953)
   - 2015: National Geospatial Intelligence Agency (2015)
4. Port Containerization Adoption Year
   - 1956–2010: *Containerisation International Yearbook* for 1968 and 1970–2010

5. Port Volume: Total imports and exports by port
    - 1948: United States Foreign Trade, January-December 1949: Water-borne Trade by United States Port, 1949, Washington, D.C.: U.S. Department of Commerce, Bureau of the Census. FT 972.
    - 1955: United States Waterborne Foreign Trade, 1955, Washington, D.C. : U.S. Dept. of Commerce, Bureau of the Census. FT 985.
    - 2008: Containerisation International yearbook 2010, pp. 8–11.
6. World Urbanization Prospects: The 2014 Revision
   These data include population counts for all urban agglomerations whose populations exceed 300,000 at any time between 1950 and 2010.
    - Downloaded from `http://esa.un.org/unpd/wup/CD-ROM/WUP2014_XLS_CD_FILES/WUP2014-F22-Cities_Over_300K_Annual.xls`

A.2 Data Choices

1. U.S. County Sample

   We drop XX counties where land area changes are greater than 35 percent. These are: [list here].

   In a future draft, we will provide a list of the county groupings we use to make the 1950 and 2010 counties geographically compatible.

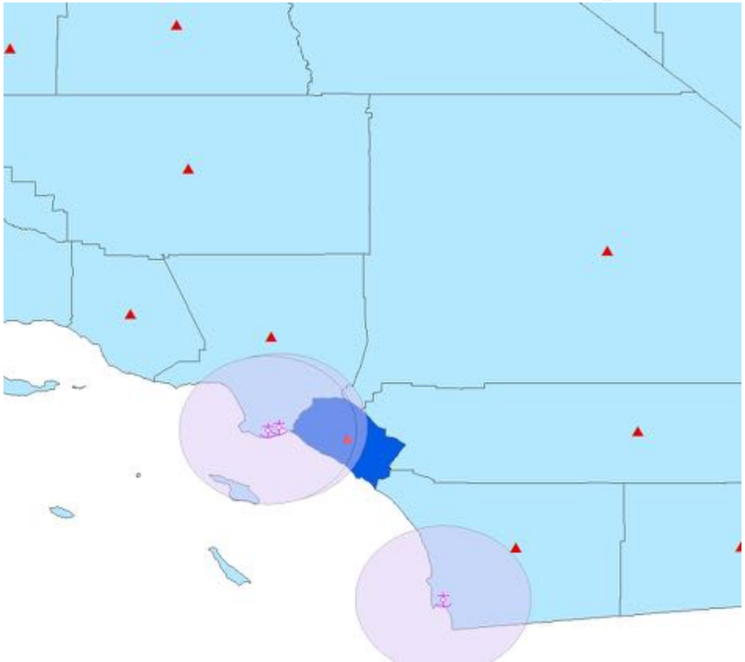
   Alaska and Hawaii were not states in 1950. We omit Alaska from our sample, because in 1950 it has only judicial districts, which do not correspond to modern counties. We keep Hawaii, where the 1950 borders are relatively equivalent to modern counties. We also keep Washington, DC, in all years.

2. Ports

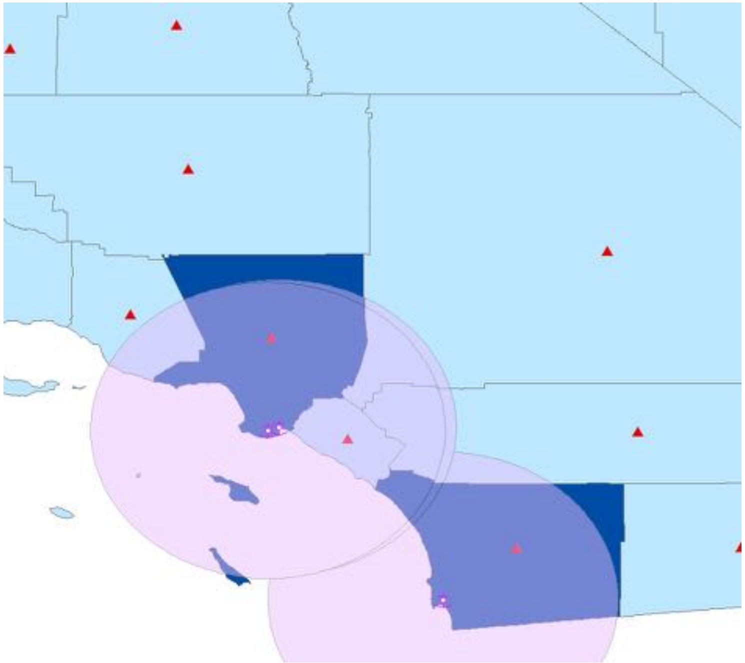
   We determine the universe of ports from the 1953 *World Port Index*.

Appendix Figure 1: Sample Construction by Distance to Port

(a) Counties within 50 kilometers of port



(b) Counties within 50 to 100 kilometers of port



Notes: Blue polygons are counties, and red triangles are the geographic county centers (centroids). The pink anchors are ports, and the grey circles show rings of 50 (top figure) and 100 (bottom figure) from the ports.
Source: See data appendix.

Appendix Table 1: Pre-Containerization International Trade by Distance to Port, United States

	Distance to Containerized Port							
	0 to 50	50 to 100	100 to 150	150 to 200	200 to 250	250 to 300	Ever Cont.	Never Cont.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1955: $ Millions of Int'l Trade in All Ports by row's distance bin								
0 to 50	960.94	136.41	266.73	159.07	72.69	107.75	103.49	0.02
	[2084.54]	[713.37]	[1229.58]	[1011.12]	[634.61]	[831.15]	[747.27]	[0.53]
50 to 100	263.24	683.12	249.84	172.95	197.05	124.46	154.96	0.04
	[1136.6]	[1716.36]	[1120.74]	[818.1]	[1086.29]	[746.85]	[863.98]	[0.61]
100 to 150	710.68	334.99	633.56	278.33	194.41	140.75	224.27	0.19
	[1927.38]	[1216.6]	[1564.15]	[1057.4]	[830.96]	[697.41]	[985.61]	[3.33]
150 to 200	332.37	371.02	337.96	586.49	267.46	252.47	267.25	0.12
	[1221.24]	[1143.06]	[1213.83]	[1483.09]	[1059.51]	[1040.41]	[1037.56]	[1.12]
200 to 250	261.98	360.8	348.85	308.51	554.01	283.61	314.91	0.55
	[1018.96]	[1297.78]	[1196.64]	[1153.69]	[1466.17]	[1121.64]	[1161.58]	[7.39]
250 to 300	549.15	389.99	471.95	385.28	361.86	610.49	389.3	1.12
	[1794.07]	[1382.2]	[1578.65]	[1289.51]	[1285.22]	[1553.58]	[1302.81]	[10.28]
1948: $ Millions of Int'l Trade in All Ports by row's distance bin								
0 to 50	812.24	109.95	228.76	138.98	60.69	93.11	87.33	0.01
	[1914.6]	[641.93]	[1118.6]	[922.04]	[575.91]	[756.12]	[678.29]	[0.34]
50 to 100	215.96	554.39	201.69	137.78	163.56	99.9	125.45	0.02
	[1005.68]	[1548.83]	[999.88]	[730.29]	[976.03]	[671.34]	[772.43]	[0.4]
100 to 150	591.27	271.95	502.97	223.56	149.21	105.63	177.95	0.12
	[1719.97]	[1093.62]	[1402.64]	[950.31]	[735.69]	[619.91]	[877.08]	[2.66]
150 to 200	261.62	286.24	272.87	462.03	215.58	201.07	209.54	0.05
	[1111.75]	[1016.17]	[1090.86]	[1334.58]	[950.47]	[930.03]	[926.23]	[0.64]
200 to 250	198.16	283.27	274.84	249.66	437.89	228.65	250.17	0.31
	[917.23]	[1162.1]	[1070.82]	[1032.86]	[1319.82]	[1005.19]	[1040.76]	[3.93]
250 to 300	452.27	317.42	387.24	309.64	293.65	485.82	311.54	0.66
	[1593.87]	[1244.47]	[1418.54]	[1157.74]	[1143.53]	[1398.74]	[1167.75]	[5.12]

Appendix Table 2: Containerization More Likely in Counties Near Deeper Ports: Alternative Depth Measures

	Dependent Variable is							
	Ever Containerized				Years Since First Containerization			
	Anchorage Depth		Channel Depth		Anchorage Depth		Channel Depth	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Depth in feet is								
30-35	0.868***		0.909***		35.869***		38.308***	
	(0.047)		(0.052)		(2.376)		(2.653)	
30-35	0.958***		0.975***		43.499***		39.255***	
	(0.049)		(0.056)		(2.458)		(2.843)	
25-30	0.961***		0.951***		40.453***		40.849***	
	(0.046)		(0.051)		(2.332)		(2.591)	
20-25	0.892***		0.897***		37.382***		36.253***	
	(0.049)		(0.051)		(2.447)		(2.613)	
15-20	0.396***		0.913***		15.494***		36.267***	
	(0.055)		(0.052)		(2.756)		(2.647)	
10-15	0.930***		-0.005		30.493***		0.085	
	(0.061)		(0.1)		(3.075)		(5.11)	
	.		-0.004		.		-0.411	
	.		(0.071)		.		(3.617)	
1{Depth $\geq$ 30 Feet}		0.211***		0.082***		10.324***		5.451***
		(0.016)		(0.011)		(0.802)		(0.523)
Mean, Dependent Variable	0.447	0.447	0.447	0.447	19.85	19.85	19.85	19.85
R-squared	0.909	0.889	0.909	0.885	0.888	0.87	0.887	0.867
Observations	2702	2702	2702	2702	2702	2702	2702	2702
F for Excluded Instrument(s)	130	170	115	58	106	166	84	109
Increase in R^2 due to instrument	0.026	0.007	0.027	0.003	0.027	0.008	0.025	0.005

Notes: We report standard errors below coefficients in parentheses. + indicates significance at the 10% level, * at the 5% level, ** at the 1% level, and *** at the 0.1 % level. Port depth is the depth of the deepest port within 300 km. All specifications control for a dummy for ever being within 300 km of a 1953 port, the number of 1953 ports in each of six distance bins to 300 km, log population 1920 to 1950, 1956 manufacturing share, and the total value of waterborne international trade in each of six distance bins to 300 km. Results using wharf depth are in Table 2.

Source: See data appendix.

Appendix Table 3: Only Counties Within 300 km of Ports: Instrument Remains Strong

	Dependent Variable is			
	Ever Containerized		Years Since First Cont.	
	(1)	(2)	(3)	(4)
Port Depth in feet is				
40 and over	1.038***		46.563***	
	(0.092)		(4.521)	
35-40	1.004***		45.742***	
	(0.091)		(4.459)	
30-35	0.890***		35.789***	
	(0.091)		(4.448)	
25-30	0.956***		41.149***	
	(0.091)		(4.451)	
20-25	0.704***		27.742***	
	(0.095)		(4.653)	
15-20	-0.017		0.647	
	(0.155)		(7.617)	
10-15	0.01		0.091	
	(0.122)		(5.973)	
1{Depth $\geq$ 30 Feet}		0.079***		3.830***
		(0.016)		(0.78)
Mean, Dependent Variable	0.933	0.933	41.384	41.384
R-squared	0.256	0.055	0.366	0.19
Observations	1296	1296	1296	1296
F for Excluded Instrument(s)	54	25	55	24
Increase in R^2 due to instrument	0.219	0.018	0.191	0.015

Notes: We report standard errors below coefficients in parentheses. + indicates significance at the 10% level, * at the 5% level, ** at the 1% level, and *** at the 0.1 % level. Port depth is the depth of the deepest port within 300 km. All specifications control for a dummy for ever being within 300 km of a 1953 port, the number of 1953 ports in each of six distance bins to 300 km, log population 1920 to 1950, 1956 manufacturing share, and the total value of waterborne international trade in each of six distance bins to 300 km.
Source: See data appendix.

Appendix Table 4: Only Cities Within 300 km of Ports: Instrument Remains Strong (World Sample)

	Ever Cont.			Years since Cont.		
	(1) 0 to 100	(2) 100 to 200	(3) 200 to 300	(4) 0 to 100	(5) 100 to 200	(6) 200 to 300
30+ ft. Depth (0-100km)	0.543^{***}	-0.094^{**}	-0.045	20.690^{***}	-3.591^{**}	-2.898^{**}
	(0.039)	(0.039)	(0.041)	(1.506)	(1.393)	(1.457)
30+ ft. Depth (100-200km)	-0.087^{**}	0.486^{***}	-0.083^{**}	-2.755^{**}	17.839^{***}	-3.238^{**}
	(0.038)	(0.039)	(0.041)	(1.367)	(1.484)	(1.420)
30+ ft. Depth (200-300km)	-0.093^{**}	-0.100^{**}	0.459^{***}	-2.809^{*}	-3.040^{*}	18.401^{***}
	(0.043)	(0.044)	(0.044)	(1.561)	(1.556)	(1.637)
Mean, Dependent Variable	0.57	0.57	0.59	19.00	18.07	19.48
R-squared	0.50	0.50	0.51	0.56	0.55	0.56
F Stat Excluded Instrument(s)	83.34	63.25	50.01	80.11	59.52	68.08

Sample: World cities over 50K in 1950, excluding cities without a port within 300km in 1953.
Dependent variable (1)-(3): Adoption of containerization between d1 and d2 km.
Dependent variable (4)-(6): Number of years since the adoption of containerization between d1 and d2 km.
Port depth is the depth of the deepest port within 300km.
All specifications control for country fixed effects, log population in 1950,
a dummy variable for having a port within 300km in 1953, and the number
of ports within 300km in 1953.
All regressions have 1051 observations.
Robust standard errors in parentheses. Stars denote significance levels: * 0.10 ** 0.05 *** 0.01.

Appendix Table 5: Specific Depth Cut-off For Instrument Not Binding

	Dependent Variable: County is d_1 to d_2 km of containerized port					
	0 to 50	50 to 100	100 to 150	150 to 200	200 to 250	250 to 300
	(1)	(2)	(3)	(4)	(5)	(6)
Depth is $\geq$ 25 feet, county is d_1 to d_2 km of any port, $t > 1956$						
0 to 50	0.478***	0.221***	-0.400***	-0.109***	-0.025	-0.014
	(0.016)	(0.022)	(0.025)	(0.027)	(0.028)	(0.029)
50 to 100	0.057***	0.512***	0.139***	-0.415***	-0.071**	-0.051+
	(0.015)	(0.02)	(0.022)	(0.025)	(0.025)	(0.027)
100 to 150	0.024+	0.100***	0.545***	0.125***	-0.532***	-0.096***
	(0.014)	(0.02)	(0.022)	(0.024)	(0.025)	(0.026)
150 to 200	0.022	0.074***	0.107***	0.549***	0.115***	-0.481***
	(0.015)	(0.021)	(0.023)	(0.026)	(0.026)	(0.028)
200 to 250	-0.007	0.007	0.060***	0.117***	0.670***	0.233***
	(0.011)	(0.016)	(0.018)	(0.019)	(0.02)	(0.021)
250 to 300	-0.01	0	0.003	0.086***	0.098***	0.660***
	(0.012)	(0.017)	(0.019)	(0.021)	(0.021)	(0.022)
Joint F test	216.9	222.8	221.2	203.1	197.8	153.3
Depth is $\geq$ 35 feet, county is d_1 to d_2 km of any port, $t > 1956$						
0 to 50	0.624***	0.168***	-0.600***	-0.119*	0.024	-0.105*
	(0.027)	(0.038)	(0.044)	(0.047)	(0.05)	(0.051)
50 to 100	0.130***	0.657***	0.055	-0.645***	-0.153***	-0.026
	(0.025)	(0.035)	(0.04)	(0.044)	(0.046)	(0.047)
100 to 150	0.007	0.174***	0.532***	-0.127**	-0.531***	-0.104*
	(0.025)	(0.035)	(0.04)	(0.043)	(0.045)	(0.046)
150 to 200	0.01	0.026	0.126**	0.617***	-0.107*	-0.492***
	(0.024)	(0.034)	(0.04)	(0.043)	(0.045)	(0.046)
200 to 250	0.009	-0.015	0.121***	0.245***	0.500***	0.121***
	(0.018)	(0.026)	(0.03)	(0.032)	(0.034)	(0.035)
250 to 300	0.01	0.068**	-0.006	0.045	0.224***	0.520***
	(0.018)	(0.025)	(0.029)	(0.031)	(0.033)	(0.034)
Joint F test	127.9	126.9	84.4	80.4	44.5	42.4
Depth is $\geq$ 40 feet, county is d_1 to d_2 km of any port, $t > 1956$						
0 to 50	0.683***	-0.075	-0.574***	-0.047	-0.05	-0.194*
	(0.048)	(0.068)	(0.077)	(0.083)	(0.085)	(0.087)
50 to 100	-0.002	0.591***	0.033	-0.393***	-0.089	0.04
	(0.04)	(0.057)	(0.064)	(0.07)	(0.071)	(0.072)
100 to 150	-0.085*	0.169**	0.547***	-0.198**	-0.317***	0.01
	(0.038)	(0.054)	(0.061)	(0.066)	(0.068)	(0.069)
150 to 200	0.090*	0.079	0.024	0.349***	-0.042	-0.352***
	(0.035)	(0.05)	(0.056)	(0.061)	(0.062)	(0.063)
200 to 250	0.052+	-0.049	-0.068	0.231***	0.376***	0.076
	(0.027)	(0.039)	(0.044)	(0.047)	(0.049)	(0.049)
250 to 300	0.012	0.008	-0.016	0.089*	0.135**	0.422***
	(0.025)	(0.035)	(0.04)	(0.043)	(0.044)	(0.045)
Joint F test	50.98	42.024	19.132	17.604	11.014	16.29

Notes: We report standard errors below coefficients in parentheses. + indicates significance at the 10% level, * at the 5% level, ** at the 1% level, and *** at the 0.1 % level. Port depth is the depth of the deepest port within 300 km. Results using wharf depth are in Table 4. Source: See data appendix.