Essays on Infrastructure and Development

Anna Tompsett

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2014

ABSTRACT

Essays on Infrastructure and Development

Anna Tompsett

Spending on infrastructure accounts for several percentage points of gross world product, reflecting its perceived importance to growth and development. Previous literature has made limited progress in providing unbiased estimates of its impacts, or causal evidence about policy changes that can alter this impact. Primarily, this is because of the selection problem: locations in which infrastructure is built differ from those in which it is not built. This dissertation provides evidence towards three important questions related to infrastructure and development. First, what role does manmade transport infrastructure play in determining and maintaining patterns of economic geography? Second, to what degree does the relocation of economic activity in response to changes in the transport infrastructure network affect estimates of the economic impact of those changes? Third, what is the effect of involving beneficiary communities in decision-making on projects to improve local infrastructure? To address the selection problem, Chapters 2 and 3 exploit quasi-experimental variation in distance to a land transport route created by the opening and location of bridges over major rivers in the historical United States, using a new dataset containing every bridge built over the Mississippi and Ohio rivers. Chapter 4 presents evidence from a randomized experiment in Bangladesh.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am grateful to my advisor Eric Verhoogen for his guidance and encouragement, and for always asking the right question.

I am also grateful to Doug Almond, Scott Barrett, Don Davis, Upmanu Lall, Leigh Linden, Malgosia Madajewicz, Vijay Modi, Frank Moretti, Suresh Naidu, Matt Neidell, Cristian Pop-Eleches, Jeffrey Sachs, Chris Small and Miguel Urquiola for suggestions, advice and support along the way, as well as to Jonathan Dingel, Jan von der Goltz, Juanita Uribe Gonzalez, Jessie Handbury, Walker Hanlon, Solomon Hsiang, Corinne Low, Gordon McCord, Kyle Meng, James Rising, Reed Walker and Tyler Williams for helpful suggestions and conversations. Many thanks to John Mutter and Mona Khalidi, as well as all the other students in the Sustainable Development, for making the last seven years a pleasure as well as a privilege.

The project described in Chapter 4 was funded by the National Science Foundation. In Bangladesh, I thank everyone at NGO Forum for Public Health for their kindness and hospitality as well as their collaborative efforts; in particular Executive Director Mr S.M.A. Rashid and my coauthor Ahasan Habib. Kabid Ahmed, Md. Shamsul Karim, Fatema Tuz-Zohora, and the indomitable Md. Rofi Uddin (Robi) provided excellent research assistance.

Thanks to the Digital Social Science Center at Columbia University for their assistance with data and technical support.

For supporting my research, I thank Columbia University, the Earth Institute, the Alliance Foundation and the Andrew W. Mellon Foundation.

For everything else, thanks to my parents, Kit and Hilary Tompsett, to Mo and Sarah Mulligan, and to Joe, Jack and Paddy: you are the best families anyone could ever wish for.

For Joe, Jack and Paddy, with love.

Chapter 1

Introduction

Two major questions can be outlined, in broad terms, about the relationship between infrastructure and development. First, we can ask what, historically, has been the impact of providing infrastructure on human development? Second, we can ask what we can learn about how infrastructure should be provided - the design and decision-making process, and how infrastructure should be paid for and managed and how this can change the impact of infrastructure provision.

My dissertation contributes to answering these questions in two ways: first, by providing robust evidence on the impact of transport infrastructure on the spatial patterns of human settlements, and to what degree this response influences estimates of the economic impact of transport infrastructure; and second, by addressing a key component of the second question outlined above: how the degree to which communities participate in decision-making affects the provision of infrastructure.

Answering questions about infrastructure has presented several important challenges to researchers. In particular, the location of infrastructure is never chosen at random; it is driven by the physical, economic and social characteristics of a location. Choices about design and decision-making processes are also correlated with other factors including which agencies implement the project and the socio-political environment in which the project is implemented. The papers that form my thesis address this identification challenge using two different approaches. In Chapters 2 and 3, I take a historical approach, using a novel dataset of bridges constructed over the Mississippi and Ohio rivers, and exploiting quasi-experimental variation in the timing and location of bridge construction to separate out the impacts of transport infrastructure on spatial and temporal patterns of population density and economic growth. Chapter 4 describes the results of a randomized field test in Bangladesh that measures the impact of varying the degree of community participation in the decision-making process on the provision of safe, community-level sources of drinking water.

In Chapter 2, I ask whether manmade transport infrastructure alters, or simply follows, pre-existing spatial patterns of human settlements. I show that that counties which experienced an improvement in access to land transport routes grew faster in the decades following the change, and that differences in density persist in the long run. The population of a county that experiences a 50% reduction in distance to a bridge grows by an additional 3% over the following 30 years, relative to a median growth rate of 15% over the same time period. The resultant differences in population density between better and worse-connected areas are persistent and larger, in the long run. Combining these results with additional evidence from bridge reconstruction and replacement and from the age of the local housing stock, I interpret these results as investments in manmade infrastructure helping to select and maintain a path-dependent equilibrium, through economies of density in transport infrastructure, increasing returns to scale and multiple waves of sunk-cost investments.

In Chapter 3, I measure the long-run impact of transport infrastructure on per-capita incomes. This long-run impact was determined in two stages: i) the direct impact of transport infrastructure on production, and ii) indirect, spatial equilibrium effects, including the relocation of firms, household and productive capital. Previous literature has been unable to separate out the two effects. I show that the direct, short-run impact of greater proximity to a transport route on per-capita production is positive. However, in the long-run, the sign of the effect is reversed: per-capita incomes are lower closer to transport routes in equilibrium. This implies that the spatial equilibrium effects are larger in magnitude, and opposite in sign, to the direct effects. I show that the likely mechanism is urbanization, followed by sorting of poorer households into more dense neighbourhoods, in response to locally reduced transport costs, and of wealthier households into less dense neighbourhoods, where housing units are larger and there are more single-unit households. These results imply that understanding the spatial equilibrium response to changes in transport infrastructure networks is critical to drawing the correct conclusions about the economic impact of transport infrastructure. Ignoring the spatial equilibrium effects would yield grossly misleading assessments of the economic impact of transport infrastructure.

Finally, in Chapter 4, co-authored with Malgosia Madajewicz and Ahasan Habib, we examine the long-standing hypothesis that participation in decision-making by intended beneficiaries of social programs improves the outcomes of those programs. We present the first experimental evidence on whether participation in project decision-making affects the outcomes of a social program. We randomly assigned participatory and non-participatory decision-making structures to communities who received an otherwise identical intervention, a package of technical advices and subsidies to provide safe drinking water sources. Participation in decision-making resulted in larger reported increases in access to safe drinking water, but only when we imposed rules on the decision-making process that were designed to limit the appropriation of project benefits by elite or influential groups or individuals. Villages in which communities participated in decision-making under rules designed to prevent appropriation reported a significantly greater increase in access to safe drinking water (an increase of 25%) relative to villages in which project staff took decisions (14%). In villages in which the communities participated in decision-making without imposed rules, the change in access to safe drinking water was the same (14%) as in villages in which project staff took decisions. We conclude that the rules we applied to limit appropriation – minimum representation requirements and decision by unanimous consensus – were effective in accomplishing their objective.

Chapter 2

Bridges: Transport Infrastructure and Economic Geography on the Mississippi and Ohio 1860-2000

Abstract

Does manmade transport infrastructure alter, or simply follow, pre-existing spatial patterns of human settlements? This paper uses local variation in distance to a land transport route around bridges over major rivers in the historical United States to show that counties which experienced an improvement in access to land transport routes grew faster in the decades following the change, and that differences in density persist in the long run. I exploit quasi-experimental variation in the timing and location of bridge construction to separate out the causal impact of the change in infrastructure from other factors that influence both decisions about infrastructure location and population growth. The population of a county that experiences a 50% reduction in distance to a bridge grows by an additional 3% over the following 30 years, relative to a median growth rate of 15% over the same time period. The resultant differences in population density between better and worse-connected areas persist, and are larger, in the long run. Combining these results with additional evidence from bridge reconstruction and replacement and from the age of the local housing stock, I interpret these results as investments in manmade infrastructure helping to select and maintain a path-dependent equilibrium, through economies of density in transport infrastructure, increasing returns to scale and multiple waves of sunk-cost investments.

2.1 Introduction

Does manmade transport infrastructure alter the location of human settlements, or do people simply build transport infrastructure in places where they already live and work? Recent studies have demonstrated path dependence in the history of human settlements, but the role of manmade transport infrastructure in determining and maintaining path-dependent equilibriums remains a largely open question. The long-standing empirical challenge is that infrastructure is built to serve pre-existing and possibly growing communities, making it difficult to separate out the causal impact of transport infrastructure access on population growth.

In this paper, I propose a novel solution to this empirical challenge by focusing on local variation in access to land transport routes created by the construction of bridges over major rivers, using a novel dataset on bridges over the Mississippi and Ohio rivers and 140 years of panel data from the United States Historical Censuses[1]. I use two empirical strategies to separate out impacts from background characteristics that affected decisions about bridge location. I first exploit variation in the timing of bridge construction to measure the transitional, 'short-run' impacts in the decades following the opening of a bridge. I then exploit variation in the location of bridge construction to measure the persistent, 'long-run' impacts of access to transport infrastructure on local population growth. I find that counties adjacent to rivers that have better access to land transport routes, proxied by distance to a bridge, grow faster in the decades following bridge construction, and that these differences persist in the long run. This approach allows me to provide the clearest causal evidence about the role of manmade land transport infrastructure in determining settlement patterns in the historical United States[2]

I use data on every bridge ever constructed over the Mississippi and Ohio rivers, a dataset that forms the longest comprehensive panel of infrastructure data available. The data were orig-

[1] Previous studies have used the construction or temporary closure of a single bridge to study the impact of a change in transport times between two locations (Åkerman, 2009; Volpe Martincus, Carballo, & Garcia, 2011) but this is the first study to use this strategy in a panel context with multiple locations and time periods.

[2] My results contrast with an earlier literature on the railroads which concluded that the railroads primarily followed, rather than led, pre-existing patterns of population growth (Fogel, 1964; Fishlow, 1965; Atack, Bateman, Haines, & Margo, 2010) and complement the findings of Bleakley and Lin (2012) who studied the long-run impacts of an obsolete natural transport advantage on economic geography; and of Duranton and Turner (2012), and Baum-Snow (2007) who studied the impact of the Interstate Highway system on growth between and within cities respectively.

inally extracted from the National Bridge Inventory, and extensively cross-referenced with other contemporary and historical sources, as well as satellite imagery. Given that the first bridge was built over the Ohio in 1849, the dataset therefore covers the rise to economic preeminence of the United States, along with the expansion of the railroads, the New Deal program of investment in infrastructure, and the creation of the Interstate Highways. The length of the panel is critical to the analysis because it allows me to measure effects that play out over several decades while accounting extremely comprehensively for unobservable characteristics that influence background trends.

To measure the 'short-run' impacts of changes in access to transport infrastructure on population growth, I exploit variation in the timing of bridge construction due to advancements in bridge technology over the time period of the study – which determine the feasibility and cost of bridge construction in a given location – and due to the time required to plan, finance, design and build a major bridge, which is on the order of several decades. For example, the Wheeling Suspension Bridge was the first bridge over the Ohio River, and the longest suspension bridge in the world when it opened; a charter to construct the bridge was issued in 1816, but the bridge itself was not completed until 1849. In contrast, the opening of a bridge creates a sharp local change in feasible journeys. Based on this historical evidence, I argue that the timing of bridge opening is exogenous to short-run deviations from local long-run growth trends, captured using county-level fixed effects and quadratic trends. Using this approach, I show that the populations of counties that experience a 50% reduction in distance to a bridge grow by an additional 3% over thirty years, relative to a median 15% growth rate over the same time period.

A natural concern regarding this approach is that bridges might be built in places that have recently experienced higher than average growth, which then continues, and that as a result this analysis would mistakenly attribute this to the causal impact of bridge construction. I show that this is not the case; there is no evidence for population growth relative to background trends in the decades preceding a change in distance to a bridge. A subsequent concern is therefore that policy-makers might build bridges in *anticipation* of higher than average growth. I rule out the possibility that policy-makers build bridges in response to county-level anticipated growth by showing the results hold even when I limit the analysis to counties that are only influenced by bridges constructed in other counties. A final concern is that bridge construction is correlated with other policies that promote growth at the state or regional level. For this to be an empirical concern, these policies would need to be timed sufficiently tightly with the construction of bridges to not affect growth in the decades before bridge construction. To allay this concern, I show that the results remain consistent in sign and timing, though are attenuated in magnitude and significance, when I include sets of progressively more conservative disaggregated time trends.

The early literature on agglomeration primarily viewed the role of transport infrastructure as a coordinating mechanism for other, unrelated economies of scale such as knowledge spillovers or labor-market pooling (see e.g. Fujita, Krugman, and Venables, 2001). However, recent theoretical work (Coşar & Fajgelbaum, 2012; Allen & Arkolakis, 2013; Armenter, Koren, & Nagy, 2013) has emphasised the direct role that economies of density associated with transport infrastructure can play in agglomeration. I show that the the short-run impacts I estimate here are a *direct* result of the change in access to transport infrastructure, and are not amplified by other, unrelated economies of scale. I demonstrate this by showing that the short-run results do not change when I narrow the comparison to places that begin a given decade with the same population density, by controlling for population density at a given time in a specification where the outcome variable is the change in log population over the following decade. My results provide the first well-identified empirical evidence to separate out the direct effect of transport infrastructure on agglomeration.

My 'short-run' empirical strategy is only valid for measuring the impacts of changes in access to transport infrastructure in the first few decades, for two reasons. First, the identifying assumption is only valid within a window of decades. Second, the county-level quadratic trends fitted in the empirical specification absorb any persistent effects. To measure the 'long-run', persistent effects, I therefore exploit variation in the location of bridges around tributary confluences. A tributary confluence is the place where a smaller river joins the main stream. As a result, the flow in the river increases sharply. Since the cost of bridge construction increases steeply with the flow rate in the river, this results in a sharp change in the likelihood of bridge construction. For example, at the confluence of the Ohio and the Upper Mississippi at Cairo, IL, there are three bridges located just upstream of the confluence, and no other bridges for another 100km downstream.

Population density increases approaching the confluence from both sides, as the tributary confluence acts as a hub for water transport routes[3]. My instrumental variables approach therefore compares places located just upstream to places just downstream of the confluence. This approach is only valid if these places only differ as a result of their access to land transport routes, proxied by distance to a bridge. To support this assertion, I show that there is no evidence for asymmetries in population density around tributary confluences prior to the era of bridge construction.

Using this approach, I show that places that are 50% closer to a bridge have population densities that are 25% higher in the long run, in my preferred specification. These results suggest that changes in manmade transport infrastructure networks lead to persistent differences in population size between better-connected and worse-connected areas, and that the long-run impacts are probably larger in magnitude than the short-run impacts.

Two plausible explanations have previously been offered for how transport infrastructure influences persistence in spatial patterns of population density. Where multiple equilibria are possible, waves of sunk costs in investments, including infrastructure, can act as a mechanism to coordinate later investment, resulting in persistence in the original equilibrium. However, sunk investments can mimic path dependency, even over long time scales, if the assets in question are sufficiently large and durable[4]. I show that although bridges themselves typically exhibit strong degrees of persistence — approximately 80% of bridges built before 1880 are still in place today — they are rebuilt many times, on average every 45 years. This provides support for the view that infrastructure's role in path dependency is unlikely to be a question of sunk costs in a specific investment, even a large and durable investment like a bridge. In further support of the importance of multiple, overlapping waves of capital investments, I show that locations closer to transport routes have higher fractions of older housing stock in the long run.

This paper therefore contributes a novel dataset and robust estimates of the impact of changes in the transport infrastructure network on the local distribution of population density to the literature on transport infrastructure and economic geography. These estimates improve on previous empirical

[3] See Fujita et al. (2001) for a discussion of the role of transport hubs in city formation.

[4] See Bleakley and Lin (2012) for a discussion.

work on the topic in this region by relying on less restrictive identification assumptions;[5] and by better accounting for spillover effects — in that a transport route passing through one county may also reduce distance to a transport route for its neighbours — and effects that take long time periods to materialize[6]. The dataset further allows me to observe the full set of infrastructure investments over time — in that I observe the universe of bridges — and also the additional investments required to maintain the bridges.

My results provide evidence for the existence of path dependency, at least among sites with similar natural advantages (in this case, access to the river). However, I also find indirect evidence for the importance of location fundamentals. Bridges, like cities, exhibit clustering characteristic of a 'Zipf's Law'-type distribution. For example, the first bridge over the Lower Mississippi — the Frisco Bridge at Memphis — was opened in 1892. This remained the only bridge over the Lower Mississippi until the opening of the Harahan Bridge in 1916, just 50m upstream. In the data, I demonstrate this clustering effect by showing that at any time, new bridges are more likely to be constructed in counties that already have at least one bridge than in counties without bridges. However, the sign of this effect is *reversed* when I account for the time-invariant probability of construction of a bridge in a given place. In other words, conditional on the fact that some places are more likely to build bridges at any moment in time, either because time-invariant local conditions are conducive to bridge construction directly, or to economic activity that then leads to investments in infrastructure. These results suggest firstly than location fundamentals remain important in explaining the distribution of infrastructure and human settlements, and that increasing returns to scale are not sufficient to dominate diminishing marginal returns from additional infrastructure (at least in the context of the possible underprovision of public goods). This result, together with the

[5]Transport infrastructure can rarely be studied in experimental contexts, a single exception being Gonzalez-Navarro and Quintana-Domeque's (2012) study of road paving in Mexico. Previous instruments for transport routes used in the literature include: straight lines or least-cost paths between major cities, which must be used with caution as instruments as they have a direct effect — in terms of being located on a natural trade route — as well as predicting infrastructure construction at all time periods e.g. rail, and later road; historical transport routes, which must also be used with caution because of the persistent effects of historical infrastructure, and because this does not remove bias due to time-invariant location characteristics; and planned infrastructure routes, which is used variously as an instrument for constructed infrastructure or as a placebo treatment, empirical applications which seemingly require opposite assumptions.

[6]In contrast to e.g. Atack et al. (2010).

finding that transport infrastructure leads to persistent differences between better-connected and worse-connected regions, also underscores the need for caution in the use of historical infrastructure as an instrument for later infrastructure.

Overall, the results illuminate the role that manmade transport infrastructure can play in determining and maintaining spatial patterns of population density. Beyond their intrinsic interest in explaining patterns of human settlement, the results are of direct relevance to policy-makers, who must plan complementary infrastructure over long time horizons at the same time as taking decisions about transport infrastructure. Understanding population mobility in response to changes in transport infrastructure networks is also key to measuring the economic impacts of transport infrastructure, given that factor mobility plays a crucial role in determining the spatial distribution of economic impacts. While the focus is on the historical United States, the results are likely to be relevant in other contexts where access to transport infrastructure remains low and there is substantial internal labour mobility, such as sub-Saharan Africa.

The paper is structured as follows. In Section 2.2 I describe the context, and in Section 2.3, the data. In Section 2.4 I describe the empirical strategy and results in the short run, and in Section 2.5 I describe the empirical strategy and results in the long run. Section 2.6 discusses the evidence regarding how manmade transport infrastructure contributes to maintaining persistence in the location of human settlements, and Section 2.7 concludes.

2.2 Historical Context: Bridges over the Great Rivers

In the early part of the 19th Century, the vast majority of inland transport in the United States was along waterways, initially the great rivers, and following the construction of the Erie Canal, via an increasingly broad network of canals[7]. By the middle part of the 19th Century, the expansion of the railroads — and the era of bridge building — had begun. River and valley crossings were expensive, and represented a significant constraint to expansion. In many cases, construction of a bridge proved a crucial final link permitting the operation of a railroad route; the Canton Viaduct

[7] The account in this section is largely based on Plowden (1974).

was completed in 1835, and the first Boston-Providence train ran 24 days later. The importance of bridges to transport journeys are sometimes evocatively captured in their names, and nicknames, such as the Short Line Bridge, between St Paul and Minnesota and the Clarksburg-Columbus Short Route Bridge[8].

In the early 19th Century, bridge construction was limited by the available materials: wood and stone. Wooden bridges typically lasted only twenty to thirty years, but it proved difficult to finance the construction of stone bridges; by 1850, only four had been constructed. The modern age of bridge construction began when economical methods of smelting iron made cast iron bridges possible. Systematic methods for truss analysis and design were put forward in the middle of the 19th Century; prior to this, bridges were designed with little or no formal attempt to calculate the loads and stresses. Human capital constraints were strongly binding; Plowden (1974) estimates that at this time there were 'probably no more than ten men in America' who were capable of designing a bridge correctly.

During the second half of the 19th Century, further developments in bridge technology made it possible to construct bridges of ever-greater spans. Cast iron was in its turn superseded; first by the development of less-brittle wrought iron, and then by steel. Other key developments included: innovations in truss design; riveted connections to replace pins; Caisson technology (compressed air boxes within which piers can be constructed below the surface of the water) and later, methods to prevent the resultant decompression sickness suffered by workers in Caissons; and the development of the suspension bridge. Bridge technology continues to evolve; the first modern cable-stayed bridges were built in Europe the 1950s and the first cable-stayed bridge over the Mississippi River was not built until 1993 (the Hale Boggs Memorial Bridge in St Charles Parish, Louisiana).

As railroad lines, and later road networks, extended westward, the Ohio, Upper Mississippi and Lower Mississippi rivers in particular represented significant obstacles to the expansion of transport routes. Figure 2.1 shows the alignment of the Ohio and Mississsippi Rivers, which cover virtually the entire North-South extent of the United States, constituting a major barrier to the creation of East-West land transport routes. The progressive improvements in bridge technology allowed these

[8]Later renamed and then replaced.

obstacles to be overcome, but the process was slow and required extensive experimentation and innovation; Plowden (1974) describes the Ohio River as a 'virtual outdoor museum of American bridge engineering.' The first bridge over the Ohio River — the John A. Roebling Bridge at Wheeling — was the longest suspension bridge in the world when it was completed in 1849. The first bridge over the Lower Mississippi — the Frisco Bridge at Memphis — had the longest span of any bridge in the United States when it was built in 1892. The earliest bridges over the Upper Mississippi were built in very specific locations conducive to bridge construction; at Nicollet Island in Minneapolis in 1855 and Rock Island, Illinois in 1856. [9]

Multiple factors affect the difficulty and expense of bridge construction at a given site. The width, depth and speed of the river all increase the cost and complexity of bridge construction. All these factors are associated with higher flow rates, but are also influenced by the gradient of the river, the shape of the side slopes and the riverbed material. The bed material also affects the difficult of constructing stable piers in the riverbed; it is straightforward to build a stable pier on rock but much more difficult in shifting sands. Navigation requirements, which determine the required clearance between maximum water level and the lowest point of a bridge, and the minimum acceptable distance between supports for the widest span of the bridge, also influence bridge design. A river in a wide, flat plain may also experience considerable movement from year to year, requiring a much larger total bridge length to accommodate potential shifts in river course. As a result, a substantial fraction of the variation in timing of bridge construction results from interactions between local factors that influence the cost and difficulty of bridge construction, and global time trends in available bridge technology and expenditure on infrastructure (such as the expansion of the railways, the New Deal Public Works Administration, or the creation of the Interstate Highway network).

However, there is characteristically a long but variable lag between the first discussions of the need for a bridge, and opening of the bridge itself. Long before construction or even design of a bridge begins, stakeholders — which often include politicians from multiple fiscal and political

---

[9] Islands reduce the cost of bridge construction by dividing the stream into two; it is much cheaper to build two shorter bridges than one longer bridge. However, islands make a poor candidate for an instrument for bridge construction, as islands also offered other advantages to potential settlers; in particular, they are highly defensible.

jurisdictions — must negotiate who will fund the bridge, and how they will raise the capital, a complex problem of collective action. Many of the bridges considered here connect not only counties, but also states, implying still greater obstacles to successful resolution of the collective action problem. Like all major civil engineering works, every bridge constructed is unique, responding to idiosyncratic local hydrogeological and geographical conditions. Both design and construction take several years, and delays are frequent. Bridges are also particularly vulnerable to damage or even destruction during construction, if exposed to extreme weather conditions.

It is difficult to document the length of these lags, as I do not in general have documentary evidence of the start of the decision-making process. The systematic search for documentary evidence is made particularly difficult by the fact that bridges are often only named after construction, meaning that identifying the first reference to a particular bridge is difficult, even where potential textual sources are digitized. Anecdotal examples are rife, and include the following: a charter to construct the Wheeling Suspension Bridge was issued in 1816 — but the bridge itself was not completed until 1849 (Plowden, 1974). The need for a bridge at St Louis was identified by 1836, but construction did not begin until 1867 and the bridge was not completed until 1874 (Plowden, 1974). The Memphis and Arkansas Bridge, completed in 1949 at Memphis, was popularly known as the 'Eleven-Year Bridge' after the time it took to construct (Cordell, 2011). A committee was formed in 1946 to discuss a bridge linking West Tennessee to Missouri, but approval for a bridge was not obtained until 1964, bridge construction began in 1969, and the Caruthersville bridge was completed in 1976 (Cordell, 2011). More recently, planning for the New Mississippi River Bridge at St Louis began in or before 1991, but construction only began in 2010 and is not expected to be completed until 2015.

In contrast to these lags between identification of a need and opening of the bridge, the impacts of the bridge — changes in feasible routes and journey times — are substantially realised in a single day, when the bridge is opened, although the impact may increase over time with the construction of complementary infrastructure.

The measure of access to land transport infrastructure that I use in this study is the distance to an extant bridge. This measure is also influenced by bridge closures, implying that the results incorporate both the effects of a bridge opening and the effects of a bridge closing. Most planned bridge closures are cases where bridges are replaced nearby, but in this case, the resultant change in distance to a bridge is small, so (correctly) these cases are not likely to greatly influence the analysis. There are a relatively small number of cases where bridges are not replaced nearby — for example, only one county of 181 has no bridge in 2000, but had a bridge at an earlier time. Historical evidence suggests that, cases where bridges are locally replaced aside, the actual timing of bridge closure or destruction is largely random and driven by concerns about safety or extreme weather conditions. For example: the Pink Bridge, at Fort Ripley, Minnesota was destroyed in 1947 by high water and an ice jam; the Silver Bridge between Point Pleasant, West Virginia and Gallipolis, Ohio collapsed in 1967 as the result of a failure of a single eyebar in a suspension chain; and in the aftermath, the Clarksburg-Columbus Short Route Bridge — just upstream and of a similar design — was closed, as the design was no longer considered safe.

I will discuss further the identifying assumptions that underlie the empirical analysis in Sections 2.4.1 and 2.5.1

2.3 Data

2.3.1 Bridge Data

The bridge dataset contains information on every bridge ever constructed across the Mississippi below Lake Winnibigoshish in North Central Minnesota, and across the Ohio below Pittsburgh, where the Monongahela joins the Allegheny to form the Ohio. I originally extracted data on bridges over the Mississippi and Ohio Rivers from the National Bridge Inventory (NBI), a dataset compiled by the Federal Highway Administration containing information on the more than 600,000 bridges and tunnels in the United States that have roads passing above or below them. I then extensively hand cross-checked the data extracted from the original database with both satellite imagery and alternative sources of information on bridges (see Appendix A for more details). Above the chosen cut-off point in Northern Minnesota, the Mississippi River meanders extensively among a series of lakes — and is no longer clearly visible as a single channel in satellite imagery — meaning that its role as a barrier to East-West land transport routes is much less clearly defined[10]. Where bridges cross the river at an island, I include only the main channel bridge in the dataset and exclude the back channel bridge or bridges.

Wherever historical bridges were mentioned that no longer exist, I added them to the dataset along with the year of demolition or collapse. To verify coverage of bridges that no longer exist, I compared the data obtained in this way to the US Army Corps of Engineers *List of bridges over the navigable waters of the United States* from 1941 (Office of the Chief of Engineers, United States Army, 1948) to ensure that bridges that had collapsed or been destroyed were included in the database.

In this study, I will focus on counties which are completely covered by the sample of bridges. In Table 2.1, I show key characteristics of the bridges included in the study, excluding those from the far northernmost extent of the original sample which only partially overlap a county, based on the 1860 boundaries. Figure 2.2 shows the geographical distribution of bridges on the Mississippi and Ohio in 1860, and in 2000. Only 4 bridges were constructed prior to 1860: the Wheeling Bridge on the Ohio; and the Rock Island Arsenal, Hennepin Avenue and Wabasha Street Bridges on the Upper Mississippi. [11]

Most of the bridges in the study are road bridges; around a quarter are either rail bridges, or of mixed use i.e. have a rail crossing and a road crossing. Before 1900, the majority of bridges constructed were rail bridges. There are later peaks in bridge construction activity during Roosevelt's New Deal programs at the end of the Great Depression, and during the construction of the Interstate Highway System. Figure 2.3 illustrates the overall patterns in the timing and location of bridge construction.

The length of the maximum span is a measure of the cost and difficulty associated with bridge construction at a given location. The maximum span increases with time consistently up until the

[10] In specification checks I will test the results of cropping the sample at an alternative, lower point on the Upper Mississippi — based on an engineer's informal assessment that this lower point represents the cut-off point of bridge structures that represent major civil engineering works (Weeks, n.d.).

[11] Bridges were also constructed at Broadway Avenue, Little Falls and Broadway Avenue, Minneapolis, in 1857, but they were both destroyed within two years and not replaced for more than twenty years, so I treat them as abortive attempts to construct a bridge.

middle of the 20th Century, reflecting improvements in bridge technology that permitted longer spans to be constructed cost-effectively. The increase is fairly consistent with time, since improvements in bridge technology have largely been incremental rather than revolutionary. The anomalous value for bridges constructed prior to 1860 is entirely driven by the Wheeling Bridge, which was at the time the longest suspension bridge in the world. For brevity, I do not show the comparisons here, but bridges over the Upper Mississippi have the shortest spans, followed by the Ohio, while bridges over the Lower Mississippi are substantially longer. The timing and frequency of bridge construction over these rivers also reflect the differences in construction difficulty associated with span length.

The total length of the structure is also a measure of the cost and difficulty associated with bridge construction, but it is more strongly influenced by the intended use of the structure. Rail bridges are longer than road bridges, as road vehicles can handle a steeper incline than trains. The changes in bridge use over time from rail towards road therefore also influence trends in total structure length. Although not shown here, road bridges increase consistently in total length, as well as in the length of the maximum span.

There is likely to be some measurement error in the data on span and bridge length, as where several bridges were constructed at the same site these values may not apply to all structures. However, bridge rebuilds often reuse parts of the same structure, particularly the piers, so the length of the maximum span and the overall length may not change much with time even when the bridge is rebuilt. Navigation requirements and construction logic (building at any time the shortest feasible span) imply that the maximum span at a given site is extremely unlikely to be shorter for an extant structure than for a previous structure in the same location.

Traffic data is only available for road bridges listed in the NBI, and is missing for 33 road bridges, including of course many of those that were no longer in place at the time of the traffic count. Traffic counts date to a particular year, usually 2005 or 2006 in this dataset. The daily traffic counts (typically in the tens of thousands) illustrate the strategic importance of these crossings on East-West transport routes across the US. The most extensively used crossings have traffic counts numbering in the hundreds of thousands. Traffic counts peak for bridges constructed in the periods around construction of the Interstate Highway System.

2.3.2 Population Data

Population data is drawn from historical censuses from the United States. Although census data has been collected in the United States since 1790, the area of coverage, and the questions asked, have varied with time. This study uses three sources of census data. The unit of analysis is the county, since this is the finest level of spatial detail available over the full historical period of interest; census blocks and tracts and zip codes were all defined at a later date than the start of this study.

First, I use aggregated data on population from the United States Censuses from the National Historical Geographical Information System (NHGIS) [12]. I also obtained shapefiles for historical county boundaries from this source. Second, I obtained aggregated data on other population variables at the county level from the Inter-university Consortium for Political and Social Research (ICPSR) [13].

However, not all variables are available consistently across time from these sources. I estimate county-level aggregate variables where they are not otherwise publicly available by using individual-level data from the Integrated Public Use Microdata Series (IPUMS)[14]. Full individual level data for a sample of households is available at the county level up until 1940, and for a subset of counties thereafter, where a Public Use Microdata Area (PUMA) coincides with a county's boundary. The PUMA is the lowest unit of geography available in the microdata files after 1950, which for confidentiality reasons is set to include at least 100,000 people. As a result, less populated counties are often aggregated together into one PUMA.

I deal with changes in county boundaries over this time period by remapping all data back to

---

[12] Minnesota Population Center. National Historical Geographic Information System: Version 2.0. Minneapolis, MN: University of Minnesota 2011 http://www.nhgis.org

[13] Haines, Michael R., and Inter-university Consortium for Political and Social Research. Historical, Demographic, Economic, and Social Data: The United States, 1790-2002. ICPSR02896-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-05-21. doi:10.3886/ICPSR02896.v3

[14] Steven Ruggles, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota, 2010.

1860 county boundaries. Where two counties have been separated, I sum the total population of the two counties and assign the information to the original county. Where counties have merged, I assign the population to the original counties according to the spatial ratio between the original county and the merged county. With the individual-level data, I deal with counties which have merged by assigning a household from the merged county to both of the original counties, with a weight corresponding to the spatial ratio between the original county and the total merged counties. I therefore use a balanced panel of counties throughout the time period, at the cost of a small increase in measurement error. The baseline year is 1860. The first bridge built on the Ohio was built in 1849, and the first bridge built on the Mississippi was built in 1855, meaning that starting from 1860 captures almost all the variation in distance to a bridge that exists after the measure can be defined. The number of counties I can include in the study increases significantly between 1850 and 1860. In robustness checks, I will show that the results remain consistent and significant when I change the start date to either 1840 (before any bridges are constructed on the rivers) or 1880 (to avoid the Civil War and the last decade of slavery), or the end date to 1960. However, the lead coefficients are significant in some specifications which begin at a later date, suggesting that specifications that start at an earlier date do better at capturing long term trends, especially since the widest variation in growth rates is observed in the earliest decades.

In Section 2.5, the unit of analysis is the census tract (in the year 2000), rather than the county. The spatial and census data on census tracts is obtained from the NHGIS. In Section 2.6, I also use data on the age of the housing stock, obtained from the same source.

2.3.3 River Data

To map the location of the river and match river characteristics to counties, I used three different spatial datasets, described in Appendix A. Each has a slightly different river alignment, reflecting the resolution of the dataset and changes in the river alignment over time. Using the three datasets enables me to best match rivers and counties, since in some cases the river no longer lines up with county boundaries which were clearly originally defined using the river alignment.

Using spatial mapping, I determine whether or not any part of the county intersects the river alignment, using a 200m buffer zone. I construct an indicator for whether or not the county is on the river based on whether the county intersects the river in any of the three datasets used. In Section 2.5, I focus on a continuous sample of census tracts, where the boundary is defined by any part of the census tract being within 10km of the river.

The most informative dataset, containing information from the National Hydrology Dataset, also contains flow characteristics. I do not have data on river width, and approximations to river width based on flow and gradient data have very limited accuracy and appear to provide little additional information. However, river width may be endogenous in urbanized areas, where river banks may have been realigned, canalized or reinforced to reduce flooding or erosion, so river flow may be the preferred measure of river size.

2.3.4 Sample

Figure 2.4 shows a map of the counties used in the study for the short-run analysis, presented in Section 2.4. I include only counties which border the river (based on one of the three river datasets described above), for which the bridge sample entirely covered the county area. There are 181 counties in the sample, from 14 states.

Using spatial mapping and hand-checking, I matched the bridges to the counties on either side of the river that they connect. Table 2.2 shows descriptive statistics covering bridge access for the counties in the sample. Columns 1) to 4) focus on those counties within the sample in which bridges are ever constructed, prior to the year 2000, of which there are 124 in total. Only two counties were connected by a bridge in the past but do not have a bridge at present; Louisa County, Iowa, and Mercer County, Illinois, were once connected by the Keithsburg Rail Bridge but have not been connected since its destruction by fire in 1981.

Many bridges that are lost (through closure, collapse or destruction) are replaced locally, if not in the identical site. A bridge reconstructed in the same site is counted in the analysis as a rebuild of the same crossing, while a bridge constructed in a different site is treated as a different bridge, even if it is explicitly constructed to replace the other bridge. The likelihood of losing a bridge increases with the number of other bridges in the county; bridges are more likely to be destroyed or moved when there are many local alternatives, although no causal relationship is established.

I measure distance to a bridge using the distance between a county's centroid and the nearest bridge. Column 4) shows this measure for counties which ever acquire a bridge; column 5) shows this measure for counties which never acquire a bridge. Both groups of counties experience important changes in distance to a bridge in the earlier period of the study; from 354 km to 17 km for counties acquiring a bridge before the year 2000, and from 592 km to 38 km for counties which have not acquired a bridge by then. The large changes in distance to a land transport route in counties which do not themselves acquire bridges suggests important spillover effects and motivates my focus throughout on distance to a bridge as a measure of access to transportation infrastructure. A simple comparison of places which acquire a bridge before and after bridge construction would not capture these potentially important effects.

Changes in distance to a bridge after 1940 are small, and changes since 1960 almost zero. I also experimented with alternative measures of bridge distance such as the distance between the nearest place on a river and the bridge site, but these measures had little predictive power and may well have been subject to more measurement error. It might seem more logical to focus on the distance between the population-weighted centroid and the river, but I do not have sub-county population information in 1860, and modern day population distribution is endogenous to transport infrastructure location.

Table 2.3 shows summary statistics for population. Panel A shows statistics for all counties; Panel B shows statistics for counties which ever acquire a bridge; and Panel C shows statistics for counties which have not acquired a bridge by 2000. Among river counties, as might be expected, counties in which bridges have been or will be constructed have higher populations and population densities in 1860 than counties which never acquire bridges, and also experience higher population growth in all time periods except the most recent decades.

There is a high level of variation in regional population density, measured using the relative variance of log population density, following Davis and Weinstein (2002). Relative variance in log population density is calculated by taking the log of population density, calculating the variance for a given time and dividing by the same measure calculated for the full sample in the year 2000.

The advantages of this measure is that it is independent of region size and invariant to average density, which rises with time (Davis & Weinstein, 2002).

In all samples, variation in regional population density is higher at present than any time covered by the study. However, variation in regional population density was decreasing in the earliest period of this study. This may reflect the spread of population to a more even distribution across the sample area during the very early part of the study period — when counties were still being settled, particularly in the northern extremes of this sample — before a period of increasing agglomeration. In the last decades of the study period, the rate of increase in the relative variance in log population has slowed substantially for counties in which bridges have already been constructed before 2000, but continues to increase in counties which have not acquired bridges by this time.

In a broader sample of counties, the relative variance of log population is higher in the river counties than in the off-river counties[15]. Since transport routes are more tightly clustered around bridge crossing sites than they are away from rivers, greater variance in population along the rivers is consistent with transport routes influencing spatial patterns of population distribution, and with bridges acting as critical nodes in the transport network.

2.3.5 Selection in bridge sites

Multiple factors influence choices about bridge location. The physical characteristics of the river at a given location affect the cost and difficulty of bridge construction. The potential benefits to be realised are determined by the social and economic characteristics of the proposed location and its surroundings. The differences shown in Table 2.3 make clear that counties in which bridges are constructed are fundamentally different from counties in which bridges are not constructed; they have higher populations (both in raw terms and in population density), and different long-term average growth rates. These differences strongly suggest that any plausible empirical strategy must account first for county size and initial population, and then for time invariant characteristics that account for differences in population growth across time.

Over the time period of the study, bridges show a significant level of clustering; the majority of

[15] Results not shown here.

new bridges constructed are built in counties which already have bridges. This raises the question of whether or not acquiring a bridge increases the likelihood of having a second bridge constructed in the future, or whether the clustering in transport routes is explained by time-invariant characteristics of the counties where the routes are sited. This seems of especial interest given that the presence of an earlier transport route has been used as an instrument for the presence of a later transport route in the literature (e.g. Duranton and Turner (2011)).

In Table 2.4 I show that, conditional on year and county fixed effects, counties in which a bridge has already been constructed are *less* likely to have a future bridge constructed. The intuition appears to be that addition of a second transport route experiences diminishing marginal returns. Clustering in bridge sites apparently reflects selection based on time-invariant physical and socioeconomic characteristics rather than a dynamic, path dependent process. This reinforces the need to account comprehensively for unobserved location characteristics as part of a convincing identification strategy.

2.4 Short-Run Impacts

2.4.1 Empirical Strategy

The short-run empirical strategy is motivated by the observation that *timing* of bridge opening at a given location is exogenous to short-run deviations from local long-run trends in population growth. The timing of bridge opening — and thus the timing of changes in distance to a bridge — is driven by several factors: 1) the physical characteristics of a place that affect the feasibility and cost of bridge construction; 2) technological developments that influence the cost and feasibility of bridge construction; 3) global trends in infrastructure spending; 4) bridge construction decisions in other counties, since the likelihood of bridge construction in a given county is reduced if a bridge is constructed in a neighbouring county; 5) the time it takes to finance, plan and construct a bridge, which may include design complications, unanticipated construction problems or accidents, and uncertainty created by political decision-making processes; and 6) factors that influence the anticipated benefits to be realised from bridge construction.

Of these factors, the concern for identification is 6), as it is possible that decisions about bridge construction could respond to recent or anticipated changes in population growth rates. The presence of 5) somewhat mitigates this concern, as these lags make it less likely that the timing of bridge opening be correlated with recent or anticipated changes in population growth. The identification strategy therefore rests on the identifying assumption that the timing of changes in distance to a bridge is uncorrelated with short-run deviations from long-run growth trends, within a window of several decades around the actual timing of the change. With county-level panel data over 140 years in the main specification, I approximate the long-run trend in a county using county fixed effects and quadratic trends.

I use the following as the main estimating equation in the short run:

$$y_{i,t} = \gamma_t + \alpha_{0i} + \alpha_{1i}t + \alpha_{2i}t^2 + \sum_{j=0}^{k} \beta_j \Delta dist_{i,t-j} + \epsilon_{i,t} \tag{2.1}$$

where $y_{i,t}$ is the outcome variable i at a time t, γ_t is a year fixed effect that flexibly captures global trends in the outcome variable and distance to a bridge, α_{0i}, α_{1i} and α_{2i} are county-specific parameters that approximate the long-term counterfactual, $\Delta dist_{t-j}$ is the change in log distance to a bridge j time periods ago, and β_j is the coefficient of interest, the cumulative effect on the outcome variable at time t of a change in distance to a bridge j periods ago.

Figure 2.5 illustrates the short-run identification strategy. Fitting a long-term quadratic trend line absorbs persistent effects, so I only estimate the first few β_j terms without significant bias, around the sharp change experienced in distance to a bridge. If long-term effects exist, they tend to bias my estimates towards zero, as long as they are of the same sign as short-term effects. The lag length k must be chosen so that k is sufficiently large to ensure that the coefficients of interest are estimated without bias, given negative serial correlation in the $\Delta dist_{t-j}$ terms[16]. In specification tests, I will vary the number of lagged measures included, to test stability of the coefficients in the period of interest.

[16] In particular, because counties acquiring bridges do not then experience large future changes in distance to a bridge.

In formal terms, the identifying assumption therefore is that:

$$\begin{aligned} E(\epsilon_{i,t}|\Delta dist_{i,t+k}, \alpha_{0i}, \alpha_{1i}, \alpha_{2i}, \Gamma) &= 0 \\ t &= 1860, 1900, ..., 2000 \\ k &= -30, -20, ..., 20, 30 \end{aligned} \tag{2.2}$$

where Γ is the vector of year fixed effects. In other words, the timing of a change in distance to a bridge is exogenous to deviations from the long-term trend within a window of 30 years either side of the date at which construction takes place. This is locally equivalent to assuming that $E(y_{it}|\alpha_{0i}, \alpha_{1i}, \alpha_{2i}, \Gamma) = \gamma_t + \alpha_{0i} + \alpha_{1i}t + \alpha_{2i}t^2$ for $t = 1890, 1900, ..., 2000$. In a robustness check, I will show that the choice of quadratic trends is not critical; the result is not altered by using an equivalently conservative piecewise linear spline instead. Once the county fixed effects are included, carrying out the analysis for population in terms of log population or log population density yields exactly equivalent results.

The identifying assumption would fail if the construction of a bridge at a given time was correlated with deviations in the growth rate of the outcome variable from the time-demeaned county average before or after the construction of a bridge. There are two particular cases which would create concerns for the identification strategy. First, my estimates would be biased upwards if policy-makers decide to build bridges in response to periods of relatively low growth in the outcome variable, or in response to the start of a 'boom' which then continues. In these cases, I might mistakenly interpret a return to the mean, or the effect of the 'boom' as a causal impact of bridge construction. In contrast, if policy-makers decide to build bridges in response to several decades of growth, this would tend to bias my estimates downwards. In robustness tests of the main specification, I include lead measures of bridge distance, to test whether contemporary population predicts future changes to bridge distance, conditional on long term trends.

Second, policy-makers may anticipate higher future growth (relative to county-level long term trends), and decide to construct bridges in response. I test whether the results hold for both counties which acquire bridges and counties which do not acquire bridges (and are therefore only effected by bridges built elsewhere). If the results hold in counties which do not acquire bridges, this suggests that county-specific anticipated growth cannot explain the results.

A further concern is that bridge construction is correlated with other policies that promote growth at the state or regional level. To test for this possibility, I carry out further robustness checks in which I vary the specification of the overall time trends, and allow them to vary with geography. I also test whether the results are internally consistent for each of the three rivers (the Upper and Lower Mississippi; and the Ohio), and for robustness to varying the start and end dates of the study period.

For the main analysis, I also test how the results vary when I include controls for lagged population density. In order to do this, I respecify Equation 2.1 in differences, in order to avoid a lagged dependent variable and deal with near-collinearity between population and lagged population. The resulting equation is:

$$\Delta y_{i,t} = y_{i,t} - y_{i,t-1} = \lambda_t + \alpha_{4i} + \alpha_{5i}t + \sum_{j=0}^{k} \tau_j \Delta dist_{i,t-j} + \epsilon_{i,t} \tag{2.3}$$

where τ_j is the effect of growth between time $t-1$ and time t of a change in distance to a bridge j periods ago such that $\beta_j = \sum_{l=0}^{j} \tau_l$.

To make the correct inference about whether differences in population growth are statistically significant, it is important to correct for serial correlation in population growth rates, and spatial correlation across counties (Bertrand, Duflo, & Mullainathan, 2004; Angrist & Pischke, 2009). Previous analysis suggests that positive serial correlation persists over two or three decades, once county fixed effects are taken into consideration, but that the fixed effects structure results in negative correlation in the residuals over a longer timescale (see Wooldridge, 2001). It seems therefore conservative to cluster standard errors at the county level, which allows for arbitrary correlation within observations from a single county, as recommended by Wooldridge (2001),

In addition, to account for the possibility of spatial correlation in the standard errors, I calculate Conley standard errors (Conley, 1999), adapting code developed in Hsiang (2010), and add these to the clustered standard errors (subtracting the robust standard error matrix to avoid double-counting within-county correlations). These allow for spatial correlation over a distance of 200km between county centroids, using a uniform kernel as recommended by Conley (2008).

The measure of distance from a bridge is calculated taking into consideration bridge closures. This helps reduce measurement error in cases where a bridge is replaced in a nearby, but not identical location; the resultant change in distance to a bridge is then minimal.

The principal alternative strategy for crossing a river is to use a ferry (or historically, to cross over the ice during the winter; railroad tracks were even laid down directly on the ice during the winter). It is possible that the locations of ferry crossings on the river may interact in some way with the locations of bridges, although it is not particularly obvious that sites well suited to ferry crossings should also be well suited to bridge crossings. Ferry crossings require shallow approach slopes to facilitate loading, while bridge crossings are cheaper where the river is narrower and the ground material more stable, which tends to be associated with steep, rocky banks. I have not been able to identify any consistent source of information on historical ferry crossings but it seems likely that the presence of ferry crossings tends to bias the estimates downwards. First, where bridges replace ferry crossings, and ferry crossings had a positive impact on growth, then the quadratic trends will reflect the positive impact of ferry crossings and bias downwards the estimated impact of the later bridge. Second, where ferry crossings relocate in response to bridge construction upstream or downstream, this would tend to improve transport access in areas further from the bridge, which would again bias my estimates downward.

2.4.2 Results

In Table 2.5, I show the results of the main short-run analysis. The coefficients in the table can be interpreted as the cumulative effect on population at time t of a change in distance to a bridge j years ago. In column 1) I show the main results: over 30-40 years, there is a gradually increasing cumulative effect on population of a change in distance to a bridge. The coefficient is negative because changes in distance to a bridge are largely negative, and an increased magnitude of a change in distance to a bridge is associated with a greater increase in population. A 50%

reduction in distance to a bridge is therefore associated with approximately 3% greater population, thirty to forty years after the change takes place. The results are statistically significant.

Column 2) shows the result of including lead (future) changes in bridge distance. A significant coefficient on future changes to bridge distance would suggest that contemporary population (conditional on long term trends) could predict future changes in bridge distance. If this were the case, this would provide evidence for a violation of the identifying assumption, but the coefficients on future changes in distance to a bridge are close to zero. I show the results from column 2) in Figure 2.6; the lead coefficients are clearly not statistically significant, while significant differences emerge after one to two decades in the lag coefficients. I reverse the y-axis in all figures that show the response to a change in distance to a bridge so that the graphs are intuitively easier to interpet; a rise in the outcome variable is shown as a rise on the figure. In analysis not shown in the paper, I show that with either county fixed effects only or county linear trends, the lead coefficients are significant, indicating that the comparison is biased by long term average growth rates or trends in growth rates. An equivalent analysis using indicators pre- and post- bridge construction results in imprecise estimates, for which none of the coefficients on time dummies pre- or post- bridge construction is significant. This results from failing to take into consideration the important spillover impacts on neighbouring counties[17].

In Table 2.5, column 3) I show the result of altering the analysis to include more lag variables. None of the coefficients on the additional lag variables is significant — which is to be expected given the long-term trend lines fitted. Overall, the coefficients of interest vary little when I introduce either lead variables or additional lag variables, as shown in Figure 2.7, and remain statistically significant when I add either lead variables or additional lag variables. The main effect of introducing irrelevant variables is to inflate the standard errors on the coefficients of interest, as shown by comparison between column 1) and columns 2) and 3) of Table 2.5. For this reason, I exclude the lead and additional lagged variables in the rest of the analysis.

[17]Results available on request.

Results by subsample

In Table 2.6, I show the main specification across different subsamples. In column 1), I show the results from the main specification for comparison. In column 2), I show the results from an alternative geographical specification which is cropped further South on the Upper Mississippi. This alternative geographical specification is based on an engineer's informal assessment that this lower point represents the cut-off point of bridge structures that represent major civil engineering works (Weeks, n.d.). The results are slightly smaller, but consistent and statistically significant.

In columns 3) and 4) I show the results from repeating the main analysis for counties which never acquire a bridge, and counties that ever acquire a bridge, respectively. The results are extremely consistent across the two specifications. This test rules out the possibility that bridge construction in response to anticipated county-level growth can explain the main results.

In columns 5) to 7) I repeat the main analysis separately for counties on the Ohio, the Upper Mississippi and the Lower Mississippi. The number of counties included in each of these subsamples is much smaller than the pooled sample (69, 66 and 45, respectively). For the Ohio and Upper Mississippi, the estimated coefficients are smaller, apparently consistent in sign and timing, but not statistically significant in the subsamples. The estimated coefficients are largest for the Lower Mississippi, consistent with the greater cost and lower density of bridge construction there. However, these results offer reassurance that the result is consistently produced across the sample, and unlikely to be driven by outliers.

Controls for lagged population density

It is possible that the effect of transport infrastructure on population growth could be amplified by unrelated agglomeration effects if, for example, transport infrastructure leads to an initial small increase in population growth, after which other agglomeration effects 'kick in'. Similarly, the effect of transport infrastructure could be attenuated if dispersal effects take effect. To assess whether this is the case, I respecify the equation in first differences (using Equation 2.3), which enables me to control for functions of lagged population density. Controlling for lagged population density effectively narrows the comparison to places with a similar size at a given moment in time. This analysis is motivated particularly by the work of Michaels, Rauch, and Redding (2012), who find a positive correlation between initial population density and subsequent population growth among areas with intermediate population densities, suggestive of agglomerative forces, in the historical United States. This analysis therefore contributes to understanding the mechanisms through which infrastructure acts on population growth.

In Table 2.7, I first compare the results obtained moving between the level specification described in Equation 2.1 and the difference specification described in Equation 2.3, and show that they are comparable. In column 1) I restate the estimates from the main specification. In column 2) I report the coefficients resulting from the difference specification described in Equation 2.3. Note that the coefficients in column 1) can be interpreted as the cumulative effect of a change in distance to a bridge j periods ago, while the coefficients in column 2) can be interpreted as the effect on growth at time t of a change in distance to a bridge j periods ago. In column 3), I report the cumulative sums of the coefficients reported in column 2). The estimates in column 3) measure the same cumulative effect as the main equation, specified in levels; the estimates are slightly larger, but within the confidence intervals of the estimates in column 1)[18].

In columns 4) and 5), I include controls for lagged population density. I follow Michaels et al. (2012) and include a cubic function of lagged population density in which I either fix the coefficients (column 4) or allow them to vary with time (column 5). I do not report the coefficients on lagged population density but I find negative coefficients on the linear and cubic terms, and a positive coefficient on the quadratic term, consistent in sign with the results from Michaels et al. (2012). The results in columns 4) and 5) show that the coefficients of interest change little when I introduce the controls for lagged population density.

The analysis in Table 2.7 shows that proximity to a transport route predicts increased population growth among counties with a similar population density at the start of the decade, with little or no reduction in the coefficient, indicating that the mechanism for increased population growth is almost exclusively attributable to the infrastructure itself and is independent of other agglomeration or

[18]The specification described in Equation 2.3 is less robust to specification tests, probably indicating a greater sensitivity to outliers, which is why I prefer Equation 2.1 as the main specification.

dispersal effects. Since population density is correlated with bridge construction, this analysis also discounts another possible alternative explanation for the short-run results i.e. that the differences are driven by differences in population growth with time across counties with initial differences in population density.

Further robustness checks

In all specifications so far, the results have controlled comprehensively for county level unobservables with a county fixed effect (accounting for different county sizes and starting populations), and a county quadratic trend (allowing each county to have its own intrinsic growth rate, and first order trend in growth rates). Table 2.8, column 2, shows that the choice of a quadratic functional form is not critical; replacing the quadratic functional form with an equivalently conservative two-part linear spline yields similar, slightly larger point estimates.

A further potential concern about the results is that bridge construction could be correlated with other time-varying unobservables which vary geographically at similar scales. For this to be an empirical concern, the timing of these unobserved changes would need to be sufficiently tightly correlated with the opening of bridges in order to create the same discontinuity in population growth. An example would be if states introduced subsidies to businesses at the same time as bridges opened. In order to address this potential concern, I test the sensitivity of the results to allowing the overall time trends to vary by river, smoothly over geography, or by state. Allowing the time controls to vary by geography absorbs the variation of interest, and likely increases attenuation as a result of measurement error.

The results remain statistically significant when I allow the time trends to vary by river (Table 2.8, column 3) and are consistent in sign and timing when I allow the time trends to vary smoothly by geography (using a quadratic polynomial in the county centroids in column 4)) or by state (in column 5)). The results are further attenuated by allowing the controls to vary further by geography — for example by interacting the year dummies with a cubic polynomial, rather than a quadratic polynomial. The results from this robustness check do not allow me to completely rule out an alternative explanation whereby state- or geographically-varying short-term shocks in population growth and bridge construction influence the main results, but these shocks would have to be precisely matched in timing against bridge construction and vary on a time scale that is not captured by the county level controls for unobservables, in order to generate these results.

Table 2.8, columns 6 to 8, show that the results are not sensitive to the choice of start or end year.

Heterogeneity of impacts

By place of birth During the first half of the study time period, up until the changes in policy before and during the Great Depression, there were extremely high rates of foreign immigration to the United States. For external validity, it is natural to ask whether these results are driven by foreign immigrants' decisions about where to settle. I obtain data on the native and foreign-born population from the ICPSR and IPUMS datasets, as described in the Appendix. The data is not available for all counties at all times, but the results change very little if I include only complete years. In 1860, counties in the study had a mean proportion foreign-born of 15%, a figure which reduces to 7% in 1910, 6% in 1960 and 1% in 1990.

Table 2.9, column 1) shows the main analysis. Column 2) shows the results for the foreign-born population only; column 3) shows the results for the native-born population only. The foreign-born population apparently responds more quickly and more strongly than the native-born population, although this result is less robust than the equivalent result for the native-born population, as I exclude a substantial fraction of observations for which zero foreign population is observed. However, the results for the native-born population are only slightly smaller than for the population as a whole. This suggests that the overall results cannot be driven only by new immigrants' decisions about where to settle.

Road vs rail The main analysis treats distance to a road bridge and distance to a rail bridge as equivalent. In columns 3) and 4) of Table 2.9, I show the results separately for distance to a road bridge and distance to a rail bridge. The effects are largely similar, and fractionally larger for rail, but the results for rail may be less robust, as the vast majority of new rail bridges are already constructed by 1920, and variation in distance to a rail bridge thereafter only stems from rail bridge closures.

Early vs late The main analysis measures the average effect across the entire study period. In columns 5) and 6) of Table 2.9, I report the results from allowing the effect to be different for observations in the first half of the study period (up to and including 1920) versus the second half of the study period (post 1920). I estimate these differences by interacting the lagged changes in distance to a bridge with a dummy for whether the observation belongs in the first or second half of the observation period. This analysis provides a rough, first order estimate of whether the effects change substantially over time; the results are similar, suggesting that to first order the results are consistent over time.

By race and gender Transport infrastructure has been considered in the past as key to understanding internal patterns of migration (see e.g. Black, Muszynska, Sanders, Taylor, and Taylor, 2011). The response in the male and female population is similar, although the male response is fractionally faster than the female response. The fraction male peaks around the time of bridge construction, which may reflect the workforce involved in bridge construction. The results when broken down by race are inconsistent; the response appears faster and stronger among the non-white population, but there is an overall increase in the fraction white. The results are probably sensitive to the inclusion or exclusion of observations with zero non-white population[19].

2.5 Long-Run Impacts

Transport infrastructure may have persistent, long-run impacts on patterns of population settlement and economic activity. However, the empirical strategy I presented in the previous section measured short-run impacts. The identifying assumption — that timing of bridge construction was uncorrelated with population growth residuals within a window around the timing of bridge construction — only enabled me to estimate consistently the impacts within that window; long-run impacts were absorbed by the quadratic trends fitted in each county. Long-run impacts are of intrinsic interest, but are also important in the context of the preceding analysis, because the

[19] Results available on request.

estimated short-run impacts are biased downwards if the long-run impacts are of the same sign, but could potentially be biased upwards if the long-run impacts were of the opposing sign. This could be the case if, for example, early growth reduced the likelihood of later growth by locking in a particular urban structure that was later on less conducive to growth.

In this section, I present estimates of the long-run impacts of transport infrastructure on population density that are of the same sign as short-run impacts, and consistently larger in magnitude.

2.5.1 Empirical Strategy

To measure long-run impacts, I use an instrumental variables approach to separate out variation in the location of bridges that is otherwise uncorrelated with other factors that influence local population densities or growth rates. I focus on discontinuities in the flow rate in the river where a tributary — or smaller river — flows into the main stream.

The volumetric flow rate is a good proxy for the difficulty and cost of bridge construction. Bridge costs are strongly convex in span length, because shorter spans can be crossed using cheaper technologies. The maximum required span of a bridge increases with: channel width, because a wider river requires a wider crossing; river depth, because this increases the cost and difficulty of constructing piers in the river; and river velocity, for the same reason. Either one or all of these three parameters — width, depth and velocity — must increase when the volumetric flow rate in the river increases.

I exploit discontinuous increases in the flow rate in the river at confluences with major tributaries. In general, the flow rate in the Mississippi and Ohio rivers increases monotically from the headwaters to the river mouth[20]. Where the flow rate increases because of rainfall runoff over land or because of small streams joining the main river, the increases in flow rate are approximately smooth. However, the changes in the flow rate are discontinuous at points where a large tributary joins the main stream, creating an abrupt change in the cost of crossing the river.

Bridges are often located just upstream of the point where a tributary joins the main stream,

[20] This is true of almost all rivers, although exceptions do exist where losses from evaporation and infiltration are larger than gains from runoff.

generating corresponding local discontinuities in the likelihood of bridge construction, even if this entails the construction of two shorter bridges over the two separate streams. Figure 2.8 illustrates this, showing the confluence of the Mississippi and the Ohio at Cairo, Illinois; bridges preferentially cross the two smaller streams of the Upper Mississippi and the Ohio, rather than the single, larger stream of the Lower Mississippi downstream of the confluence. As a result, there are three bridges in the immediate upstream neighbourhood of the confluence, and no bridges for approximately 100km downstream. This effect is shown more generally in Panels a) to c) of Figure 2.9; the flow rate increases sharply at the tributary, resulting in a sharp reduction in the local probability of bridge construction, and a corresponding increase in distance to a bridge.

The confluence of a tributary and a river is in itself a place which attracts human settlements, as it forms a natural transport hub, as the place where two natural water transport routes — the river and its tributary — coincide (see Fujita et al. (2001)). Panel d) shows this effect; population density increases from both directions with reducing distance from the tributary, but the upstream-downstream asymmetry at the tributary remains. In my empirical analysis, I control for distance to the nearest tributary, and use an indicator for whether the section of the river is upstream or downstream of the nearest tributary as an instrument for distance to a bridge. The identification strategy focuses on the discontinuities at the tributary confluence, so I also include interaction terms between the upstream indicator and distance to the nearest tributary as control variables.

Since the flow rate generally increases upstream to downstream — with corresponding gradients in bridge construction and distance to a bridge — I also include distance from the river mouth, interacted with river dummies, to avoid confounding the upstream-downstream comparison with overall north-south trends. Conditional on these controls, I argue that tributary confluences create natural experiments in bridge location, since there is no reason to think that location just upstream or just downstream of a tributary should influence population density through any other channel except the location of bridges.

Local discontinuities around the tributary confluence are attenuated at the county level. I focus instead on census tracts, the next level of aggregation available, selecting census tracts for which any part of the tract is within 10km of the river; I also include tracts which are completely enclosed by tracts meeting the preceding criteria. Census tracts are constructed to have an average population of 4000 individuals. Figure 2.10 illustrates the distribution of the census tract level observations, relative to tributary confluences. Where population densities are highest, there are greater numbers of observations, particularly just upstream of the tributary, which is the area of interest of this study.

Formally, I use the following as the estimating equation in this section:

$$y_{i,2000} = \alpha + \Omega X_i + \pi dist_{i,2000} + \epsilon_{i,2000} \tag{2.4}$$

where $y_{i,2000}$ is the outcome variable in census tract i in the year 2000, and $dist_{i,2000}$ is the distance to a bridge in the year 2000. π is the coefficient of interest. X_i is a vector of control variables which includes: nearest-tributary fixed effects, a quadratic function of distance from the nearest tributary and its interactions with the upstream indicator, and distance from the mouth of the river, interacted with an indicator for each river. Ω is the vector of coefficients on these control variables. I focus on log population density as the main outcome variable of interest.

Since infrastructure was built where people lived, or where expected to live, the coefficient π is estimated with bias in the cross-section. I therefore instrument for distance to a bridge using the following first stage equation:

$$dist_{i,2000} = \alpha + \Omega X_i + \gamma \mathbb{1}\left(upstream = 1\right) + \nu_{i,2000}$$

I construct the instrument using data on river reaches from the National Hydrology Dataset. I define a major tributary as a node connecting two reaches at which the local increase in flow is greater than 7.5% of the upstream flow rate or 10,000 cubic feet per second. For each reach, I calculate the distance in river km to the nearest tributary, and the distance between the river reach centroid and the nearest bridge. I then assign census tracts to the nearest river reach[21]. I drop the

[21]This is not quite equivalent to the approach taken in the previous section, where I used the distance between

ends of the rivers, so I only include tributaries where I observe river reaches both upstream and downstream of the tributary confluence[22].

Using a cross-sectional instrument to estimate a dynamic process is problematic in the presence of dynamic effects. Present-day population in a given location is influenced by the full history of access to transport infrastructure in the location, and not just by current access. However, the instruments are fundamentally cross-sectional in nature, as they relate to permanent geographical features[23]. This analysis therefore essentially treats distance from a bridge in the year 2000 as a proxy for the cumulative history of distance to a bridge at a given location over time. In formal terms, the term $\pi dist_{i,2000}$ in Equation 2.4 approximates the cumulative effects on population in the year 2000 of distance to a bridge at all previous times j in the area's history or $\sum_{j=0}^{\infty} \pi_j dist_{i,2000-j}$. Recall from Table 2.2 that at the county level, mean distances to a bridge fall very little after the year 1940.

For simplicity, I limit the sample to census tracts matched to river segments within 50km of a tributary[24]. Table 2.10 shows the cross-sectional relationships between the outcome variables and distance to a bridge. In columns 1) and 2), I show that the cross-sectional relationship between population density and distance to a bridge does not alter significantly when I reduce the sample to census tracts within 50km of a tributary, suggesting that limiting the sample in this way does not make a difference in terms of the relationship of interest. After limiting the analysis to census tracts matched to river segments within 50km of the nearest tributary, the sample comprises a total of 28 tributary 'neighbourhoods', to which river segments are assigned to be either upstream or downstream of the nearest tributary. Since population density is strongly spatially correlated between nearby census tracts, I cluster standard errors throughout the analysis at the level of the upstream-nearest tributary interaction, resulting in 56 clusters. This avoids mistakenly interpreting

the county centroid and the nearest bridge, but this approach makes better intuitive sense when there are no fixed effects in the analysis to remove the raw effect of proximity to the river from the measure of distance to a bridge.

[22]Full details included in Appendix A.

[23]At least, permanent over the time frame of the study; river locations do evolve considerably over longer frames.

[24]Including the full sample does not change the nature of the relationship around tributaries, but a more complex specification is required to isolate the local effects around the tributary.

an upstream-downstream difference as a causal relationship, when it is in fact the result of chance. In column 3) of the same table, I show that the correlation between population density and distance to a bridge is reduced when I include the geographic controls, consistent with spatial correlation in both variables that does not reflect a causal relationship.

For the upstream-nearest tributary distance interactions to be valid instruments for distance to a bridge, they must first be reasonably strong predictors of distance to bridge. In column 4) of Table 2.10, I show that the relationship in the first stage is consistent with intuition: the distance to a bridge is shorter upstream of a confluence, and the relationship diminishes with distance to a bridge. I do not report the coefficients on distance from a tributary here, but distance to a bridge locally reduces with distance from a tributary, downstream of the confluence; the intuition, as seen in Figure 2.9 is that increasing the likelihood of bridge construction in one place decreases the likelihood of bridge construction elsewhere, as the marginal benefit of constructing a bridge is less when there is a substitute bridge nearby. The F-statistic on the instrument, the upstream indicator, is 17.6, which is well above the value of 10 suggested as a rule of thumb for a sufficiently strong first stage (Angrist & Pischke, 2009).

The instruments must also satisfy the monotonicity assumption i.e. being upstream of a tributary should affect the likelihood of bridge construction and therefore distance to a bridge in the same way everywhere; there should be no places in which being upstream of the tributary makes it less likely that a bridge is constructed. It seems reasonable to suppose that this is the case as it is difficult to construct an alternative scenario where a local increase in flow rate would make construction of a bridge more likely.

In column 5) of Table 2.10, I show the reduced form relationship in the sample within 50km of a tributary confluence; consistent with increasing distance from a bridge resulting in a lower population density, being upstream of a tributary confluence is associated with a locally increased population density. Figure 2.9, panel d) shows the same relationship.

Finally, the instruments must satisfy the exclusion restriction; conditional on the geographical controls, they must be uncorrelated with population density through any other avenue than distance to a bridge. In formal terms:

$$E\left(\epsilon_{i,2000}|X_i, \mathbb{1}\left(upstream = 1\right) = 0\right.$$

In other words, being upstream of a tributary, rather than downstream, must not have any other effect on population density, except to alter the likelihood of bridge construction. Three possible objections occur. First, the upstream-downstream comparison might be biased by differential likelihoods of flooding; rivers tend to flood upstream of confluences, as the river backs up and overflows its banks. Increased flood risk would tend to bias my estimates downwards, although if increased flood risk also influences sediment deposits and thereby increases agricultural productivity, this would tend to bias estimates downwards. Second, the location of ferry crossings might also be influenced by location relative to a tributary confluence. However, the effect might act in the opposite direction; if loading costs are a large component of the cost of crossing a river by ferry, ferry crossings might tend to locate downstream of tributaries; this would tend to bias my estimates downwards. Third, the location of river ports might vary around tributaries; river ports might be located downstream of the confluence, rather than upstream, as river traffic going in both directions must pass downstream of the confluence. In general this would also tend to bias my estimates downwards.

In support of the claim that the instruments satisfy the exclusion restriction, I show that the upstream-downstream asymmetries emerge only after the era of bridge construction begins. I am able to match a subset of year 2000 census tracts to 1840 counties. As a result, for these census tracts, I have a measure of county-level population density in 1840, prior to the construction of any bridges on these rivers. For the remainder of the census tracts in the year 2000 sample, county boundaries were not defined in 1840 and census data from 1840 is therefore not available.

In Figure 2.11, panel a) and Table 2.10, column 7), I show that the asymmetries around the tributary confluence are not present in 1840, and that if anything, population density is locally higher just downstream of the tributary confluence. In Figure 2.11, panel b) and column 6) of Table

2.10, I confirm that this is not a product of the change in sample associated with limiting the analysis to census tracts that can be matched to 1840 counties; the estimated reduced form relationship between population density in 2000 and the instruments is larger in magnitude, although less precisely measured, in this subset of census tracts. Measurement error introduced by assigning county-level variables to census tracts might partially explain the absence of discontinuities in the 1840 data, but the reduction in magnitude of the coefficient (from 0.96 to 0.02) is quite compelling.

2.5.2 Results

In Table **??**, I show a series of estimates of the impact of distance to a bridge on population density. All estimates show the expected sign, are statistically significant, and larger in magnitude than the estimated short-run impacts. In column 1) of Table **??**, I show the main specification. These results suggest that a county which is 50% closer to a bridge has a population that is 25% higher in the long run. The point estimate is lower than the corresponding, raw correlation from the cross-section, without controls (Table 2.10, column 2), reflecting the expected positive selection effect.

Given the comprehensive set of nearest-tributary fixed effects and the fact that most bridges connect multiple states, it is unlikely that the result could be driven by state-level differences in infrastructure construction policy and population growth. However, I check whether the analysis is robust to the inclusion of state-level fixed effects in column 2); the results change very little, either in magnitude or precision.

For the subset of the census tracts for which a matching county existed in 1840, I can test whether the results are altered by controlling for population density at the county level in 1840. The reduced sample excludes the northernmost reaches of the Upper Mississippi, where censuses were not carried out in 1840. As we might expect, given the pattern of the short-run impacts described in the previous section, excluding the Upper Mississippi increases the estimates; column 3) shows the same regressions as column 1), but for the reduced sample for which 1840 population data is available. Column 4) shows the results when a control for 1840 population density is included. The correlation between log population density in 1840 and in 2000 is close to unity, reflecting the general persistence of patterns of human settlement. However, the estimated impact of distance to a bridge alters very little.

These results are consistent with the hypothesis that changes in access to manmade transport infrastructure result in long-run, persistent effects on the geography of human settlements. The estimated long-run impact is almost an order of magnitude larger than the estimated short-run impacts, and a test of whether the estimated long-run impact is equal to the estimated short-run coefficient is rejected with p-value 0.03. However, it is worth noting that there are several differences between the two analyses that may contribute to explaining this difference in magnitude. First, as noted, when long-run impacts are of the same sign as short-run impacts, these will tend to bias down the estimated short-run impacts in the analysis presented here. Second, the 'long-run' sample is more tightly fitted to the river, and more geographically disaggregated. Either of these factors might reduce measurement error and result in larger estimates. Third, the instrumental variables approach estimates the local average treatment effect at low distances from a transport route, while the short-run approach estimates the treatment effect at a broader range of distances. If there are important nonlinearities in the response, this might explain the difference in magnitude in the coefficients. Finally, although I previously showed that the short-run impacts were direct, other agglomeration effects and economies of scale may 'kick in' on a longer time-scale.

2.6 Persistence

The proposed mechanism by which transport infrastructure influences long-run growth patterns is by 'locking in' patterns of human settlement and economic activity, as a result of sunk costs either directly in the infrastructure capital itself, or in waves of complementary private and public capital. In this section, I will provide suggestive evidence regarding this mechanism. First, I describe one of the most striking features of the bridge dataset: the persistence of crossings in the same location across time. Around 80% of early crossings persist to the present day, in that there is still a bridge in the location. However, this is not driven by sunk costs in the bridges themselves; the actual structure of the bridge is rebuilt on average every 45 years. Second, I examine the age profile of the housing stock in the long run, and show that places that are closer to transport routes have a larger fraction of older housing stock, consistent with a persistence mechanism involving multiple waves of investment in physical capital.

2.6.1 Transport infrastructure

Table 2.12 shows some measures of persistence. Crossings built prior to 1860 have a bridge lifetime, measured as the time to date that a bridge has existed in the same site, of 157 years; crossings built between 1860 and 1880 have a bridge lifetime of 131 years. Around 80% of these early crossings persist to the present day, in that there is still a bridge in the location. This persistence vastly exceeds the lifetime of a given structure; the bridges in location today are in many cases quite different from the original bridges constructed, since almost all bridges built before 1880 have undergone at least one rebuild, and many have undergone multiple rebuilds.

To construct the data on the number of rebuilds, I used textual sources describing the history of the bridges (see the Appendix for further details), defining a rebuild as a substantial replacement of or alteration to a large part of the bridge structure.The number of rebuilds is measured with significant error, since there is no centralized source of data on rebuilds, and it is difficult to define a rebuild conclusively. The number of rebuilds included in the dataset is probably therefore a lower bound on the true number of rebuilds, particularly for crossings that have undergone multiple rebuilds.

Using this measure, the average lifetime of a given structure is around 45 years. Bridges are rebuilt or replaced in some cases due to weather damage from extreme floods or winds, but are often strategically rebuilt to reflect an increase or change in use or obsolescence of the current structure[25]. These measures somewhat underestimate the true persistence in bridge location, as many of the bridges that are closed or destroyed in the sample are also replaced very nearby[26].

---

[25]This seems consistent with typical design lifespans of 50 years — noting however that a design life of 50 years means designing to withstand the most extreme loading conditions expected within a 50 year time period, and does not reflect a particular expectation that the structure will only last 50 years.

[26]I do not break down the results by rail and road here, but 90% of road bridges ever built persist to the present day, in that a structure stands in the same place. In comparison, only 75% of rail bridges persist to the present day. However, the persistence of contemporaneous road and rail bridges is roughly equal, and the difference over all

The extent of rebuilding and the disparity between structure and crossing lifespans strongly suggests that sunk costs associated with the bridges themselves cannot explain the persistence of bridge crossings. However, generations of physical capital overlap; by the time a bridge reaches obsolescence, other investments in physical capital, both public (e.g., connecting highways) and private (e.g., housing stock, factory location) have been made. Conditional on other capital investments made during the bridge lifetime, the net marginal benefit of rebuilding the obsolete bridge is evidently much higher than the net marginal benefit of relocating the structure. Other mechanisms may be important, such as self-reinforcing beliefs — investment decisions are taken assuming the long run persistence of the bridge crossing, which in turn increases the value of the bridge crossing — and local politics; the incumbent beneficiaries of a bridge crossing may be reluctant to relinquish their local advantage.

2.6.2 Housing stock

If investments in other complementary components of physical capital play a role in maintaining persistence, then there should be evidence of a historical sequence of capital investments in the long run. To examine this, I look at the distribution of structure age in the housing stock in the long run. Figure 2.12 illustrates the main results. I estimate the effect of distance to a bridge on the age of the housing stock using Equation 2.4. It may seem counterintuitive that I find a causal relationship between distance to a bridge in the year 2000 on the history of the building stock, but distance to a bridge was largely determined early on in the study period, and as previously noted I am taking distance to a bridge in the year 2000 as an approximation of distance to a bridge over time. Places that are closer to transport route have higher percentages of the oldest housing stock, and lower percentages of the newest housing stock, than places that are further away. These results are also shown in Table 2.13, with additional nonparametric evidence in support of this relationship shown

bridges ever constructed may reflect the fact that no new rail bridges have been constructed since 1940. Reliable data on the use of rail crossings is not consistently available. However, informal estimates collated by bridge enthusiasts and published online suggest that many of these crossings, although nominally still in use, are very infrequently used, with daily train counts numbering in single digits. If this data were available, it would likely contribute to a picture of declining rail transport services in infrastructure across the US. It seems probable that many of the rail bridges listed in the dataset will not still be in use in twenty years time.

for each age category of housing stock in Appendix Figure 2.13. Overall, these results are consistent with transport infrastructure helping to coordinate generations of investments in physical capital, including housing, leading to persistence in infrastructure location and contributing to persistence in human settlement location.

2.7 Summary and Discussion

This paper has focused on understanding the impact of access to transport infrastructure on spatial patterns of human settlements. Using historical census data and a novel dataset containing information about all the bridges ever constructed over the Mississippi and Ohio rivers, I exploit quasi-experimental variation in the timing and location of bridge construction to separate out the causal impact of transport infrastructure on population growth.

I show that the population of counties which experience a 50% reduction in distance to a bridge grows by an additional 3% over the following 30 years. At the median population growth rate, the county would have grown by 15% over the same time period. By showing that the effect persists when the comparison is narrowed to counties with similar population densities at the start of a decade, I show that the measured effect is largely attributable directly to the construction of infrastructure, and in the short term at least is not significantly amplified by other unrelated forces of agglomeration or economies of scale. This supports the hypothesis that economies of density associated with transport infrastructure can contribute directly to agglomeration effects.

I compare these 'short-run' results to long-run instrumental variables estimates, exploiting discontinuous local changes in the likelihood of bridge construction around tributary confluences — places where a smaller river flows into the main stream. Using this approach, I show that places that are 50% closer in distance to a bridge have population densities that 25% higher, in the long run. The results are consistent with the presence of path dependency, in that local changes in access to transport infrastructure result in long-run differences in population density.

The study covers both rail and road bridges. My results contrast somewhat with an earlier literature on the railroads which concluded that the railroads primarily followed, rather than led, pre-existing patterns of population growth (Fogel, 1964; Fishlow, 1965; Atack et al., 2010). Atack et al. (2010), in particular, concluded that the impact of the railroads on population growth between 1850 and 1860, while positive, was statistically and economically insignificant. The difference between our results may be explained by the presence of spillover effects (which my analysis allows for) and the time frames considered. My results complement the findings of Bleakley and Lin (2012) who studied the long-run impacts of an obsolete natural transport advantage on economic geography; and of Duranton and Turner (2012), and Baum-Snow (2007) who studied the impact of the Interstate Highway system on growth between and within cities respectively; however, my results rely on less restrictive identifying assumptions that these studies, and in particular, avoids relying on the assumption that historical infrastructure is a valid instrument for present-day infrastructure.

Labour and capital mobility in the United States has historically been exceptionally high, particularly during the late 1800s, when there was a large influx of migrants and foreign capital into a region where the spatial distribution of modern economic activity was still being determined. These characteristics make this context particularly interesting to study the question of how transport infrastructure affects the spatial distribution of economic activity, but raise concerns for the external validity of the results. I show that the results are not being driven by the location choices of foreign immigrants, improving the external validity of the results to contexts where internal migration is high, but foreign immigration is low.

The results are important for researchers designing studies that aim to measure the growth or welfare impacts of transport infrastructure provision. Structural approaches (e.g. Donaldson, 2010; Donaldson and Hornbeck, 2013) typically depend on strong assumptions about labour mobility. For example, Donaldson and Hornbeck (2013) assume full labour mobility in their evaluation of the economic impact of the railroads. My results suggest that although labour may be fully mobile in the very long run, there are important frictions which delay this response. Donaldson and Hornbeck (2013) may underestimate the overall economic impacts if they focus on immediate changes in outcomes such as land values. Assuming full labour mobility in this case may result in mistakenly attributing part of the real economic impact to the movement of population, when in fact labour mobility takes several decades to play out.

Further, the results can guide policy-makers and planners. Where labour is mobile, the results suggest that planners should expect changes in population growth rates in response to transport infrastructure over several decades, and should design complementary long-term infrastructure plans accordingly. In the absence of labour mobility, or in the presence of labour market frictions, the corollary to this finding is that transport infrastructure construction may have more heterogeneous effects, and may possibly exacerbate or create inequalities between better- and worse-connected regions.

2.8 Tables and Figures

Figure 2.1: The Mississippi and Ohio Rivers

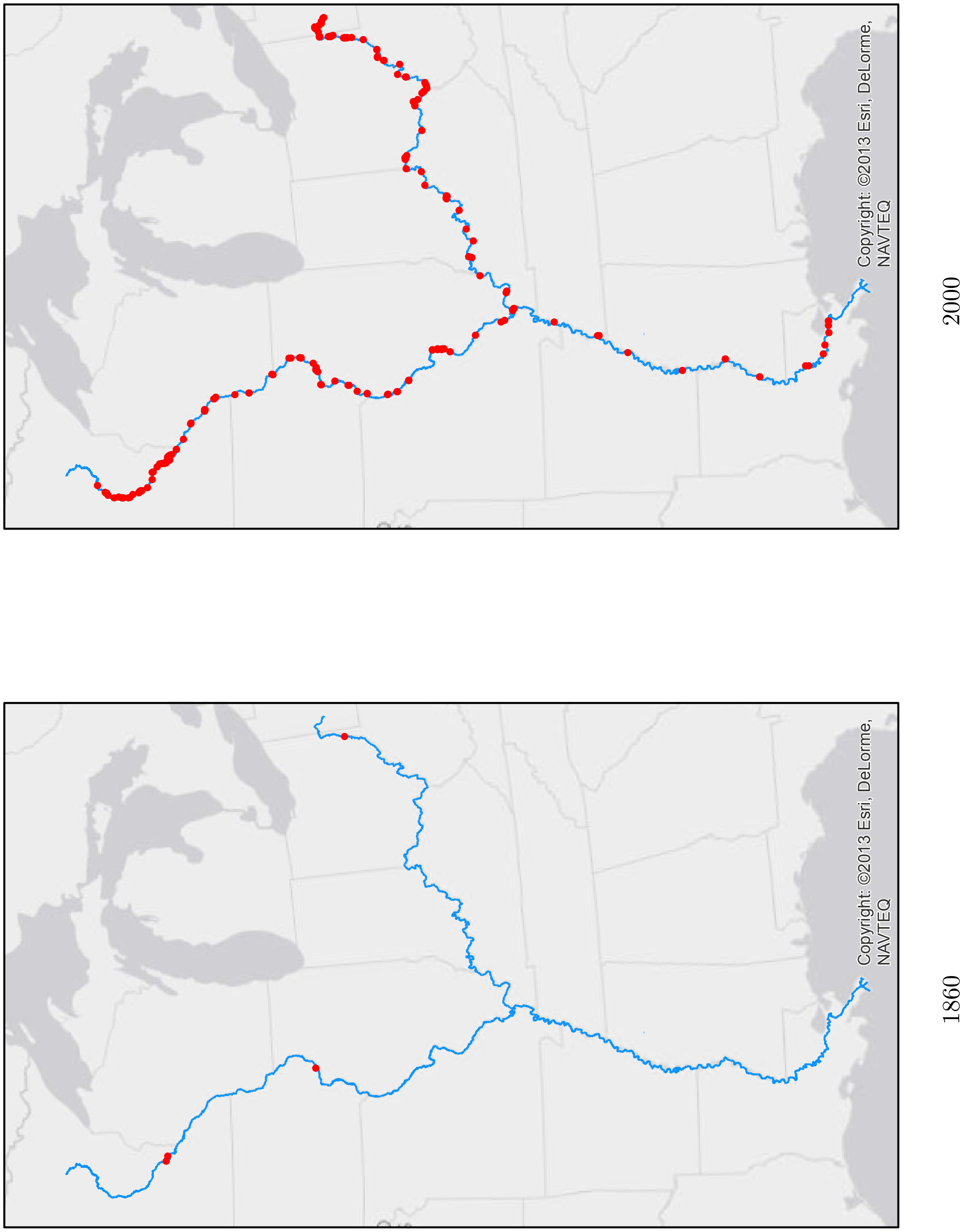
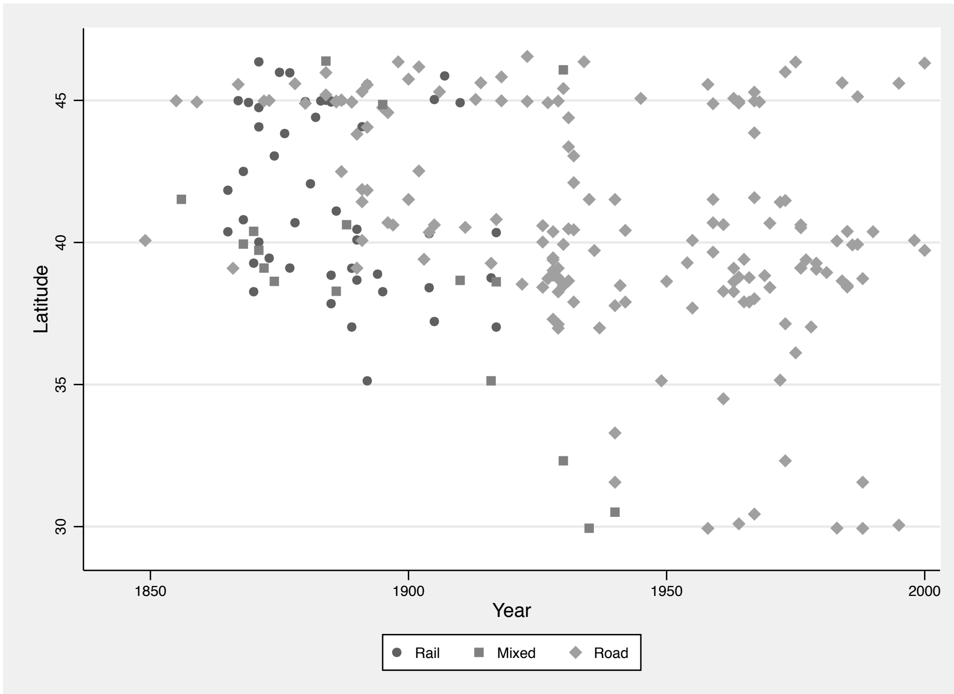


Figure 2.2: Distribution of bridges across the Mississippi and Ohio Rivers

Figure 2.3: Timing of bridge construction



Notes Bridges shown are those included in main analysis. Markers show year of original construction. N = 229.

Figure 2.4: 1860 County Boundaries for counties on the Mississippi or Ohio River

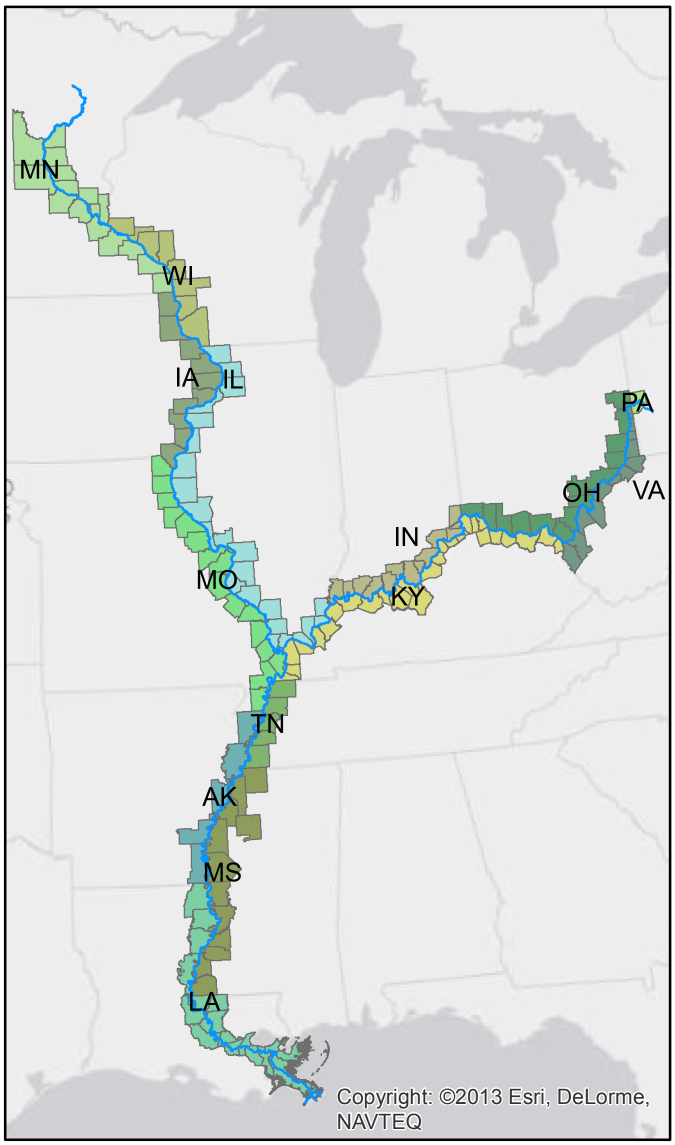


Figure 2.5: Illustration of short-run identification strategy

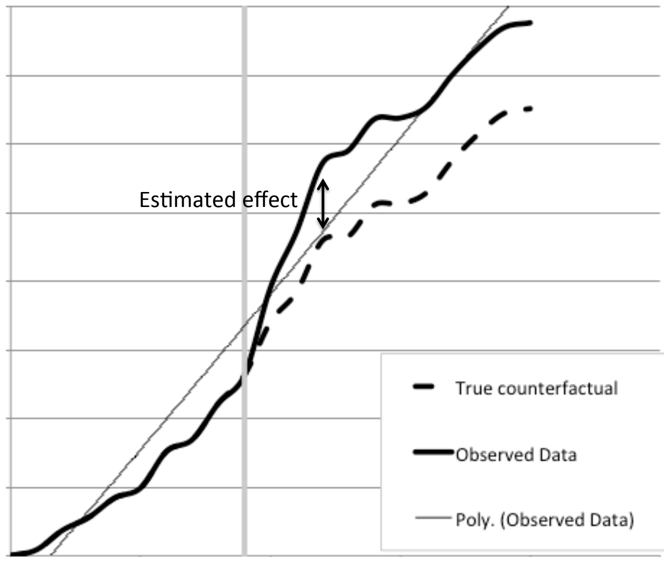
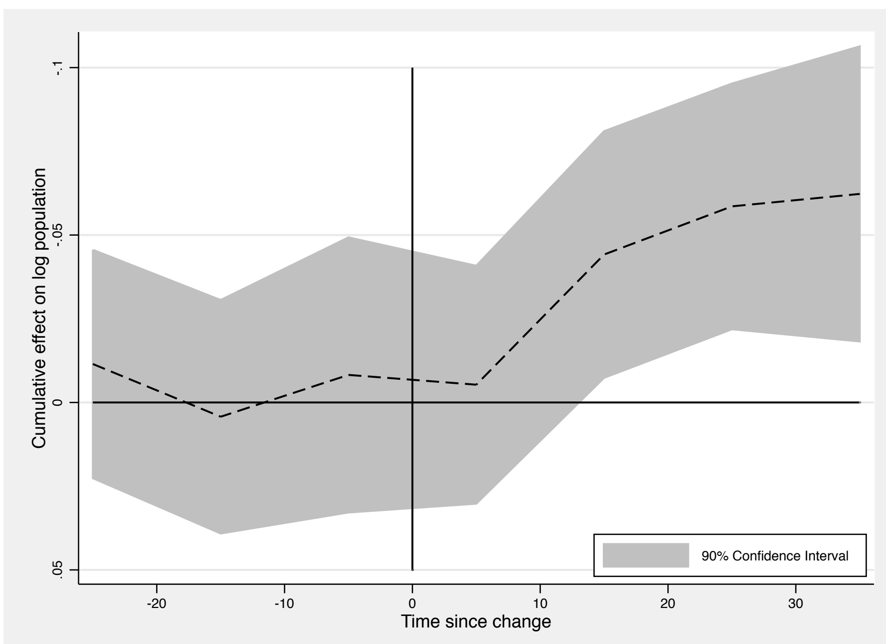
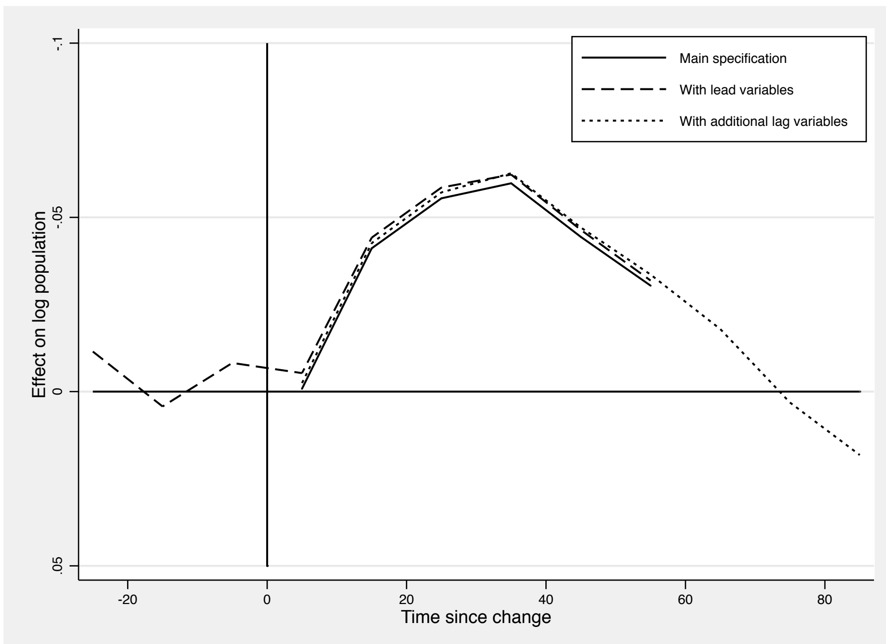


Figure 2.6: Population density before and after a change in distance to a bridge



Notes Coefficients from a regression of log population on lead and lagged changes in log bridge distance, year fixed effects and county quadratic trends. Time zero is defined as the beginning of the decade in which the change in distance takes place. Coefficients on future changes in bridge distance are shown to the left of the black line, and coefficients on lagged changes in bridge distance are shown to the right of the black line. Sample consists of counties on the rivers Mississippi and Ohio, unweighted, using constant 1860 boundaries. Standard errors are clustered at the county level and robust to spatial correlation within a 200km radius. Y-axis is reversed.

Figure 2.7: Effect on main estimates of varying leads and lags included in regression

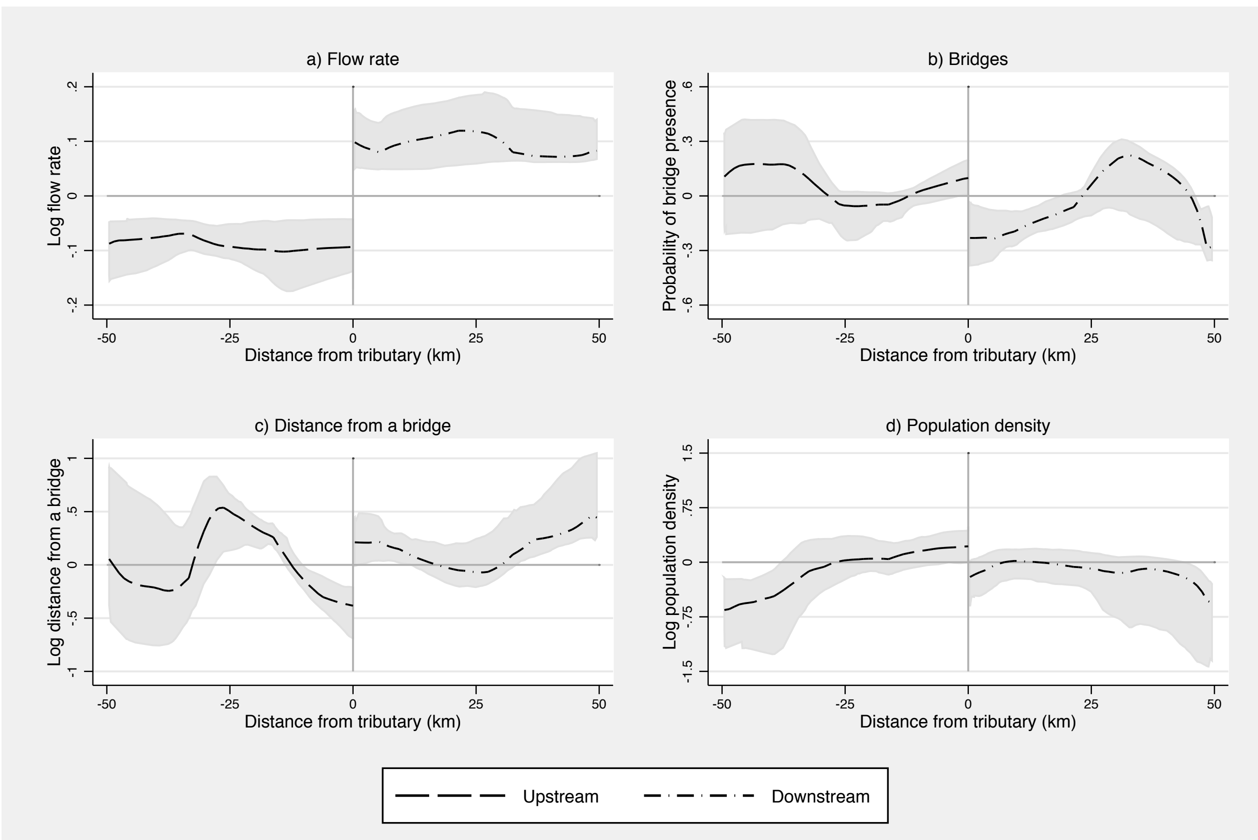


Notes Coefficients from a regression of log population on lead and lagged changes in log bridge distance as indicated, year fixed effects and county quadratic trends. Time zero is defined as the beginning of the decade in which the change in distance takes place. Coefficients on future changes in bridge distance are shown to the left of the black line, and coefficients on lagged changes in bridge distance are shown to the right of the black line. Sample consists of counties on the rivers Mississippi and Ohio, unweighted, using constant 1860 boundaries. Y-axis is reversed.

Figure 2.8: Location of bridges and town at Cairo, IL, confluence of the Mississippi and Ohio

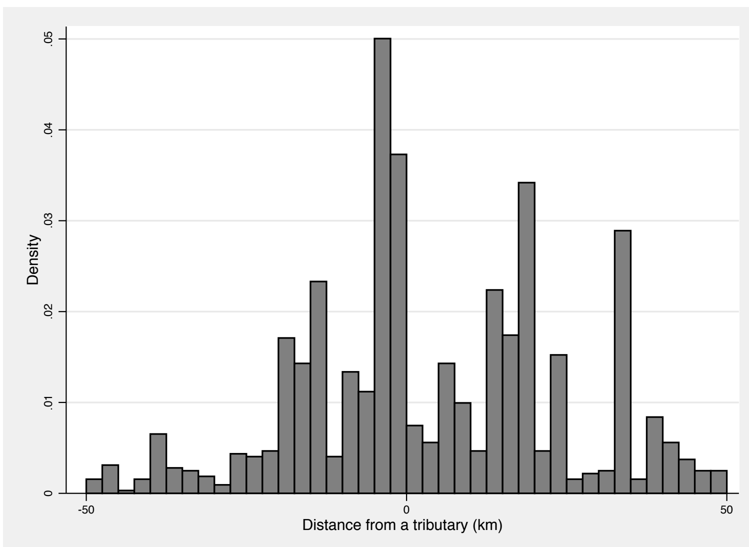
Imagery ©2013 TerraMetrics; Map data ©2013 Google.

Figure 2.9: Bridge location and population density around tributaries (2000)



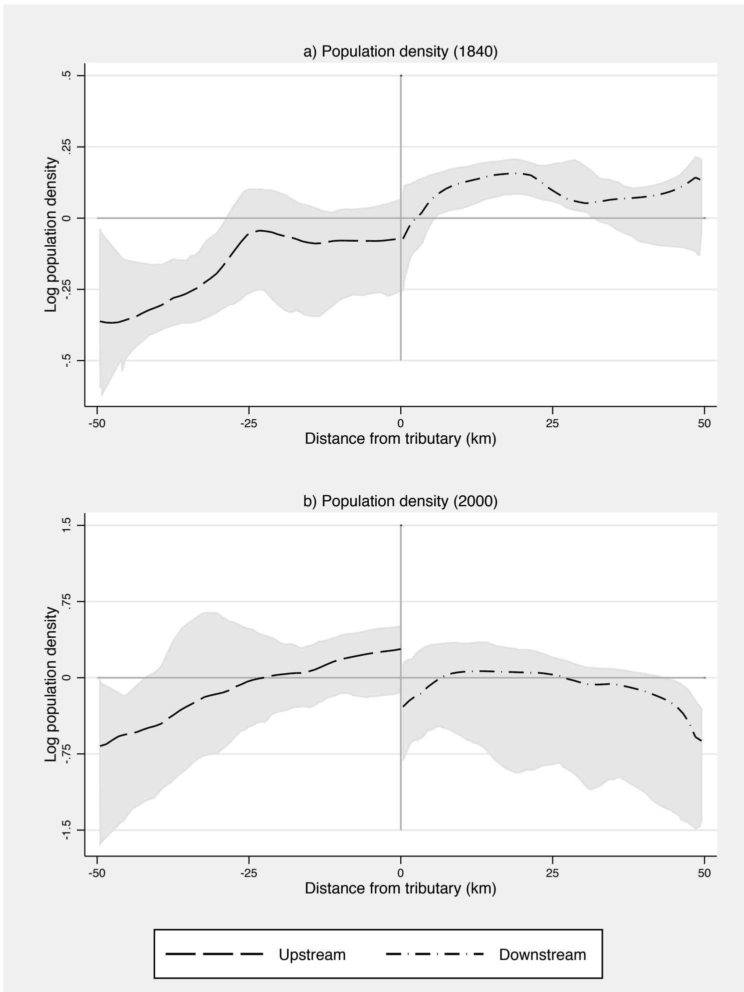
Graphs show results from a local linear regression of the outcome variable on distance from a tributary, for census tracts less than 50km from a tributary confluence. Outcome variables were previously demeaned around the nearest tributary. Upstream tracts and downstream tracts are pooled separately. Bootstrapped 90% confidence intervals are shown in grey, clustered at the nearest tributary.

Figure 2.10: Location of census tracts relative to tributaries (2000)



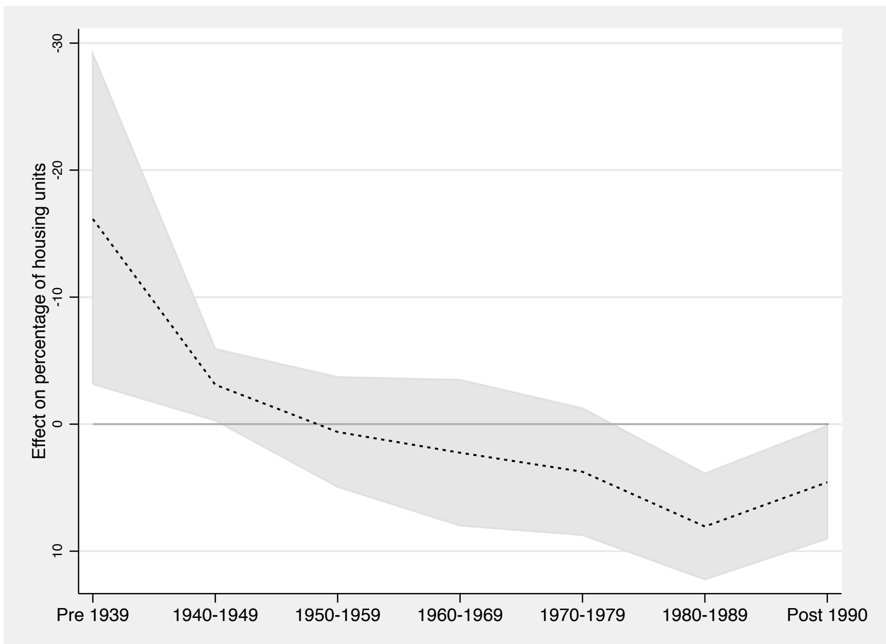
Notes Histogram shows distribution of census tracts relative to local tributary confluences, for all census tracts less than 50km from a confluence. N = 1287.

Figure 2.11: Variation in population density around tributaries: Pre (1840) and post (2000) era of bridge construction



Notes Graphs show results from a local linear regression of the outcome variable on distance from a tributary, for census tracts less than 50km from a tributary confluence that can be matched to an 1840 county. Outcome variables were previously demeaned around the nearest tributary. Upstream tracts and downstream tracts are pooled separately. Bootstrapped 90% confidence intervals are shown in grey, clustering at the nearest tributary.

Figure 2.12: Variation in age of housing stock with distance from a bridge (2000)



Note Graph shows coefficients from a regression of the percentage of housing units of a given age on log distance to a bridge, instrumented by an indicator for being upstream of a confluence, nearest tributary fixed effects and geographical controls. Sample consists of year 2000 census tracts where any part of the tract is within 10km of the Mississippi or Ohio rivers and within 50km of a tributary. Standard errors are clustered by nearest tributary and upstream status ($N_c = 56$). 90% confidence intervals are shown in grey. Y-axis is reversed.

Table 2.1: Summary Statistics: Bridges

	Number built	Rail	Max span length (m)	Total length (m)	Daily traffic c. 2005
Pre-1860	4	0.25	174	415	11950
1860-1880	29	0.83	111	843	13899
1880-1900	49	0.51	121	766	19870
1900-1920	27	0.44	141	1061	9744
1920-1940	40	0.08	189	935	13221
1940-1960	20	0.05	221	1053	28757
1960-1980	42	0.00	190	975	50400
1980-2000	18	0.00	225	762	20865
All	229	0.28	169	923	26105

Note: Data on bridges covers range of counties described in text. Bridges included are those which intersect the county sample. Traffic counts are from 2001-2006 and 33 road bridges have missing data. I dropped traffic data from one bridge where the count was from 1993, and from a rail bridge which appeared to have road traffic data listed in error.

Table 2.2: Summary Statistics: Bridge access by county

	Counties with bridges	No. of bridges ever built	No. of extant bridges	Distance to bridge (km)	
Bridges?	Ever	Ever	Ever	Ever	Never
1860	0.06	0.06	0.06	354	592
1880	0.36	0.52	0.52	162	306
1900	0.56	1.25	1.25	82	143
1920	0.60	1.67	1.67	78	140
1940	0.73	2.26	2.23	25	46
1960	0.85	2.57	2.48	23	44
1980	0.96	3.21	2.94	18	38
2000	0.98	3.49	3.09	17	38

Note: All counties are on the Missisppi or Ohio Rivers. Counties included in the sample are those which completely overlap the bridge dataset. N = 181 of which: 124 counties ever acquire bridges; 57 counties never acquire bridges.

Table 2.3: Summary Statistics: County population

	Population	Population Density /km^2	Annual average growth	Relative variance in log pop. density
Panel A: All counties (N = 181)				
1860	16866	16		0.74
1880	25921	24	2.5%	0.46
1900	35459	33	1.3%	0.44
1920	43126	40	0.3%	0.54
1940	52618	49	0.5%	0.64
1960	67642	64	0.4%	0.83
1980	77766	73	0.7%	0.92
2000	83604	78	0.2%	1.00
Panel B: Counties acquiring bridges (N = 124)				
1860	19686	19		0.93
1880	30773	29	2.7%	0.54
1900	43023	41	1.4%	0.52
1920	53801	52	0.5%	0.65
1940	66754	65	0.6%	0.76
1960	88532	85	0.7%	0.95
1980	101197	97	0.8%	0.98
2000	107888	103	0.2%	1.05
Panel C: Counties never acquiring bridges (N = 57)				
1860	10731	9		0.27
1880	15366	13	2.0%	0.18
1900	19004	15	1.0%	0.14
1920	19902	15	-0.1%	0.13
1940	21864	16	0.3%	0.14
1960	22198	16	-0.2%	0.21
1980	26795	20	0.4%	0.34
2000	30776	23	0.2%	0.45

Note: Counties included in the sample are those on the Missisppi or Ohio Rivers which completely overlap the bridge dataset. Growth refers to the average annual growth in the preceding twenty years, approximated by taking the log difference and dividing by 20. Relative variance in log population density is calculated by taking logs of population density, calculating the variance for a given time and dividing by the same measure calculated for the full sample in the year 2000.

Table 2.4: Probability of bridge construction and location characteristics

		Outcome variable: Indicator for construction of a new bridge				
		(1)	(2)	(3)	(4)	(5)
Already has bridge?	Coefficient	0.143***	0.159***	-0.047	-0.124***	-0.379***
	s.e.	0.021	0.025	0.032	0.046	0.055
	Year FE	No	Yes	Yes	Yes	Yes
	County FE	No	No	Yes	Yes	Yes
	County-specific time trends	No	No	No	Linear	Quadratic

Note: Coefficients from regressions of an indicator for new bridge construction in a county on an indicator for whether or not the county has a pre-existing bridge and year and county fixed effects and county-specific time trends as specified in the table. Sample consists of counties on the rivers Mississippi and Ohio, unweighted, using constant 1860 boundaries. Standard errors clustered by county and robust to spatial correlation within 200km. N = 181, T = 18. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2.5: Cumulative effect on log population following change in distance to a bridge

RHS variable: Change in log bridge distance between times:		Outcome Variable: Log population at time t (1)	(2)	(3)
t+20 to t+30	Coeff.		-0.011	
	s.e.		0.021	
t+10 to t+20	Coeff.		0.004	
	s.e.		0.021	
t to t+10	Coeff.		-0.008	
	s.e.		0.025	
t-10 to t	Coeff.	-0.001	-0.005	-0.002
	s.e.	0.017	0.022	0.018
t-20 to t-10	Coeff.	-0.041***	-0.044**	-0.043**
	s.e.	0.016	0.022	0.018
t-30 to t-20	Coeff.	-0.055***	-0.059***	-0.057**
	s.e.	0.018	0.022	0.022
t-40 to t-30	Coeff.	-0.060***	-0.062**	-0.063**
	s.e.	0.023	0.027	0.028
t-50 to t-40	Coeff.	-0.044**	-0.046**	-0.047*
	s.e.	0.019	0.021	0.025
t-60 to t-50	Coeff.	-0.030**	-0.032**	-0.034
	s.e.	0.012	0.014	0.022
t-70 to t-60	Coeff.			-0.018
	s.e.			0.020
t-80 to t-70	Coeff.			0.003
	s.e.			0.019
t-90 to t-80	Coeff.			0.018
	s.e.			0.020
	N	2715	2715	2715

Note: Coefficients from regressions of log population on lead and lag changes in log distance to a bridge as specified, year fixed effects and county fixed effects and quadratic trends. Sample consists of counties on the rivers Mississippi and Ohio, unweighted, using constant 1860 boundaries. Standard errors clustered by county and robust to spatial correlation within 200km.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2.6: Cumulative effect on log population following change in distance to a bridge: Results by subsample of counties

RHS variable: Change in log bridge distance between times:		Outcome Variable: Log population at time t						
		All (1)	Alternative (2)	Never (3)	Ever (4)	Ohio (5)	U.Miss (6)	L. Miss (7)
t-10 to t	Coeff.	-0.001	0.002	-0.007	0.003	-0.008	0.009	0.023
	s.e.	0.017	0.016	0.024	0.018	0.016	0.044	0.019
t-20 to t-10	Coeff.	-0.041***	-0.038**	-0.047*	-0.037**	-0.020	0.008	0.010
	s.e.	0.016	0.015	0.024	0.017	0.019	0.043	0.025
t-30 to t-20	Coeff.	-0.055***	-0.047***	-0.058**	-0.055***	-0.022	-0.014	-0.043
	s.e.	0.018	0.017	0.025	0.020	0.023	0.043	0.030
t-40 to t-30	Coeff.	-0.060***	-0.050**	-0.059**	-0.063**	-0.010	-0.042	-0.080**
	s.e.	0.023	0.021	0.027	0.026	0.021	0.043	0.031
	N	2715	2640	855	1860	1035	990	675

Note: Coefficients from regressions of log population on lag changes in log distance to a bridge, year fixed effects and county quadratic trends. Two additional lags included beyond those reported in the table. Samples consists of subsamples of counties, indicated in the table and described in the text, on the rivers Mississippi and Ohio, unweighted, using constant 1860 boundaries. Standard errors clustered by county and robust to spatial correlation within 200km.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2.7: Robustness to controls for lagged population density

RHS variable: Change in log bridge distance between times:		Population	Population Growth			
		β	τ	$\Sigma\tau$	τ	τ
		(1)	(2)	(3)	(4)	(5)
t-10 to t	Coefficient	-0.001	-0.027	-0.027	-0.026	-0.021
	s.e.	0.017	0.018	0.018	0.016	0.014
t-20 to t-10	Coefficient	-0.041***	-0.044***	-0.071***	-0.037***	-0.047***
	s.e.	0.016	0.014	0.024	0.012	0.012
t-30 to t-20	Coefficient	-0.055***	-0.017	-0.088**	-0.020	-0.013
	s.e.	0.018	0.017	0.034	0.013	0.011
t-40 to t-30	Coefficient	-0.060***	-0.009	-0.097**	-0.018	-0.016
	s.e.	0.023	0.013	0.041	0.012	0.010
	N	2715	2534	2534	2534	2534
Controls for lagged pop.?		No	No	No	Fixed	Flexible

Note: Coefficients from regressions of log population or growth as noted in the table on lag changes in log distance to a bridge, year fixed effects, county quadratic or linear trends and controls for lagged population density as specified. One or two additional lags included beyond those reported in the table. Sample consists of counties on the rivers Mississippi and Ohio, unweighted, using constant 1860 boundaries. Standard errors clustered by county and robust to spatial correlation within 200km. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2.8: Short-run Results: Further Robustness Checks

RHS variable: Change in log bridge distance between times:		Outcome Variable: Log population at time t (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
t-10 to t	Coeff.	-0.001	0.008	0.005	-0.009	-0.003	-0.081	0.001	-0.014***
	s.e.	0.017	0.030	0.016	0.014	0.014	0.030	0.013	0.017
t-20 to t-10	Coeff.	-0.041***	-0.045*	-0.006	-0.015	-0.011	-0.093*	-0.023***	-0.054***
	s.e.	0.016	0.026	0.016	0.016	0.016	0.031	0.015	0.017
t-30 to t-20	Coeff.	-0.055***	-0.058**	-0.030	-0.023	-0.017	-0.083**	-0.035***	-0.063***
	s.e.	0.018	0.024	0.018	0.017	0.017	0.019	0.019	0.018
t-40 to t-30	Coeff.	-0.060***	-0.067**	-0.046**	-0.023	-0.021	-0.071***	-0.043**	-0.062***
	s.e.	0.023	0.026	0.020	0.017	0.017	0.018	0.017	0.025
	N	2715	2715	2715	2715	2715	2431	2353	1991
	Global time trends	Year	Year	River-year	XY Quad. - Year	State-Year	Year	Year	Year
	Local time trends	Quadratic	Splines	Quadratic	Quadratic	Quadratic	Quadratic	Quadratic	Quadratic
	Study Period	1860–2000	1860–2000	1860–2000	1860–2000	1860–2000	1840–2000	1880–2000	1860–1960
	County boundaries	1860	1860	1860	1860	1860	1840	1880	1860

Note: Coefficients from regressions of log population on lag changes in log distance to a bridge, county quadratic trends (or linear splines, where specified), and global time trends as specified in the table. Sample consists of counties on the rivers Mississippi and Ohio, unweighted, using constant boundaries from the starting year indicated. Standard errors clustered by county and robust to spatial correlation within 200km. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2.9: Short-run Results: Heterogeneity of Impacts

RHS variable: Change in log bridge distance between times:		Outcome Variable: Log population at time t						
		All (1)	Foreign (2)	Native (3)	Road (4)	Rail (5)	Early (6)	Late (7)
t-10 to t	Coeff.	-0.001	-0.093*	0.0	0.011	-0.026	-0.004	-0.002
	s.e.	0.017	0.055	0.017	0.017	0.020	0.024	0.018
t-20 to t-10	Coeff.	-0.041***	-0.181***	-0.045***	-0.024	-0.035*	-0.048**	-0.030
	s.e.	0.016	0.062	0.017	0.015	0.021	0.020	0.021
t-30 to t-20	Coeff.	-0.055***	-0.168***	-0.059***	-0.039**	-0.054**	-0.056***	-0.060**
	s.e.	0.018	0.058	0.019	0.017	0.026	0.021	0.025
t-40 to t-30	Coeff.	-0.060***	-0.149***	-0.050**	-0.048**	-0.044*	-0.057*	-0.070***
	s.e.	0.023	0.052	0.022	0.023	0.026	0.031	0.026
	N	2715	1810	2038	2715	2715	2715	2715

Note: Coefficients from regressions of log population on lagged changes in log distance to a bridge, year fixed effects and county quadratic trends, except in columns 4) and 5) where the RHS variables are lagged changes in log distance to a road and rail bridge, respectively. Results in columns 6) and 7) from a regression of log population on lagged changes in log distance to a bridge and the same controls, where the lagged distance variables are interacted with an indicator for whether the observation is in the first or second half of the time period studied. Samples consists of subsamples of counties, indicated in the table and described in the text, on the rivers Mississippi and Ohio, unweighted, using constant 1860 boundaries. Standard errors clustered by county and robust to spatial correlation within 200km. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.10: Long-run Results: Cross-sectional correlations, first stage and reduced form (2000)

		Log pop. dens. (2000)	Log pop. dens. (2000)	Log pop. dens. (2000)	Log bridge dist. (2000)	Log pop. dens. (2000)	Log pop. dens. (2000)	Log pop. dens. (1840)
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log distance from a bridge (2000)	Coeff.	-0.69***	-0.67***	-0.31***				
	s.e.	0.07	0.12	0.08				
Upstream	Coeff.				-1.21***	0.62*	0.96	0.02
	s.e.				0.29	0.37	0.58	0.12
	N	2417	1285	1285	1287	1285	746	747
	First stage F-stat				17.6			
	Controls	No	No	Yes	Yes	Yes	Yes	Yes
	Tributary F.E.	No	No	Yes	Yes	Yes	Yes	Yes
	Sample	All	50km	50km	50km	50km	1840	1840

Note: Coefficients from OLS regressions of outcome variable listed on variable listed, and nearest-tributary fixed effects and geographical controls as indicated. Controls comprise a quadratic function of distance to the nearest tributary, interactions of this function with the upstream indicator, and pathlength from the river mouth, interacted with river indicators. Full sample consists of year 2000 census tracts where any part of the tract is within 10km of the Mississippi or Ohio rivers. Subsamples are i) tracts within 50km of a major tributary confluence ii) tracts within 50km of a major tributary confluence that can be matched to an 1840 county. Standard errors are clustered by nearest tributary and upstream status ($N_c = 56$ for sample within 50km of tributaries.). *** p<0.01, ** p<0.05, * p<0.1.

Table 2.11: Long-run Results: Estimates

		Outcome variable: Log population density (2000)			
		(1)	(2)	(3)	(4)
Log bridge distance (2000)	Coeff.	-0.51**	-0.48**	-0.78**	-0.76**
	s.e.	0.24	0.23	0.35	0.32
Log population density (1840)	Coeff.				0.97***
	s.e.				0.22
	N	1285	1285	746	746
	First stage F-stat	17.6	15.7	9.7	9.6
	Sample	50km	50km	1840	1840
	State F.E.	No	Yes	No	No

Note: Coefficients from regressions of log population density in 2000 on log distance from a bridge (instrumented by an indicator for being upstream of the nearest tributary confluence); nearest tributary fixed effects; tributary distance and tributary distance squared, and their interactions with the upstream indicator; and pathlength from the river mouth interacted with river indicators. State F.E. and controls for 1840 population density at the county level are included where indicated. Main sample consists of year 2000 census tracts where any part of the tract is within 10km of the Mississippi or Ohio rivers. Standard errors are clustered by nearest tributary and upstream status ($N_c = 56$ for sample within 50km of tributaries.). *** p<0.01, ** p<0.05, * p<0.1.

Table 2.12: Persistence of bridge crossings

	Number built	Rail	Bridge Lifetime	Structure Lifetime	Persistence Probability
Pre-1860	4	0.25	157	28	1.00
1860-1880	29	0.83	131	42	0.79
1880-1900	49	0.51	113	53	0.73
1900-1920	27	0.44	98	77	0.85
1920-1940	40	0.08	77	50	0.80
1940-1960	20	0.05	61	42	0.90
1960-1980	42	0.00	43	36	1.00
1980-2000	18	0.00	25	24	1.00
All	229	0.28	81	45	0.86

Note: Data on bridges covers range of counties described in text. Bridges included are those which intersect the county sample. Bridge lifetime is defined as the time to date that a bridge has existed in the same site. Structure lifetime is the average length of time between rebuilds. Persistence probability is the probability that a bridge still exists in the same location today.

Table 2.13: Investments in housing stock

LHS variable: % of housing units built:		RHS variable: Log distance to a bridge (2000)				
		OLS (1)	IV (2)	IV (3)	IV (4)	IV (5)
Before 1939	Coeff.	-4.03***	-16.15**	-17.38**	-19.08**	-18.99**
	s.e.	1.19	7.89	8.28	8.88	8.79
	N	1275	1275	1275	744	744
1940-1949	Coeff.	-0.48	-3.10*	-3.35*	-2.55	-2.54
	s.e.	0.40	1.71	1.81	2.83	2.83
	N	1275	1275	1275	744	744
1950-1959	Coeff.	-1.02**	0.62	0.61	3.81**	3.80**
	s.e.	0.50	2.62	2.76	1.83	1.83
	N	1275	1275	1275	744	744
1960-1969	Coeff.	0.25	2.25	2.34	6.23**	6.23**
	s.e.	0.28	3.48	3.68	3.10	3.09
	N	1275	1275	1275	744	744
1970-1979	Coeff.	1.62***	3.75	4.07	6.59*	6.56*
	s.e.	0.54	3.03	3.18	3.58	3.56
	N	1275	1275	1275	744	744
1980-1989	Coeff.	1.30**	8.06***	8.74***	2.02	2.01
	s.e.	0.51	2.53	2.64	2.60	2.59
	N	1275	1275	1275	744	744
After 1990	Coeff.	2.36***	4.58*	4.97*	2.98	2.93
	s.e.	0.74	2.69	2.89	2.74	2.71
	N	1275	1275	1275	744	744
	Sample	50km	50km	50km	1840	1840
	Nearest tributary F.E.	No	Yes	Yes	Yes	Yes
	Geographical controls	No	Yes	Yes	Yes	Yes
	State F.E.	No	Yes	No	No	No
	Controls for 1840 pop.density	No	No	No	Yes	Yes

Note: Coefficients from regressions of outcome variable listed on log distance from a bridge, instrumented by an indicator for being upstream of the nearest tributary confluence where indicated. Main sample consists of year 2000 census tracts where any part of the tract is within 10km of the Mississippi or Ohio rivers and within 50km of a tributary. Standard errors are clustered by nearest tributary and upstream status ($N_c = 56$ for sample within 50km of tributaries.).
*** p<0.01, ** p<0.05, * p<0.1.

2.9 Appendices to Chapter 2

Appendix A: Data

Bridge Data

Locational information In the version compiled by the Federal Highway Administration, the NBI normally contains locational information. However, the information is often missing or unverified. The version of the data used was conflated to a common reference network [27] by the Research and Innovative Technology Administration's Bureau of Transportation Statistics (RITA/BTS) to create a shapefile with locational information where available. However, 25% of locational information was missing. Information on bridges carrying railways was also missing from the dataset by design (this information is stored in a separate database).

Extract of original sample and initial checking Bridges over the Mississippi or Ohio Rivers were extracted either through string matching (using a text field describing the feature 'under' the bridge) or through proximity mapping to river shapefiles using ArcGIS. The dataset of possible matches was then hand checked using both satellite imagery (accessed through Google Earth) and other contemporary lists of bridges, to ensure completeness with respect to extant bridges. Since the NBI is compiled from listings created separately by state and county administrative bodies, bridges that spanned two counties or states were often listed twice. The process of hand checking also enabled me to remove duplicate listings of the same bridge.

Supplementary sources of bridge data The data was also cross-checked with alternative sources of information on bridges, typically compiled by amateur bridge enthusiasts. These alternative sources of information included: (Costello, 1995); (Costello, 2002); and various websites including: http://bridgehunter.com/; http://www.johnweeks.com/river_mississippi/; http://www.bridgemeister.com/; http://bridgestunnels.com/; and Wikipedia. In particular, the alternative sources of information were used to establish the original date of first construction at the site, since there is a high occurrence of rebuilding of bridge structures, and the date listed in the NBI often corresponded to the most recent rebuild. Where extant bridges were missing from the dataset, I obtained the required information about the bridge from these alternative sources. In a few cases where alternative sources had significantly conflicting or unclear information about location or number of bridges, I contacted local historical societies to ask them for information from their archives.

Length of maximum span and total length of structure The data on length of maximum span and total length of structure was either cross-checked with or obtained from the alternative sources listed above. Where data on a structure that still exists was missing from all other sources, I obtained the measurements from satellite images using Google Earth. Where data on a historical structure was missing, I obtained the data wherever possible from Office of the Chief of Engineers, United States Army (1948), which lists the width of the navigable channel. This is usually a slight underestimate of the maximum span, but represented the best available estimate.

Census Data

Mapping bridges to counties The bridge data was spatially mapped to the county data, using at most a 5000m radius around the coordinates attributed to the bridge. (The bridges are assigned

[27] TeleAtlas's DynaMap for Transportation

a single point location in the data, but in reality the length of the bridge may be significant). In almost all cases, a bridge treats two counties on either side of the river by connecting them, although there are a few cases, mostly in Louisiana, when both sides of the river are within the same county boundary. Where a bridge is located close to the boundary between two counties, the spatial mapping resulted in more than one or two matches between bridge and county. I cross-checked the original spatial mapping with a spatial mapping with a smaller radius to remove some spurious matches, and hand-checked the remaining matches to ensure that the bridges only match one county on either side of the river.

Population foreign-born or native-born Aggregate statistics on the proportion of the population in a county that is foreign-born are available from the ICPSR dataset for 1870-1900, 1940, 1960 and 1990. I estimate the fraction of the county population that is foreign-born using IPUMS data for all study counties for 1860 and for 1910-1930, and for the subset of counties that coincide with a PUMA for 1950, 1970, 1980 and 2000. I am therefore able to recover the total foreign born and native population, and fraction foreign born, for all counties in all years except 1950, 1970, 1980 and 2000, for which the information is only available for the largest counties.

Population by race and gender Data on the composition of the population by race (black, white and other) and gender is available for all years from the ICPSR dataset, although the definitions used of race are not stable over time.

River Data

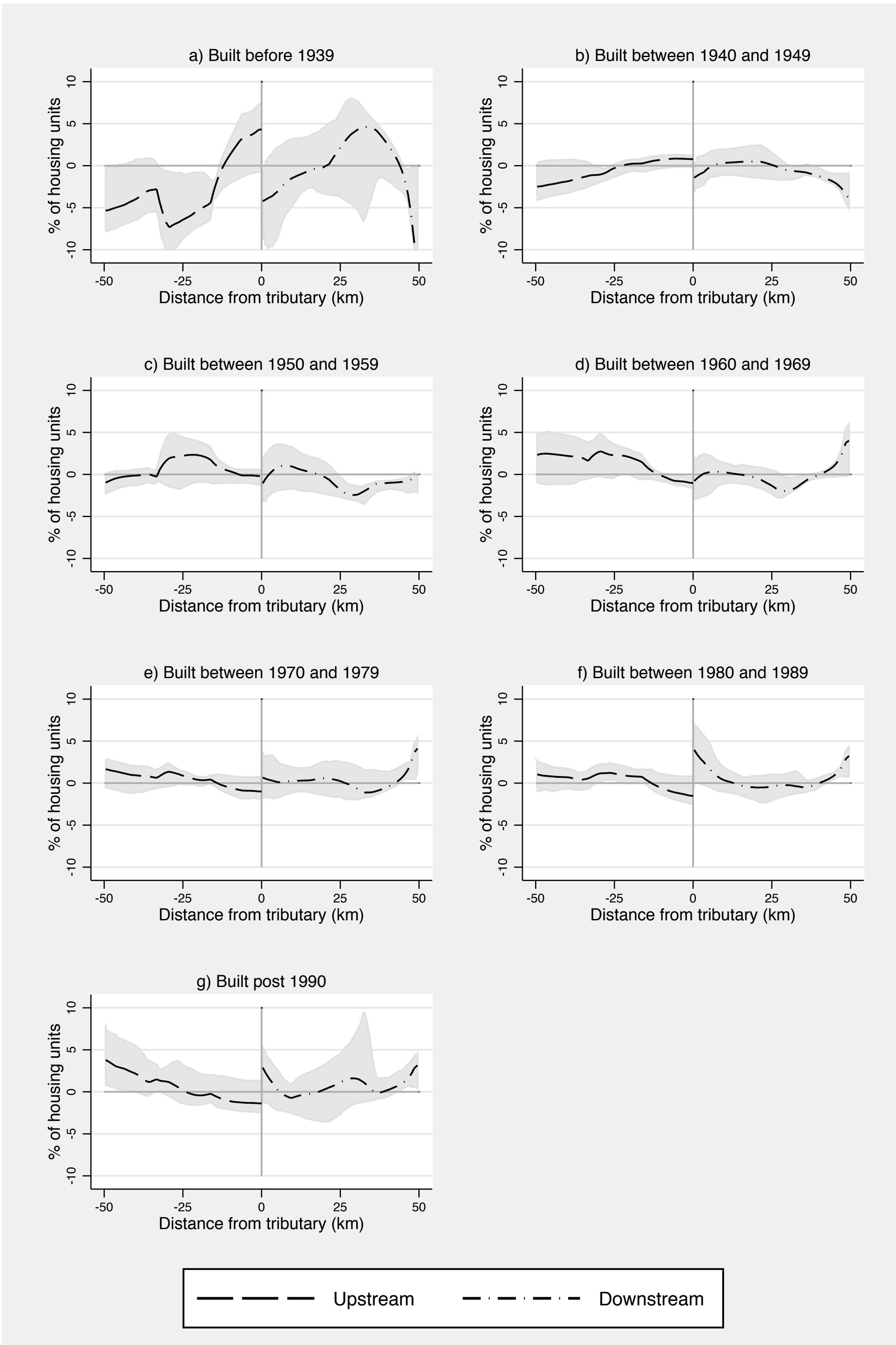
Datasets The three datasets used to map river data to county data are: 1) Flowlines from the USGS National Hydrography Dataset (NHD); 2) Environmental Systems Research Institute, Inc (ESRI) US Rivers (Generalized); and 3) Environmental Systems Research Institute, Inc (ESRI) US Waterbodies.

Mapping rivers to counties I code the county as being on one of the rivers if the county intersects a 200m buffer zone around any of the three river shapefiles, since the shapefiles have slightly different alignments which may reflect changes in river alignment over time. At the same time, I record the minimum distance between the county centroid and the river shapefile to obtain the distance from the county to the river. The river dataset which contains most information about river characteristics is the NHD. To obtain more details about the local characteristics of the river, I assign each reach in the dataset (reaches range between 0.01 and 40 km in length) to the county on each side of the river with which it shares the greatest overlap by creating a buffer zone around the river. I then use the characteristics of these reaches to calculate local river characteristics e.g. the mean flow across these reaches, the total length of the river that intersects the county.

Instrument construction The river is modelled in the dataset in reaches of a mean length of 1.9 km. I first removed river reaches where the data was missing or otherwise clearly misassigned e.g. much larger or smaller than the neighbouring reaches. Each reach is assigned a flow rate. A second variable describes the incremental increase in the flow rate along the reach. The difference between the flow rate in a downstream segment and the flow rate in the upstream segment plus the incremental addition, is the increase in flow at the junction between the reaches. I define a major tributary as a point at which the increase in flow is greater than 7.5%, or greater than 10,000 cubic feet per second. For each reach, I then construct a variable that measures the distance in river km from the nearest tributary. I then assign bridges to the nearest river segment; and calculate the distance between the river segment centroid and the nearest bridge. I then assign census tracts to the nearest river segment.

Appendix B: Figures and Tables

Figure 2.13: Age of housing stock around tributaries (2000)



Graphs show results from a local linear regression of the outcome variable on distance from a tributary, for census tracts less than 50km from a tributary confluence. Outcome variables were previously demeaned around the nearest tributary. Upstream tracts and downstream tracts are pooled separately. Bootstrapped 90% confidence intervals are shown in grey, clustered at the nearest tributary.

Chapter 3

The long-run economic impacts of transport infrastructure: Urbanization, sorting and spatial equilibrium

Abstract

The long-run impacts of transport infrastructure are determined in two stages: i) the direct impact of transport infrastructure on production, and ii) indirect, spatial equilibrium effects, including the relocation of firms, household and productive capital. Previous literature has been unable to separate out the two effects. In this paper, I exploit quasi-experimental variation in distance to a land transport route created by the opening and location of bridges over major rivers in the historical United States to measure the economic impacts of proximity to transport infrastructure both in the following decades, and in the very long run. I show that the direct, short-run impact of greater proximity to a transport route on per-capita production is positive. However, in the long-run, the sign of the effect is reversed: per-capita incomes are lower closer to transport routes in equilibrium. This implies that the spatial equilibrium effects are larger in magnitude, and opposite in sign, to the direct effects. I show that the likely mechanism is urbanization, followed by sorting of poorer households into more dense neighbourhoods, in response to locally reduced transport costs, and of wealthier households into less dense neighbourhoods, where housing units are larger and there are more single-unit households. These results imply that understanding the spatial equilibrium response to changes in transport infrastructure networks is critical to drawing the correct conclusions about the economic impact of transport infrastructure. Ignoring the spatial equilibrium effects would yield grossly misleading assessments of the economic impact of transport infrastructure.

3.1 Introduction

Global spending on transport infrastructure accounts for between 1 and 2% of global GDP (International Energy Agency, 2013), reflecting its perceived importance to policy-makers in promoting and sustaining growth and development. However, there is very little robust evidence regarding the economic impacts of improvements to transport infrastructure. The primary obstacle to establishing a reliable evidence base was to distinguish the impacts of transport infrastructure construction from other factors that affected decisions about infrastructure location. Recent papers have made some progress in addressing this problem. However, even if researchers can empirically resolve the selection problem, the resultant impact measured — particularly over long time horizons — consists of both the direct impact of transport infrastructure, and indirect spatial equilibrium effects, including the relocation of and sorting among workers and households. The importance of these spatial equilibrium effects remains an open question, as previous studies have not been able to separately measure the direct and spatial equilibrium effects.

In this paper, I exploit local variation in distance to a land transport route created by the timing and location of bridge construction over major rivers to compare the economic impacts of changes in the land transport infrastructure network immediately after the change, and in the very long run. Separating out the impacts over time in this way allows me to compare the direct impacts and the impacts in spatial equilibrium, because the spatial equilibrium effects take several decades to play out in full. In the long run, per-capita incomes are 9% lower in places that are 50% closer to a transport route, while population density is 25% higher. A naive interpretation of this result might conclude that transport infrastructure lowers incomes. However, this conclusion would be misguided; in contrast, per-capita production rises immediately when a place experiences a reduction in distance to a land transport route; but this rise in per-capita production attracts an influx of population and leads to urbanization and structural transformation. In the long run, the difference in per-capita income between better-connected and worse-connected areas is likely explained by sorting of poorer households into more dense areas, and richer households into less dense areas, in response to transport costs and the housing market. These spatial equilibrium effects eventually outweigh the direct impacts of transport infrastructure, resulting in lower per-capita incomes closer to transport routes.

Previous studies (Banerjee, Duflo, & Qian, 2012; Faber, 2013) have considered the role of capital mobility in determining the equilibrium economic impacts of transport infrastructure in contexts where labour mobility is highly restricted, but this study is the first to address the degree to which factor mobility influences estimates of the economic impacts of transport infrastructure in a context of high labour mobility[1] This is study is the first to separate empirically the direct (short-run) and indirect (long-run) economic impacts of transport infrastructure, enabling me to evaluate the importance of spatial equilibrium effects. The study also joins the relatively small number of estimates of the economic impact of transport infrastructure which take seriously the endogeneity of location choice in transport infrastructure[2].

I exploit two identification strategies, previously described in Tompsett (2014), henceforth referred to as *Bridges*. To measure the long-run impacts of access to transport infrastructure, I exploit variation in the location of bridges around tributary confluences, which are places where a smaller river connects into a larger river. The associated discontinuous increase in flow rate in the river produces a similar jump in the cost and feasibility of bridge construction, resulting in more bridges being built just upstream of the confluence than just downstream of the confluence. To measure the short-run impacts of changes in the transport infrastructure network, I exploit variation in the timing of bridge construction that is driven by innovations in bridge construction and the time taken to plan, finance, design and build a major bridge, which is on the order of several decades. As in *Bridges*, I use a comprehensive dataset containing details on all the bridges ever built over the Mississippi and Ohio rivers, and 140 years of panel data from the United States Historical Censuses. I control for county-level unobservable characteristics that influence both eco-

[1]The closest study in spirit is Chandra and Thompson (2000) who study the relocation of economic activity by comparing growth trajectories in counties that are connected to the Interstate Highway and their immediate neighbours with a control group of counties located further away. However, their analysis does not consider the movement of population, and also treats highway location as exogenous in non-metropolitan counties. My empirical strategy advances this analysis by addressing selection in transport route location and spillover effects.

[2]These include: the only experimental study in this literature, which focuses on road paving in a Mexican city (Gonzalez-Navarro & Quintana-Domeque, 2012); Donaldson's (2012) study of the impact of the railroads in Colonial India; Banerjee, Duflo, and Qian's (2012) study of the highways in contemporary China.

nomic growth and decisions about bridge construction using county-level quadratic trends. The identifying assumption is that the timing of bridge opening is uncorrelated with other short-term deviations from these local long-run trends.

The results from the two analyses can be interpreted as yielding the direct impacts in the short run — before the movement of population begins in response — and the equilibrium impacts in the long run, with some caveats regarding other differences in what is estimated. In particular, the long-run instrumental variables approach estimates the local average treatment effect at low distances from a transport route, for locations that already have higher probabilities of city formation, as a result of their proximity to the tributary confluence, which acts as a hub on the water transport network[3]. The short-run approach estimates the treatment effect at a broader range of distances.

In the absence of direct measurements of production or income in the panel, I study changes in the value of agricultural land, following precedents in the literature including Donaldson and Hornbeck (2013). Agricultural land in a county that experiences a 50% reduction in distance to a transport route rises in value by an additional 5.5% over thirty years. This rise can be interpreted with additional assumptions as a rise in local production. However, the rise in production precipitates a rise in population density — described in *Bridges* — and is larger in magnitude, leading to an apparent positive difference in production per capita. This result is in sharp contrast to the long-run impact on per-capita income: improvements in access to transport infrastructure increase per-capita production initially, but in the long run, per-capita incomes are lower. This implies that the direct impacts on economic growth are positive, but that the indirect, spatial equilibrium impacts produce differences in income between better and worse-connected areas that are of the opposite sign, and larger in magnitude, than the direct effects.

Multiple mechanisms might contribute to these spatial equilibrium effects. First, I show that thirty years after a 50% reduction in distance to a bridge, the percentage of the population in urban areas is 0.9 percentage points higher, and the percentage of the workforce engaged in agriculture is 1.5 percentage points lower, with corresponding increases in the percentage of the population engaged in manufacturing, construction, retail and wholesale and services. These results suggest that

[3] See Fujita et al. (2001) for a discussion of the role of transport hubs in city formation.

improvements in transport infrastructure increase the local likelihood of structural transformation and urbanization.[4]

In the long run, the local average treatment effect I estimate compares places that all essentially urban; within this subsample, census tracts that are 50% closer to bridges have slightly more urban populations, but the differences are insignificant. However, the areas closer to transport routes are denser, and have lower per-capita incomes. Previous literature has proposed two mechanisms by which poorer people sort into denser areas. First, richer households can afford to live in larger, single unit houses and therefore live where land is cheaper. Second, poorer households are more sensitive to transport costs and live in higher density areas where there is more provision of public transport and where journey times to work are shorter[5].

I find evidence that indicates that both of these mechanisms may be important. First, housing units are larger, and more likely to be single units (as opposed to apartment buildings) at greater distances from bridges. Housing rents and values are correspondingly higher at greater distance from a bridge, but this difference disappears and may even reverse after conditioning on observable housing unit characteristics. Second, people living closer to transport routes have shorter journeys to work, are more likely to commute by public transport and less likely to commute by car, and are less likely to own a vehicle (although the differences are small and insignificant after conditioning on per-capita income.) As such, these results also contribute to a related literature which studies the role of transport infrastructure in determining the distribution of population and income within cities (e.g. Baum-Snow, 2007).

These results imply that estimates of the economic impact of transport are highly contingent on the timing of the measured impacts, and on local conditions that affect the degree to which spatial equilibrium effects are likely to matter, such as labour mobility. For example, these results contrast

[4]These results complement those of Michaels (2008) and Duranton, Morrow, and Turner (2014), who showed that counties and cities which experienced improvements in access as a result of the Interstate Highway system experienced increases in trade-related activities, and trade in heavy goods, respectively; those of Duranton and Turner (2012) who showed that cities in the United States in which the stock of highways increased during the expansion of the interstate highway system grew faster in the following years; and those of Jedwab and Moradi (2013) and Berger and Enflo (2013) who showed that railroad access in colonial Africa and Sweden, respectively, affected local patterns of urbanization, with persistent impacts on the economic landscape.

[5]See Glaeser, Kahn, and Rappaport (2008) for a fuller discussion.

with those of Banerjee, Duflo, and Qian (2012), who find a long-run weakly positive impact of improved access to transport routes on incomes, in a context with highly restricted labour mobility. The results further underscore the sensitivity of estimates of the economic impact of transport infrastructure — particularly those based on structural models — to assumptions, and violations of assumptions, regarding factor mobility, particularly that of labour and firms. The results are also relevant to policy-makers planning improvements in transport infrastructure, who should anticipate influxes of poorer households to the area in their plans for complementary infrastructure. While the focus is on the historical United States, the result is relevant to other parts of the world today, particularly sub-Saharan Africa, where access to transport infrastructure remains low, but there is substantial internal labour mobility, and where rapid, unplanned urbanization presents social, environmental and public health challenges.

The paper is structured as follows. In Section 3.2, I describe the empirical strategies to identify the impacts of transport infrastructure in both the short and long run and in Section 3.3, I describe the data. In section 3.4 I describe the results. Section 3.5 concludes.

3.2 Empirical Strategies

Throughout this paper, I focus on local variation in access to land transport routes created by the construction of bridges over major rivers. In the following section, I summarize the characteristics of bridges that are critical to explain the identification strategies, the identifying assumptions, and the empirical methodologies that I use to measure the impacts of changes in the transport infrastructure network, in both the short and long run analyses. In all cases, I refer the reader to Tompsett (2014) for a fuller account.

The right hand side variable I focus on throughout is distance to the nearest bridge. In all cases, in the absence of comprehensive historical data on the rest of the transport network, I take distance to a bridge as a proxy for distance to a land transport route. The focus on distance to a bridge is motivated by spillover effects; a bridge built in one county may reduce the distance to a land transport route for its neighbouring counties as well as itself. In *Bridges*, I showed that simple pre-post comparison of counties acquiring bridges to those not acquiring bridges ignores these effects, and biases downwards the resulting comparisons; and that even counties that never construct their own bridges experience population growth in response to changes in their distance to other counties' bridges.

3.2.1 Short-run empirical strategy

My short-run empirical strategy — where in this case, the 'short run' refers to the decades following a change in distance to a bridge — relies on several important properties of bridges, the historical evidence to support which was presented in detail in *Bridges*. First, the timing of bridge construction in a given location is strongly influenced by the available technology. Second, it typically takes several decades to plan, finance, design and build a major bridge. Third, once the bridge is opened, the change in local transport routes is realized over a very short time period: often a single day. These properties are critical for my short-run identification strategy, for which the identifying assumption requires that the timing of bridge opening be uncorrelated with short-term deviations from local long-run trends in the outcome variables of interest.

In the panel, I calculate the measure of distance from a bridge taking into consideration bridge closures. This reduces measurement error and helps to accurately capture cases where a bridge is replaced in a nearby, but not identical location. Although there are relatively few cases where a bridge is not replaced nearby, the identifying assumption needs further that the timing of bridge closure be uncorrelated with short-term deviations from local long-run trends in the outcome variables of interest. Since closures are often driven by safety concerns or extreme weather conditions, this assumption also seems plausible.

In estimating the short-run impacts, the estimating equation is as follows:

$$y_{i,t} = \gamma_t + \alpha_{0i} + \alpha_{1i}t + \alpha_{2i}t^2 + \sum_{j=0}^{k} \beta_j \Delta dist_{i,t-j} + \epsilon_{i,t} \tag{3.1}$$

where $y_{i,t}$ is the outcome variable i at a time t, γ_t is a year fixed effect that flexibly captures global trends in the outcome variable and distance to a bridge, α_{0i}, α_{1i} and α_{2i} are county-specific parameters that approximate the local long-term counterfactual trend in the outcome variable, $\Delta dist_{i,t-j}$ is the change in log distance to a bridge j time periods ago, and β_j is the coefficient of interest, the cumulative effect on the outcome variable at time t of a change in distance to a bridge j periods ago.

Fitting a long-term quadratic trend line absorbs persistent effects, so I only estimate the first few β_j terms accurately, around the sharp change experienced in distance to a bridge. If long-term effects exist, they will tend to bias my estimates towards zero, as long as they are of the same sign as short-term effects. I include 5 lags of the change in distance, but only consider the estimates of the coefficients on the first 3 lags to be reliably estimated, given serial autocorrelation in the $\Delta dist_{i,t-j}$ terms[6]. In formal terms, the identifying assumption therefore is that:

$$\begin{aligned} E\left(\epsilon_{i,t}|\Delta dist_{i,t+k}, \alpha_{0i}, \alpha_{1i}, \alpha_{2i}, \Gamma\right) = 0 \\ t = 1860, 1900, ..., 2000 \\ k = -30, -20, ..., 20, 30 \end{aligned} \tag{3.2}$$

where Γ is the vector of year fixed effects. In other words, that the timing of a change in distance to a bridge is exogenous to deviations from the long-term trend within a window of 30 years either side of the date at which construction takes place. This is locally equivalent to assuming that $E(y_{i,t}|\alpha_{0i}, \alpha_{1i}, \alpha_{2i}, \Gamma) = \gamma_t + \alpha_{0i} + \alpha_{1i}t + \alpha_{2i}t^2$ for $t = 1890, 1900, ..., 2000$. Most of the outcomes I consider are scale-invariant (e.g. the percentage of the workforce in a given industry), but where there is a scale consideration, log transformations and county-level fixed effects remove the scale effects (e.g. with county fixed effects, it is exactly equivalent to consider log population and log population density).

In *Bridges*, I showed an extensive series of robustness checks in support of this identifying assumption, where population density was the outcome of interest. In particular, I showed that

[6] In *Bridges*, I previously showed that with log population as an outcome variable the estimated coefficients on the first 3 lags were not sensitive to inclusion of additional lags, and including additional lags only reduced precision.

there was no evidence for population growth in the decades preceding a change in distance to a bridge; that the results held for counties in which bridges were never constructed, ruling out a potential alternative explanation in which bridges are built in anticipation of county-level growth; that the results could not be driven by changes in the relationship between growth rates and population density over time; that the results were consistent across all three river basins, and across a range of alternative time specifications; and that the results were consistent to varying the way in which the long-run trends were modelled. I also showed that the results remained consistent in sign and timing, though were attenuated in magnitude and significance, when I included sets of progressively more conservative disaggregated time trends. In general, I do not repeat this full set of robustness checks for each of the outcome variables considered, largely taking as given the validity of the identification strategy based on the analysis in *Bridges*, and indicate in the text where I am unable to replicate these tests for certain outcome variables, typically where there are highly skewed distributions or systematic missing values.

To make the correct inference about whether relationships between outcomes and changes in distance to a bridge are statistically significant, it is important to correct for serial correlation in population growth rates, and spatial correlation across counties (Bertrand et al., 2004; Angrist & Pischke, 2009). As in *Bridges*, I cluster standard errors at the county level, to allow for arbitrary levels of serial correlation, and add Conley standard errors, allowing for spatial correlation over a distance of 200km.

In *Bridges*, I discussed in detail how ferry crossings might influence the estimates, as these were the principle alternative strategy for crossing rivers prior to bridge construction, but are unobserved in my data. I concluded that ferry crossing sites might well be orthogonal to bridge crossings (due to the different geophysical requirements), but would otherwise bias the estimates downward (by increasing my counterfactual estimate of growth in the absence of a change in distance to a bridge, or by relocating to compensate areas further from a bridge in response to bridge construction).

3.2.2 Long-run empirical strategy

The previous empirical strategy only allows me to measure impacts in the first few decades following changes in distance to a bridge, for two reasons. The first is that the validity of the identifying assumption only extends to a window of decades around the actual timing of the change; the second is that the long-run quadratic trends fitted absorb persistent effects.

In order to measure long-run impacts, I use an instrumental variables approach to separate out variation in the location of bridges that is otherwise uncorrelated with local population densities or growth rates. To do this, I focus on discontinuities in the flow rate in the river where a tributary — or smaller river — joins the main stream. The higher the flow rate in the river, the more difficult and costly is bridge construction, because either some or all of the depth, width, and velocity of the water in the river must increase to accommodate the higher flow. At a tributary confluence, the flow increases sharply, resulting in a local asymmetry in the probability of bridge construction, and a corresponding sharp change in distance from a bridge. Figure 3.1[7], panels a) to c), illustrate this discontinuity.

The confluence of a tributary and a river is in itself a place where people choose to live, as it forms a hub on the water transport network, as the place where two natural water transport routes — the river and its tributary — coincide (see Fujita et al., 2001). Panel d) of Figure 3.1 illustrates this; population density increases from both directions with reducing distance from the tributary, but the upstream-downstream asymmetry remains. I therefore control for distance to the nearest tributary, and use an indicator for whether the section of the river is upstream or downstream of the tributary as an instrument for distance to a bridge. Since the flow rate generally increases upstream to downstream — with corresponding gradients in bridge construction and distance — I also include distance from the river mouth, interacted with river dummies, to avoid confounding the upstream-downstream comparison with overall north-south trends.

Using a cross-sectional instrument to estimate a dynamic process is problematic in the presence of dynamic effects. Present-day population in a given location is influenced by the full history of

[7] Replicated from Tompsett (2014).

access to transport infrastructure in the location, and not just by current access. However, the instrument is fundamentally cross-sectional in nature. As in *Bridges*, I treat distance from a bridge in the year 2000 as a proxy for the cumulative history of distance to a bridge at a given location over time.

Since local discontinuities around the tributary confluence are attenuated at the county level, I focus on census tracts, the next level of aggregation available, selecting census tracts for which any part of the tract is within 10km of the river; I also include tracts which are completely enclosed by tracts meeting the preceding criteria. Formally, I use the following as the estimating equation in this section:

$$y_{i,2000} = \alpha + \Omega X_i + \pi dist_{i,2000} + \epsilon_{i,2000} \tag{3.3}$$

where $y_{i,2000}$ is the outcome variable in census tract i in the year 2000, and $dist_{i,2000}$ is the distance to a bridge in the year 2000. π is the coefficient of interest. The term $\pi dist_{i,2000}$ approximates the cumulative effects on the outcome variable in the year 2000 of distance to a bridge at all previous times j in the area's history or $\sum_{j=0}^{\infty} \pi_j dist_{i,2000-j}$. X_i is a vector of control variables, including : nearest-tributary fixed effects, a quadratic function of distance from the nearest tributary, and distance from the mouth of the river, interacted with an indicator for each river. Ω is the vector of coefficients on these control variables.

Since decisions about infrastructure location are driven by knowledge about where economic activity takes place and where it is anticipated to grow, the coefficient π is estimated with bias in the cross-section. The location of a tributary creates a natural experiment whereby bridges are more likely to be built upstream than downstream, but there is no reason to think that location just upstream or just downstream of a tributary should otherwise influence economic activity, conditional on nearest-tributary fixed effects, distance from a tributary and an overall north-south gradient. I therefore use an indicator which takes the value 1 upstream of the nearest tributary and 0 downstream of the nearest tributary to instrument for distance to bridge, including interactions between this indicator and a quadratic function of distance from the nearest tributary as additional exogenous controls. Full details on the instrument construction are included in *Bridges*. I therefore instrument for distance to bridge using the following first stage equation:

$$dist_{i,2000} = \alpha + \Omega X_i + \gamma \mathbb{1}\left(upstream = 1\right) + \nu_{i,2000} \tag{3.4}$$

Since I expect the instrument only to be valid in the vicinity of the tributary, I limit the sample to census tracts matched to river segments within 50km of a tributary[8]. To make the correct inference about the significance of observed relationships in the context of high levels of spatial correlation, I cluster standard errors throughout the analysis at the level of the upstream-nearest tributary interaction, resulting in 56 clusters in the main sample (census tracts less than 10km from the river and less than 50km from a tributary).

For the upstream indicator to be a valid instrument for distance to a bridge, it must first be a sufficiently strong predictor of distance to bridge. In column 1) of Table 3.1, I show that the relationship in the first stage is consistent with intuition: the distance to a bridge is shorter upstream of a confluence. The F-statistic on the upstream instrument is 17.6 or 15.7 (depending on the other controls included) in the main sample (less than 50km from a tributary); and is 8.3 in a restricted sample, for which I can observe population density at the county level in 1840. The rule of thumb for a sufficiently strong first stage is an F-statistic of 10 Angrist and Pischke (2009).

The instrument must also satisfy the exclusion restriction: being just upstream of a tributary, rather than just downstream, must not have any other effect on the outcome variables, except to alter the likelihood of bridge construction. In formal terms:

$$E\left(\epsilon_{i,2000} | X_i, \mathbb{1}\left(upstream = 1\right)\right) = 0 \tag{3.5}$$

[8] I showed previously in *Bridges* that the full sample did not change the nature of the relationship of interest around tributaries, but a more complex specification is required to isolate the local effects around the tributary.

In support of this assertion, I showed in *Bridges* that asymmetries in population density upstream and downstream of tributary confluences only emerged after the era of bridge construction, and that pre-bridge construction differences were if anything opposite in sign. Finally, the instruments must also satisfy the monotonicity assumption, which in this case seems reasonable; it is difficult to construct an alternative scenario where a local increase in flow rate would make construction of a bridge more likely.

3.3 Data

Bridges

The bridges dataset contains essentially all the bridges ever constructed over the Mississippi and Ohio rivers, below Pittsburgh on the Ohio and a cut-off below Lake Winnibigoshish in North Central Minnesota[9]. I originally extracted the data from the National Bridge Inventory (NBI), a dataset compiled by the Federal Highway Administration, and cross-referenced the data with historical and contemporary sources and satellite imagery; the dataset is described in full in Bridges. Figure 3.2 shows the geographical distribution of bridges on the Mississippi and Ohio in 1860, and in 2000. Only 4 bridges were constructed prior to 1860, but the geographical area for which a county-level panel can be constructed reduces significantly using an earlier starting date; I showed in *Bridges* that changes in the time frame did not significantly alter the estimated coefficients of interest.

Most of the bridges in the dataset are road bridges; around a quarter are either rail bridges, or of mixed use i.e. have a rail crossing and a road crossing. Before 1900, the majority of bridges constructed were rail bridges. There are later peaks in bridge construction activity during Roosevelt's New Deal programs at the end of the Great Depression, and during the construction of the Interstate Highway System.

[9] I showed in Tompsett (2014) that the results were not sensitive to changes in this cut-off point.

Census Data

Outcome data is drawn from historical censuses from the United States. Although census data has been collected in the United States since 1790, the area of coverage, and the questions asked, have varied with time. The unit of analysis in the short-run analysis is the county, since this is the finest level of spatial detail available over the full historical period of interest, and in the long-run analysis is the census tract. As in *Bridges*, I focus in the short-run analysis on counties which are completely covered by the bridge data.

The study uses three sources of census data. I use aggregated data at the county level from two sources: the United States Censuses from the National Historical Geographical Information System (NHGIS) [10] and the Inter-university Consortium for Political and Social Research (ICPSR) [11]. I also obtained shapefiles for historical county boundaries from the NHGIS, along with census tract data for the year 2000. Where aggregated data is not available from either of these two sources, I estimate county-level aggregate variables using individual-level data from the Integrated Public Use Microdata Series (IPUMS)[12]. I deal with changes in county boundaries over this time period by remapping all data back to 1860 county boundaries, as it *Bridges*.

Rivers

As in *Bridges*, I used three different spatial datasets to map counties and census tracts to rivers. I use data on river flow from the National Hydrology Dataset to construct the instrument for the long-run analysis.

---

[10] Minnesota Population Center. National Historical Geographic Information System: Version 2.0. Minneapolis, MN: University of Minnesota 2011 http://www.nhgis.org

[11] Haines, Michael R., and Inter-university Consortium for Political and Social Research. Historical, Demographic, Economic, and Social Data: The United States, 1790-2002. ICPSR02896-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-05-21. doi:10.3886/ICPSR02896.v3

[12] Steven Ruggles, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota, 2010.

Sample

Figure 3.3 shows a map of the county sample used in the short-run analysis. There are 181 counties in the sample, from 14 states. I measure distance to a bridge using the distance between a county's centroid and the nearest bridge at any moment in time. The sample used in the long-run analysis is a further refinement of this sample, selecting census tracts for which any part of the census tract is within 10km of the river, or census tracts that are completed enclosed by census tracts selected on the basis of preceding criteria. I then assign properties from the nearest river segment to each census tract.

3.4 Results

In this section, I describe the results. I first describe the apparent overall dichotomy; per-capita production appears to rise initially in places experiencing a 50% reduction in distance to a bridge, but per-capita incomes are 9% lower in the long run. This suggests that while the direct effect of transport infrastructure on economic growth is likely to be positive, spatial equilibrium effects result in differences between better and worse-connected areas that opposite in sign to the direct effects. I then provide evidence for a potential mechanism for these spatial equilibrium effects. I first show that there are increased rates of urbanization and structural transformation in counties receiving improvements in access to transport. I then provide evidence for sorting within urban areas, with at least two potential channels: lower transport costs in more dense, more central areas, and higher value housing constructed in less dense, peripheral areas.

3.4.1 Production and income

In this section, I examine the impacts of improved transport infrastructure on economic development in the short-run and the long-run. In general, I am unable to observe the equivalent of GDP at the appropriate level of disaggregation in either the short or the long run. In the short run, I show instead that the value of agricultural land rises by approximately 5.5% in the decades following a 50% reduction in distance to a bridge. Under additional assumptions, this can be interpreted as a local rise in the value of production. In the long run, I show that in contrast, local per-capita incomes are 9% lower in census tracts that are 50% closer to a bridge.

Short run: Value of agricultural land

Whereas mobile factors of production — including labour — shift in location in response to a local increase in their returns, increased productivity is capitalized as an increase in the rental rate of immobile factors of production, such as land. Structural approaches have therefore previously used the value of agricultural land as a proxy for production (e.g. Donaldson and Hornbeck,2013).

The argument proceeds as follows. Land rental for a unit of land is equal to αY, where α is the factor share of land and Y is total production per unit land area, and land rental is capitalized in land value V_z as $\frac{1}{r}\alpha Y$, where r is the interest rate. If α and r do not vary with distance from a bridge d then $\frac{\partial \ln V_z}{\partial \ln d} = \frac{\partial \ln Y}{\partial \ln d}$, and the elasticity of land values with respect to distance from a bridge is equivalent to the elasticity of production with respect to distance from a bridge. The required assumptions are that the factor share of land and the returns to capital be constant across space; I discuss the impact of relaxing these assumptions at the end of this section.

Continuing this line of reasoning, and conditional on the same assumptions, the elasticity of per capita production with respect to distance from a bridge can therefore be estimated using the elasticity of the difference between the log value of agricultural land and log population to distance from a bridge. Since per capita production y is equal to $\frac{Y}{L}$, where Y is production per unit land area and L is population per unit land area, then:

$$\frac{\partial \ln y}{\partial \ln d} = \frac{\partial \ln Y - \ln L}{\partial \ln d} = \frac{\partial \ln Y}{\partial \ln d} - \frac{\partial \ln L}{\partial \ln d} = \frac{\partial \ln V_z}{\partial \ln d} - \frac{\partial \ln L}{\partial \ln d} = \frac{\partial \ln V_z - \ln L}{\partial \ln d}$$

In Figure 3.4, I show how land values evolve in the decades following a change in distance to a bridge, showing the effect on population and the difference between the two for comparison. I focus on log average land values per ace. Taking the log reduces the asymmetry of the data, and normalizes the effect size in percentage terms over time, as the data is in nominal values. Overall changes in the value of a dollar are captured by the year fixed effects. Figure 3.4, panel b) illustrates that coefficients on the lead values of the variables of interest are not statistically different from zero, and are much smaller in magnitude than the coefficients on the lag variables; I exclude them from the main specification, which improves precision without altering the estimates.

Table 3.3 shows these results. Column 1) replicates the results from *Bridges*. Column 2) shows the results for agricultural land values: the strongest effect measured is equivalent to a 5.5% increase in the values of agricultural land for a 50% reduction in distance to a bridge. The effect on agricultural land values precipitates the population response, and is larger in magnitude. As a result, a difference between the two opens up, which can be interpreted as a difference in per-capita production. Column 3) shows that a reduction in distance to a bridge leads to a significant increase in per-capita production over the first few decades[13].

Four important points are worth noting. First, the rise in per-capita production, and the lag between an effect on land values and an effect on population, are consistent with a wage differential opening up between better and worse connected areas, which in turn drives migration. Second, the immediate impact on production is smaller than the impact after two or three decades, because the component of the difference in production between better- and worse-connected areas that is attributable to movement in the labour force materializes more slowly than the immediate, direct increase in production. Third, after thirty years, the net effect of production per-capita is not statistically different from zero, implying that essentially all of the observed difference in production between better and worse-connected areas comes from the relocation of economic activity. Finally, total production probably reduces overall in worse-connected areas, relative to a counterfactual scenario, but per-capita production may actually rise. However, aggregate impacts on per capita production will be indistinguishable empirically from overall trends.

If we relax the assumptions on factor shares and returns to capital, then the estimated coefficient might be larger than the true effect on production. For example, if α increases with distance from a bridge (because agricultural activities become less capital intensive at greater distance from a transport route), i.e. $\frac{\partial \ln \alpha}{\partial \ln d} > 0$, and r decreases with distance from a transport route (consistent

[13] These differences are insignificant, however, if I include the lead values, which reduces precision.

with Banerjee, Duflo, and Qian (2012), and assuming that capital flows towards urban areas at a cost) i.e. $\frac{\partial \ln r}{\partial \ln d} < 0$, then $\frac{\partial \ln V_z}{\partial \ln d}$ becomes an upper bound for the impact of transport infrastructure on production. The estimate may be further inflated because the measure of the value of agricultural land incorporates improvements to land including fencing, irrigation and buildings. I am unable to correct this data to recover purely the value of land (as in Donaldson and Hornbeck (2013)) as these corrections are not available over the entire time period of interest. Also, if the area of agricultural land reduces in a county over time, and the best land is selected, the average value of land per acre would tend to also reflect the selection effect. These effects — which tend to lead the measured effect on land values to be larger than the true effect on production — must however be offset against measurement error; a downward bias created by the inclusion of the long-run trends which bias downwards the estimates of the short-run effect if the long-run effect is of the same sign; and the fact that while the value of agricultural land must equal the value of urban land at the geographical margin between urban and rural areas, it may understate the value of land in urban areas.

Since the value of agricultural land reflects the expected value of future land rents, we might expect land rents to rise in anticipation of bridge construction. There is no evidence of increases in land values prior to bridge construction, which may reflect uncertainty about whether bridges will be built or completed over the time-scales in question; data is only available at decadal intervals. However, this does raise some further questions regarding whether or not the value of agricultural land incorporates anticipated population growth. If this is the case, then this might be an alternative factor which drives the effect on the value of agricultural land to overstate the contemporaneous effect on production.

Long run: Incomes

Figure 3.5 illustrates the main effect. Per-capita incomes are lower upstream of tributary confluences, where distances to a bridge are lower; this effect only partially offsets the population density effect, so that total incomes are still higher. In Table 3.3, I show a series of estimates of the impact of distance to a bridge on incomes. In the first row, I show the results from *Bridges* for reference: population density is approximately 25% higher in census tracts that are 50% closer to a bridge. In the second row, I show the results for per-capita incomes: per-capita income is approximately 9% lower in census tracts that are 50% closer to a bridge. Finally in the third row, I show the results for total incomes — calculated by combining the data on per-capita incomes and population density. The preceding two effects partially offset each other, so that total income remains higher closer to a transport route, although the resultant differences are not statistically significant across all samples.

Table 3.3, column 1), shows the reduced form relationships between the outcome variables and log distance to a bridge: population and total income are strongly decreasing with distance to a bridge, while per-capita incomes are weakly decreasing with distance to a bridge. In other words, population and overall economic activity are clustered around transport routes, and places with slightly better access to transport routes are slightly richer, if anything, in the cross-section.

In column 2), I present the results when I instrument for distance to a bridge using the upstream indicator. The coefficient for population density and total income is reduced in magnitude, reflecting the expected positive selection effect: at least some infrastructure is constructed to serve locations where population and economic activity is located for other reasons. However, the coefficient on per-capita income flips in sign: once I remove bias due to the fact that richer places can afford to build more infrastructure in general, the relationship is negative: per-capita incomes increase with distance to a transport route.

In columns 3), 4) and 5) I show how the estimates vary with robustness checks. In column 3), I introduce state fixed effects. In column 4), I reduce the sample to the subset of census tracts which I can match to an 1840 county, which excludes the northernmost parts of the Upper Mississippi. In column 5), I show the results controlling for 1840 population density at the county level. Altering the sample (between columns 2) and 4)) tends to increase the magnitude and significance of the results, consistent with evidence shown in Tompsett (2014) that the effect sizes were weakest for the Upper Mississippi, where contemporary bridge density is highest. However, neither introducing state fixed effects nor controlling for pre-bridge era population density alters the results.

The long-run gradient in per-capita income is particularly surprising given the evidence that production per-capita rises in the first decades after bridges are constructed. To explore the mechanisms by which this gradient emerges in spatial equilibrium, I now examine the impact of transport infrastructure on urbanization and structural transformation, and then consider how these processes, combined with local sorting, lead to the reversal of the estimated impact in the long run.

3.4.2 Urbanization and structural transformation

In this section, I present more suggestive evidence to show that urbanization increases following a change in distance to a bridge, and that structural transformation also takes place, with a decrease in the fraction of the workforce in agriculture. A county experiencing a 50% reduction in distance to a bridge sees a 0.9 percentage point increase in urbanization, and a 1.5 percentage point decrease in the relative proportion of the population working in agriculture. In the long run, I use non-parametric evidence to illustrate that the local average treatment effect I estimate applies at distances from a transport route which (in the long run) are already urbanized. As a result, I find no significant impact on urbanization or agriculture, but that increasing distance from a transport route is associated with an increased fraction of the workforce in wholesale and transportation, consistent with typical differences between city centres and the surrounding urbanized area.

3.4.2.1 Short run: Urbanization and industrial composition in the workforce

In Table 3.4, I show the results from estimates of the impact of changes in distance to a bridge on the percentage of population living in urban areas, and the percentage of the workforce in various industries, using Equation 3.1. In column 1), I show that the percentage of the population living in urban areas increases by approximately 0.9 percentage points following a 50% reduction in distance to a bridge, 30 years later. This proportional reduction is driven by relatively high growth in the urban population, and not by absolute shrinkage in the rural population: the rural population also increases, but by a smaller amount[14].

In the absence of county-level data on production, I treat industrial composition in the workforce as a proxy for industrial composition in production. This may overstate the importance of labour-

[14] Results available on request.

intensive industries to the economy. In columns 2) to 7), I show that the percentage of the workforce in agriculture in a county experiencing the same decrease in distance to a bridge declines by 1.5% over the same time period, with offsetting increase in the fraction of the population employed in manufacturing, construction, retail and wholesale, and services. These results suggest that, consist with theory and prior empirical evidence, transport infrastructure increases the rate of urbanization and structural transformation, probably by facilitating the import of cheap substitutes and the export of goods for which the area has a comparative advantage in production.

In general, these results are less robust than the results on land values. In particular when I include lead variables in these regressions, some lead coefficients are significant and for some industries, similar in magnitude to the coefficients on the lagged variables. The coefficients on the lagged variables are more unstable than for other outcome variables. However, the quality of the data is not as good, as there are missing values for some counties and decades, and the variable definitions may not be completely stable over time (for example, industries included as services may vary over time). The distributions of the variables are often highly skewed, and in particular, concentrated around zero. These less robust results do not generally invalidate the identification strategy, as whenever a variable has close to complete observations and a relatively symmetric distribution (e.g. log rural population), I can replicate the robustness checks on lead variables. However, they do suggest the need for caution in interpreting these results, as the presence of spurious lead coefficients suggests instability in the estimates and possible sensitivity to outliers.

3.4.2.2 Long run: Urbanization and industrial composition in the workforce

In Figure 3.6 and Table 3.5, I evaluate the long-run effects of access to transport infrastructure on urbanization and industrial composition.

First, it is helpful to discuss what the local average treatment effect is in the long-run analysis. Figure 3.1, panel d) is helpful here: as previously noted, population increases as we approach a tributary confluence from both sides, reflecting the role that tributary confluences play (or played) as a node on the water transport network. Similarly, as shown in figure 3.6, panel a), urbanization increases from both sides, and agriculture similarly decreases (panel b)). Manufacturing (panel d)) and wholesale (panel e)) peak at medium distances from the tributary confluence, while typically urban industries such as services (panel h)) peak around the confluence. The local average treatment effect measured in the long run analysis evaluates the impact of approximately halving the distance to a bridge, but at short distances from a bridge (typically within a few miles). At this distance to a bridge, in the year 2000, the neighbourhoods considered are typically all relatively urbanized. As a result, the long-run effect measured is the local average treatment effect which essentially compares places very close to the bridge with their immediate neighbours.

Table 3.5, panel a), shows that as might be intuitive from this discussion, the effect on urbanization is of the expected sign (i.e. places closer to transport routes are more highly urbanized) but small and insignificant. In panel b), I repeat the analysis for the percentage of the workforce in seven industries. Essentially no effect is seen for agriculture — in all likelihood a result of the local average treatment effect estimated, as Figure 3.6, panel b) shows that agricultural employment is relatively low on both sides of the tributary confluence. The other effects are characteristic of a comparison between peri-urban and urban areas: services decrease, while manufacturing, wholesale and transportation (which includes communication and utilities) increase. However, only the effects on wholesale and transportation are significant.

The preceding results show that improvements to transport infrastructure most likely play a role in the urbanization process, probably as a result of helping to select between possible multiple local equilibria. In the short run, improvements to transport infrastructure are followed by urbanization and structural transformation from an agricultural economy into an urban, industrialized economy. In the long run, places close to a transport route exhibit more strongly urban industrial profiles, while places slightly further away exhibit more peri-urban characteristics.

Previous studies have noted that the most dense areas of cities attract the poor (see references in Glaeser et al.,2008), proposing as mechanisms the land market (since rich people consume relatively more land and larger houses, they are attracted by the lower land prices in the suburbs) and access to transportation (since more dense areas are better-served with public transport). The second explanation is particularly salient in this context, where the urban development in question is spurred by the initial construction of transport infrastructure. In the following section, I provide evidence that supports an explanation whereby sorting among the urban population follows urbanization, to produce the observed income gradient in the long run.

3.4.3 Sorting: transport expenditures and the housing market

In this section, I examine two possible explanations for sorting on incomes between more dense areas, closer to transport routes, and less dense areas, further from transport routes. I find evidence for both mechanisms proposed by Glaeser et al. (2008). First, I show that transport behaviour is different: households living further from a transport route own more vehicles, have longer commuting times, are more likely to commute by car, and are less likely to commute by public transport. Second, I show that the housing market is different. Housing units are larger, and more likely to be single unit structures (as opposed to multi-housing unit structures such as apartment buildings). Rents and housing values are higher, but these differences are attenuated and may even reverse, once I condition on observable housing characteristics. Together these patterns provide suggestive evidence in favour of both mechanisms.

3.4.3.1 Long run: Transport choices

Figure 3.7 and Table 3.6 show the main results. In Table 3.6, row 1, I show in column 3) that greater distance from a bridge is associated with a larger number of vehicles per capita. I note in column 4) however, that when I condition on per capita income, the difference is attenuated and insignificant. In other words, places with the same approximate incomes only exhibit slightly higher levels of vehicle ownerships at greater distance to a bridge. This result is also shown in Figure 3.7, panel a). However, places that are further from a bridge also exhibit other variations in their transport behaviour: they have slightly (insignficantly) higher commuting times, and are more likely to commute by car than by public transport. A 50% reduction in distance to a bridge is associated with 0.04 vehicles per person, a 2% shorter commute in minutes (mean commuting time across census tracts is 22 minutes), a 3.5 point decrease in the percentage of workers commuting by car (mean percentage of workers commuting by car 88.5%), and a 1.9 point increase in the percentage of workers commuting by public transport (mean percentage of workers commuting by public transport 4.3%.) The relatively small differences in commuting time are possibly a product of two opposing effects i.e. commuting distances may be shorter, but commuting by public transport may be slower.

These results are consistent with sorting of poorer people into more densely populated areas that are better served by public transport, particularly as the more densely populated areas in question are located close to major transport routes. It seems likely therefore that local transport routes connect to these areas. For example, it is rare for local train or bus services not to connect to the central train station. However, I am unable to confirm that the causality runs in this direction. To be clear, I find that in the long run, closer to bridges, i) people are poorer and ii) people use more public transport, own fewer vehicles, and have slightly shorter commuting times. However, I am unable to rule out the possibility that poor people have different transport habits because they are poor, and sort into areas that are closer to transport routes for other reasons. Indeed, evidence from the housing market supports the existence of other, possibly complementary mechanisms.

3.4.3.2 Long run: Housing market

Figure 3.8 and Table 3.7 illustrate the main results. Housing units are larger, and more likely to be single household structures, at greater distance from bridges. Rents and housing values are correspondingly higher. However, conditional on observable housing unit characteristics, the differences in rents and housing values are attenuated, and may even reverse. In other results described in *Bridges*, I showed that the housing stock is also on average older in areas closer to bridges. Important unobservables remain — in particular, the size of plot. Taken together, these results are consistent with sorting of richer and poorer households between areas of higher and lower density as product of increased housing size and quality in less dense areas.

An alternative explanation for sorting on incomes relative to the location of transport infrastructure follows from the question of whether transport infrastructure acts as a consumption amenity or disamenity. In particular, if consumption acts as a disamenity through pollution, congestion or traffic accidents, then rents might fall in the vicinity and poor households might sort into these neighbourhoods to take advantage of lower rents. This mechanism does not appear to be salient, as this would predict that rents and housing values would be lower closer to transport routes, conditional on housing quality. The results show that conditional on observable housing quality, rents and housing values are the same, or higher, closer to bridges.

3.4.4 Education

An alternative mechanism to explain the gradient in per-capita income would focus on sorting in the labour force, depending on whether transport infrastructure complements skilled or unskilled labour. Theory and prior empirical evidence does not give a clear prediction as to whether migration in response to changes in the transport infrastructure network should primarily come from low- or high-skilled labour. On the one hand, Duranton, Morrow, and Turner (2014) show that industries specializing in heavy goods are more likely to relocate in areas with better access to transport infrastructure, which might favour low-skilled labour. On the other hand, if transport improvements lead to urbanization and structural transformation, these changes might favour more highly skilled workers.

To study the changes in skill composition resulting from the growth in population in the short run, I create a composite variable that ranks the average education level in the population within the sample at a given time, and examine whether a county's relative skill level increases or decreases following a change in distance to a bridge. The results, shown in Appendix Table 3.8 are ambiguous: in the first half of the study period, the relative education level of the population falls — reflecting an influx of relatively less well-educated migrants — and in the later half of the study period, the relative education level of the population rises. In the long-run, the evidence is also inconclusive; Appendix Table 3.9 and Appendix Figure 3.9 show that education levels increase with distance from a transport route, but less so than the corresponding income gradient would predict. Taken together, this evidence does not provide strong evidence that sorting amongst types of workers leads to the observed gradients in per-capita incomes.

3.5 Summary and Discussion

This paper has focused on comparing the direct, short-run economic impacts of transport infrastructure to the long-run impacts in spatial equilibrium. I exploit quasi-experimental variation in the timing of bridge openings to estimate the short-run impacts, and in the location of bridges around tributary confluences (places where a lesser river flows into a larger river) to estimate the long-run impacts. These strategies were previously applied in an earlier study which focused on the role of transport infrastructure in determining patterns of economic geography (Tompsett, 2014), herein referred to as *Bridges*.

In the short run, I showed that places that experienced a 50% reduction in distance to a bridge experienced a 5.5% increase in the value of agricultural land, which under additional assumptions can be interpreted as an increase in production. Combined with previous estimates on changes in population density, this implies an increase in per-capita production of 3.2%. In the long run, however, I show that per-capita incomes are 9% lower in places that are 50% closer to a bridge. I interpret the short-run results as showing the direct impacts, and the initial changes that anticipate the spatial equilibrium response; and the long-run results as the impacts in equilibrium. The results suggest that the spatial equilibrium response reverses the gradient of per-capita economic impacts between better- and worse connected areas.

I show that the likely mechanism for this reversal is urbanization, followed by sorting of poorer households into denser, more central urban areas. Counties which experience a reduction in distance to a bridge of 50% also experience a 0.9 percentage point increase in urbanization, and a 1.5 percentage point decrease in the relative proportion of the population working in agriculture, driven by relatively higher growth in the urban, non-agricultural population, rather than by absolute shrinkage in the rural, agricultural population. These results are however less robust to different specifications than the short-run results on land values. In the long run, the local average treatment effect I measure shows concentrations of typically urban activities closer to transport routes, and more characteristically peri-urban activities at slightly greater distances.

I then provide evidence in support of two complementary mechanisms by which poorer households sort into more dense, more central urban areas: lower transport costs in more dense, central areas cause poorer households to move in, while richer households move towards less dense, less central areas in order to consume higher quality housing (larger, single household units). I show that the evidence is on the other hand not consistent with two other potential explanations. There is no evidence for transport infrastructure acting as a disamenity, locally lowering housing prices and thereby attracting poorer households, as conditional on housing quality, housing prices are if anything higher in more dense areas. The evidence is also inconsistent with a mechanism of sorting amongst higher and lower-skilled workers, as transport infrastructure initially attracts relatively less educated migrants, then relatively well educated migrants; and finally households in less dense areas are better educated, but the differences are less than the differences that the income gradient would predict.

Since urban poverty is historically a peculiarly American phenomenon, these exact processes may not be replicated elsewhere. However, the results indicate that estimates of the economic impact of transport infrastructure are likely to be strongly driven by indirect, spatial equilibrium effects. The degree of local factor mobility will mediate the strength of these indirect effects. For example, in contrast to the results described here, Banerjee, Duflo, and Qian (2012), find a negative association between per capita income and distance to a transport route in contemporary China, a context where labour mobility is highly controlled. The impacts of transport infrastructure are likely to demonstrate considerable heterogeneity, depending on initial local conditions. However, I am unable to explore this heterogeneity further in this context, as even if I observe a different effect size in places with different initial characteristics, I am unable to attribute a causal explanation to this change in the effect size.

Particularly during the late 1800s, labour and capital mobility in the United States was exceptionally high, with a large influx of migrants and foreign capital into a region where the spatial distribution of modern economic activity was still being determined. However, these characteristics may make lessons from this time period applicable to contemporary developing countries undergoing structural transformation with significant rural-urban migration and inflows of foreign direct investment (FDI), particularly in sub-Saharan Africa; in comparison, for example, to other recent estimates of the economic impacts of transport infrastructure that focus on contexts where labour mobility was highly constrained (e.g. Banerjee, Duflo, and Qian, 2012; Donaldson, 2012).

My results are consistent with other recent literature which has studied the role of transport infrastructure in urbanization and structural transformation (Michaels, 2008; Duranton, Morrow, & Turner, 2014; Duranton & Turner, 2012; Jedwab & Moradi, 2013; Berger & Enflo, 2013; Atack et al., 2010), although in contrast to these studies my emphasis is on the way in which these processes mediate the overall economic response. The results also contribute to a related literature which examines the distribution of population within cities, and in particular how transport infrastructure helps to determine this equilibrium (e.g. Baum-Snow, 2007; Glaeser et al., 2008).

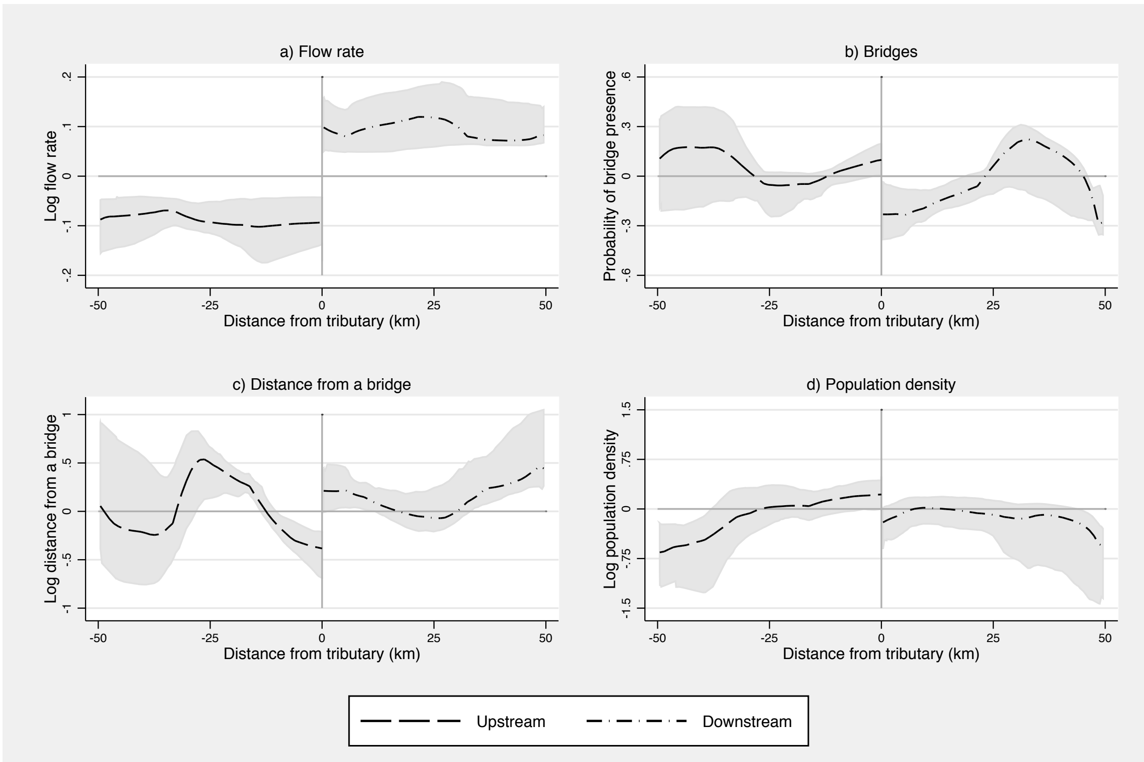
An increase in the value of agricultural land is consistent, although not directly comparable, with the results in Donaldson and Hornbeck (2013), who find an increase in value of agricultural land associated with an increase in market access through expansion of the rail network. The results on the value of agricultural land are also consistent in modern times and in a different context with the only experimental results in this literature; Gonzalez-Navarro and Quintana-Domeque (2012) find increases in housing values on streets which are randomly assigned to be asphalted. Housing, like land, is at least in the short term considered a fixed factor of production. The findings are also consistent with the work of Chandra and Thompson (2000) who found that the construction of the interstate highways resulted in shifts in industrial composition. In contrast, my analysis is silent on Chandra and Thompson's conclusion that the net impact on economic growth of transport infrastructure is zero; in this framework (as with all panel approaches) I am unable to separate out aggregate effects from overall global trends.

The results are important for researchers designing studies that aim to measure the growth or welfare impacts of transport infrastructure provision. Where labour is mobile, comparisons of differences in per capita GDP between locations with and without good access to transport infrastructure, may be biased downwards or even reversed in sign by spatial equilibrium effects including endogenous migration, urbanization and sorting. Timing also matters; structural approaches that assume full labour mobility (e.g Donaldson and Hornbeck (2013)) may underestimate overall impacts if they focus on immediate changes in outcomes such as land values, because spatial equilibrium effects take time to emerge. Assuming full labour mobility would mistakenly attribute some of the direct impacts to migration, when frictions meant that the migration response takes time to play out.

Further, the results can guide policy-makers and planners. The findings that the provision of transport infrastructure results in migration, where the labour force is free to migrate, and in increases in the value of immobile factors of production, suggest that in the absence of labour mobility, transport infrastructure construction may have more heterogeneous effects, and may possibly exacerbate or create inequalities. Where labour is free to migrate, the results suggest that planners should expect changes in population growth rates, in urbanization processes and possibly, the influx of poorer households, in response to transport infrastructure. Prior knowledge of these potential impacts could help avoid the costs of rapid, unplanned urbanization.

3.6 Tables and Figures

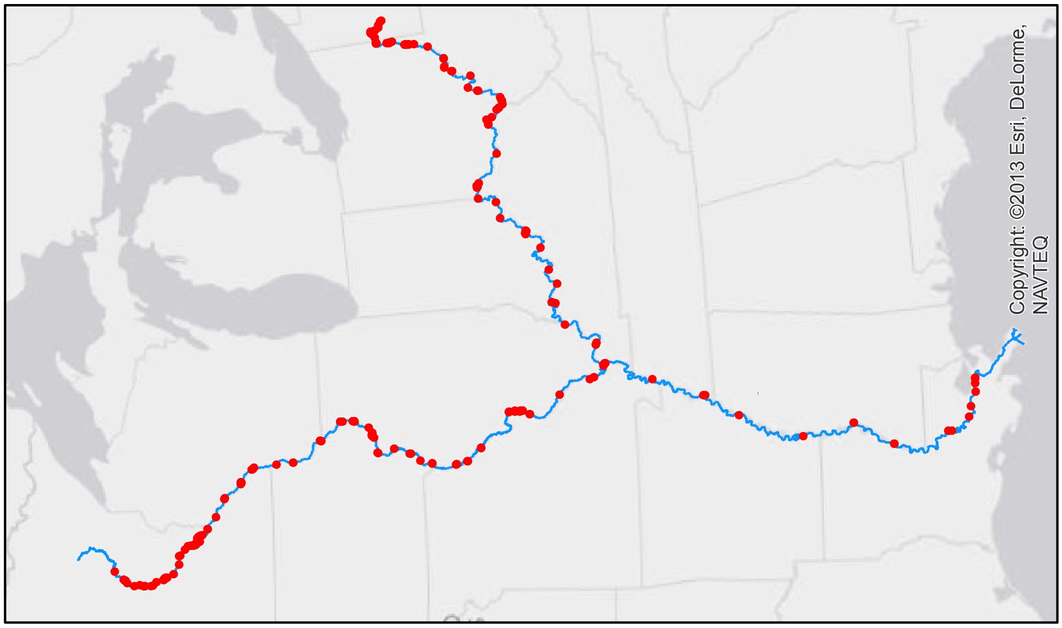
Figure 3.1: Bridge location and population density around tributaries (2000)



Graphs show results from a local linear regression of the outcome variable on distance from a tributary, for census tracts less than 50km from a tributary confluence. Outcome variables were previously demeaned around the nearest tributary. Upstream tracts and downstream tracts are pooled separately. Bootstrapped 90% confidence intervals are shown in grey, clustered at the nearest tributary.

Figure 3.2: Distribution of bridges across the Mississippi and Ohio Rivers

1860



2000

Figure 3.3: 1860 County Boundaries for counties on the Mississippi or Ohio River

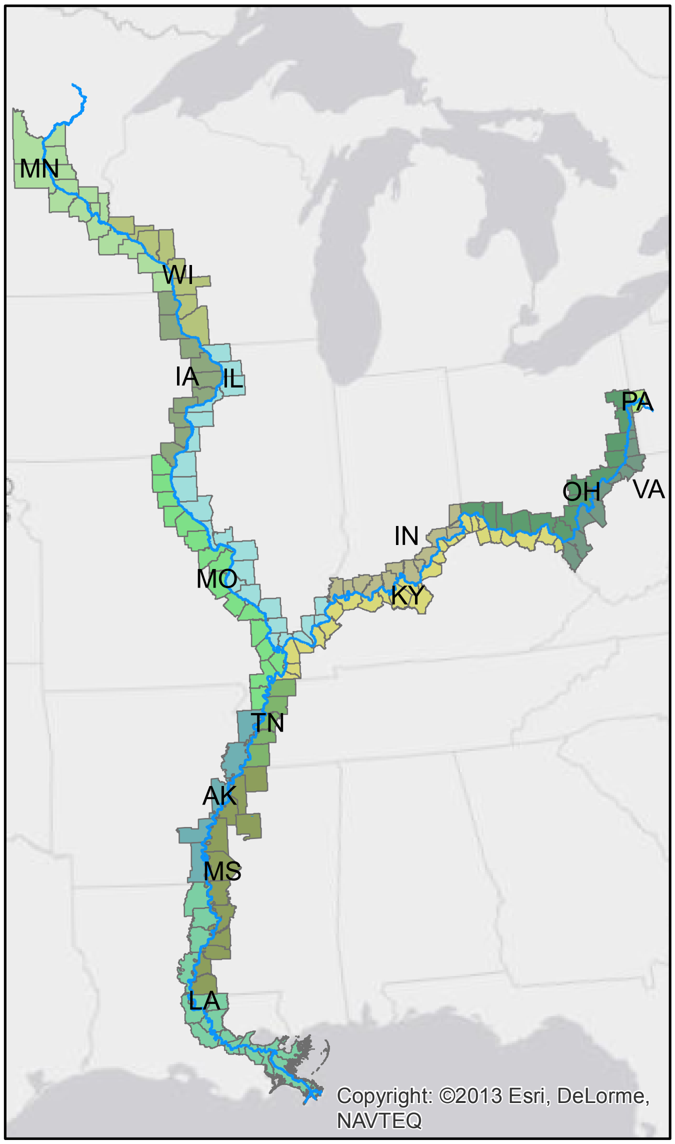
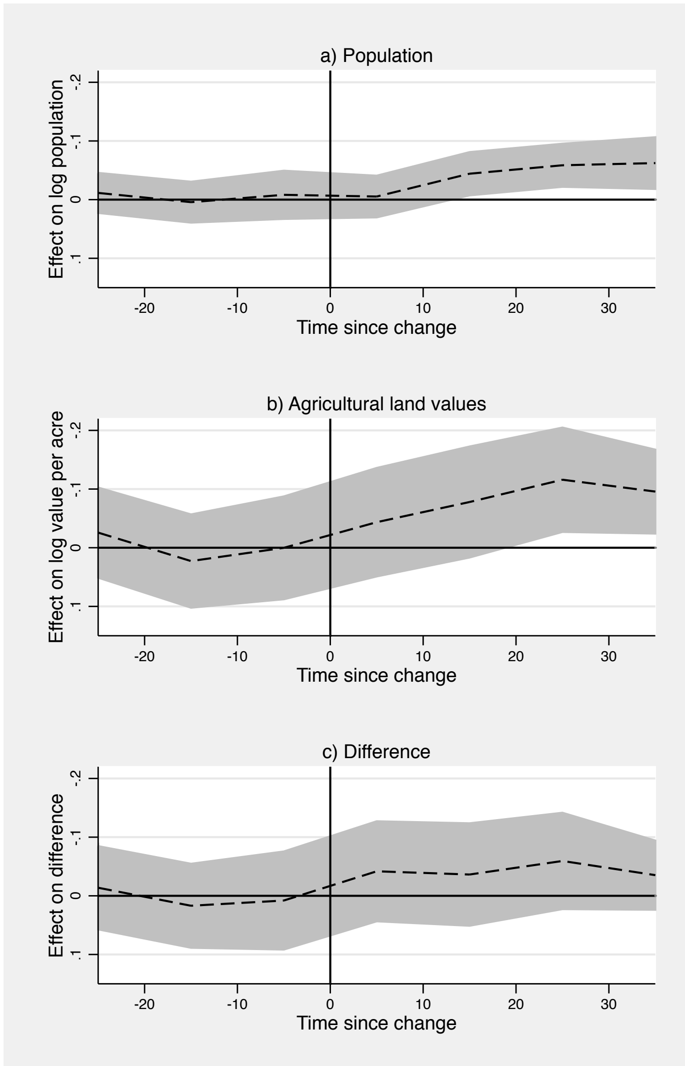
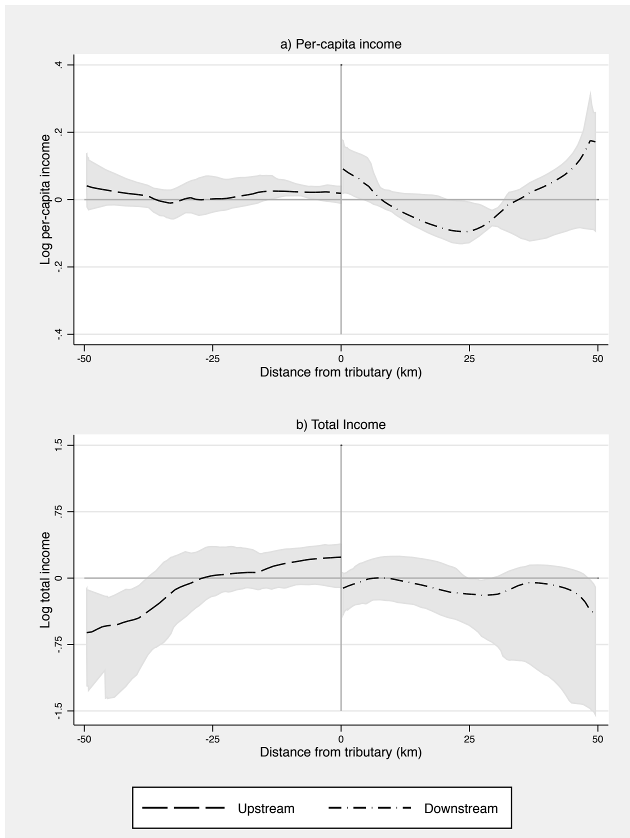


Figure 3.4: Changes in population density and value of agricultural land following a change in distance to a bridge



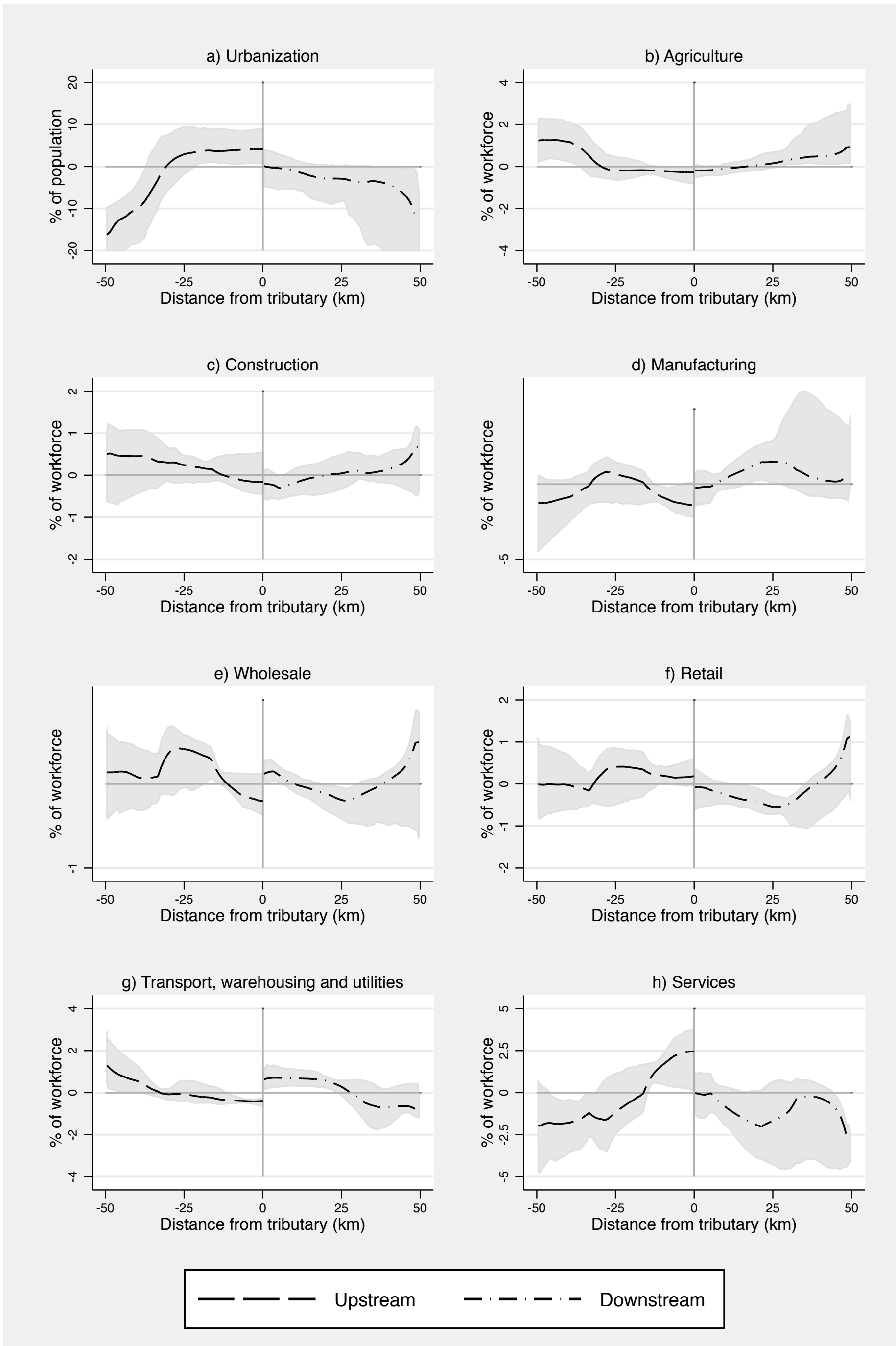
Coefficients from a regression of outcome variable on lead and lagged changes in log bridge distance, year fixed effects and county quadratic trends. Time zero is defined as the beginning of the decade in which the change in distance takes place. Coefficients on future changes in bridge distance are shown to the left of the black line, and coefficients on lagged changes in bridge distance are shown to the right of the black line. Sample consists of counties on the rivers Mississippi and Ohio, unweighted, using constant 1860 boundaries. Standard errors are clustered at the county level and robust to spatial correlation within a 200km radius. 90% confidence intervals are shaded in grey. Y-axis is reversed.

Figure 3.5: Per-capita and total income around tributaries (2000)



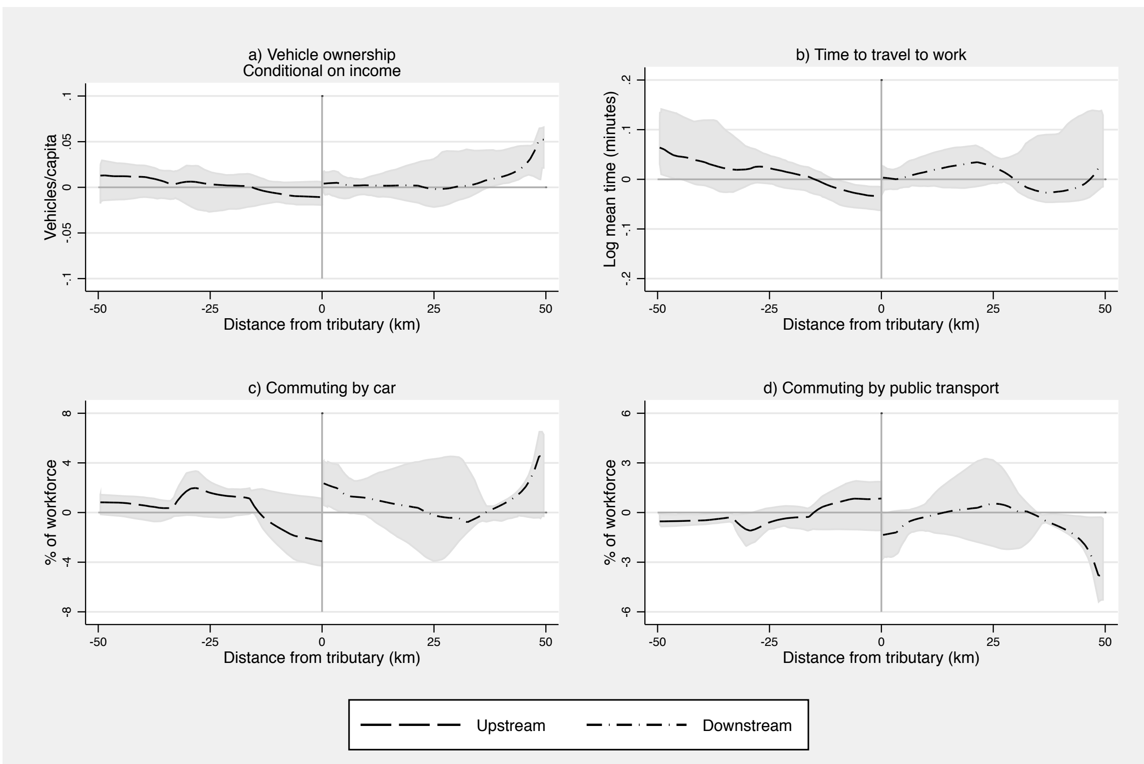
Graphs show results from a local linear regression of the outcome variable on distance from a tributary, for census tracts less than 50km from a tributary confluence. Outcome variables were previously demeaned around the nearest tributary. Upstream tracts and downstream tracts are pooled separately. Bootstrapped 90% confidence intervals are shown in grey, clustered at the nearest tributary.

Figure 3.6: Urbanization and industrial composition around tributaries (2000)



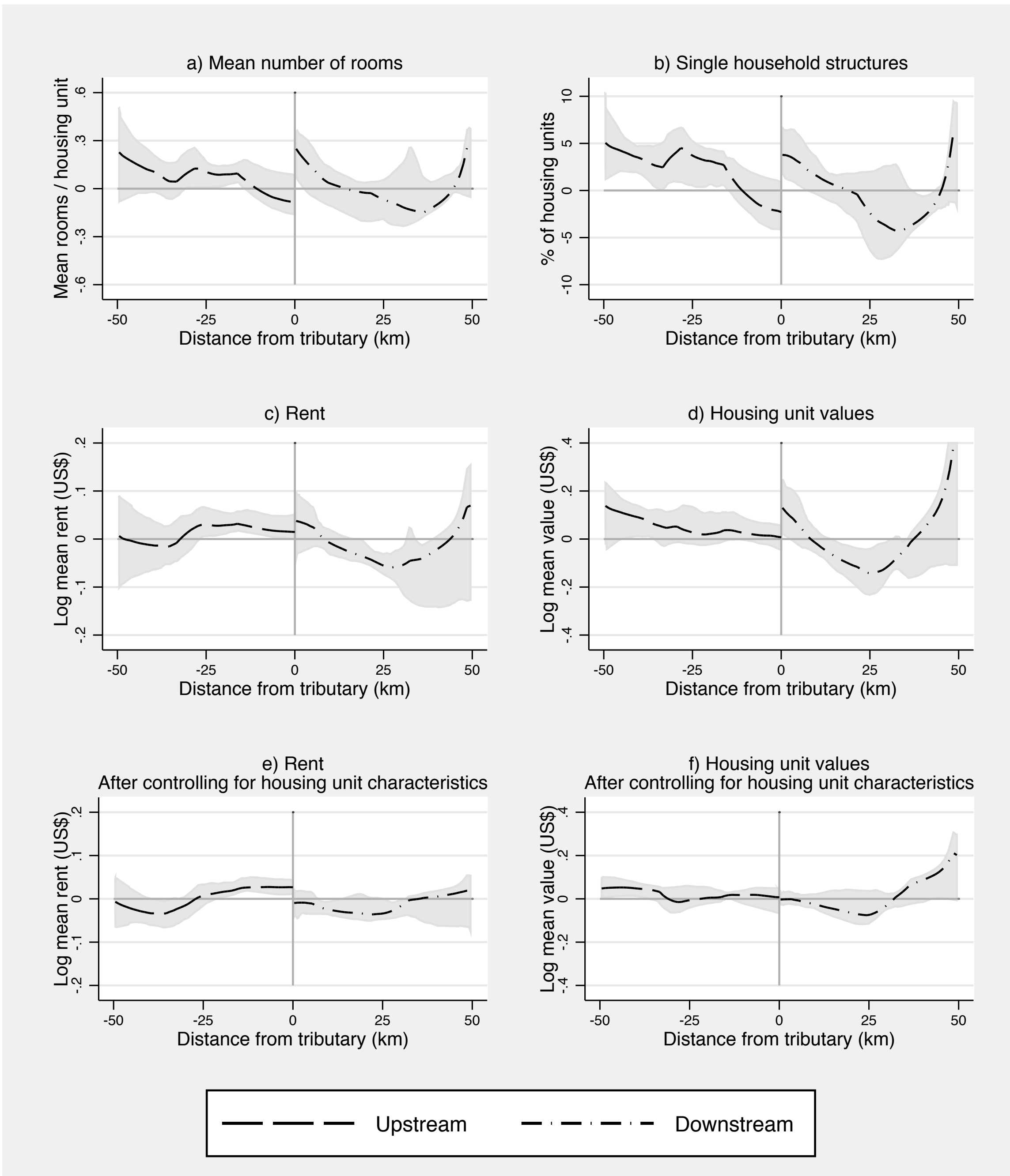
Graphs show results from a local linear regression of the outcome variable on distance from a tributary, for census tracts less than 50km from a tributary confluence. Outcome variables were previously demeaned around the nearest tributary. Upstream tracts and downstream tracts are pooled separately. Bootstrapped 90% confidence intervals are shown in grey, clustered at the nearest tributary.

Figure 3.7: Transportation patterns around tributaries (2000)



Graphs show results from a local linear regression of the outcome variable on distance from a tributary, for census tracts less than 50km from a tributary confluence. Outcome variables were previously demeaned around the nearest tributary. Upstream tracts and downstream tracts are pooled separately. Bootstrapped 90% confidence intervals are shown in grey, clustered at the nearest tributary.

Figure 3.8: Housing market characteristics around tributaries (2000)



Graphs show results from a local linear regression of the outcome variable on distance from a tributary, for census tracts less than 50km from a tributary confluence. Outcome variables were previously demeaned around the nearest tributary. Upstream tracts and downstream tracts are pooled separately. Bootstrapped 90% confidence intervals are shown in grey, clustered at the nearest tributary.

Table 3.1: Long-run Results: First stage

		Outcome Variable: Log Distance to a bridge (2000)		
		(1)	(2)	(3)
Upstream	Coeff.	-1.21***	-1.15***	-1.17***
	s.e.	0.29	0.29	0.41
	N	1287	1287	747
	F Stat	17.6	15.7	8.3
	Sample	50km	50km	1840
	State F.E.	No	Yes	No

Note: Coefficients from regressions of log distance to a bridge in year 2000 on an indicator for being upstream of the nearest tributary confluence; nearest tributary fixed effects; tributary distance and tributary distance squared, and their interactions with the upstream indicator; and pathlength from the river mouth interacted with river indicators. State F.E. included where indicated. Main sample consists of year 2000 census tracts where any part of the tract is within 10km of the Mississippi or Ohio rivers, and within 50km of a tributary. Standard errors are clustered by nearest tributary and upstream status ($N_c = 56$ for sample within 50km of tributaries.). *** p<0.01, ** p<0.05, * p<0.1.

Table 3.2: Cumulative effect on log population, log average value of agricultural land, and difference

RHS variable: Change in log bridge distance between times:		Log population (1)	Log value ag. land (2)	Difference (3)
t-10 to t	Coeff.	-0.001	-0.045	-0.047*
	s.e.	0.017	0.029	0.027
t-20 to t-10	Coeff.	-0.041***	-0.081**	-0.042
	s.e.	0.016	0.037	0.034
t-30 to t-20	Coeff.	-0.055***	-0.117***	-0.064*
	s.e.	0.018	0.037	0.034
t-40 to t-30	Coeff.	-0.060***	-0.096***	-0.038
	s.e.	0.023	0.032	0.025
	N	2715	2710	2710

Note: Coefficients from regressions of outcome variables on on lagged changes in log distance to a bridge, year fixed effects and county fixed effects and quadratic trends. Sample consists of counties on the rivers Mississippi and Ohio, unweighted, using constant 1860 boundaries. Standard errors clustered by county and robust to spatial correlation within 200km.
*** p<0.01, ** p<0.05, * p<0.1.

Table 3.3: Long-run Results: Income

		RHS variable: Log distance to a bridge (2000)				
		OLS (1)	IV (2)	IV (3)	IV (4)	IV (5)
Log population density (2000)	Coeff.	-0.67***	-0.51**	-0.48**	-0.78**	-0.76**
	s.e.	0.12	0.24	0.23	0.35	0.32
	N	1285	1285	1285	746	746
Log per capita income (2000)	Coeff.	-0.03	0.18***	0.19***	0.19**	0.20**
	s.e.	0.02	0.06	0.07	0.09	0.09
	N	1284	1284	1284	745	745
Log total income/unit area (2000)	Coeff.	-0.70***	-0.35	-0.30	-0.60*	-0.59**
	s.e.	0.13	0.22	0.21	0.31	0.28
	N	1284	1284	1284	745	745
	Sample	50km	50km	50km	1840	1840
	Nearest tributary F.E.	No	No	No	Yes	Yes
	Geographical controls	No	No	No	Yes	Yes
	State F.E.	No	Yes	No	No	No
	Controls for 1840 pop. density	No	No	No	Yes	Yes

Note: Coefficients from regressions of outcome variable listed on log distance from a bridge, instrumented by an indicator for being upstream of the nearest tributary confluence where indicated. Main sample consists of year 2000 census tracts where any part of the tract is within 10km of the Mississippi or Ohio rivers and within 50km of a tributary. Standard errors are clustered by nearest tributary and upstream status ($N_c = 56$ for sample within 50km of tributaries.).
*** p<0.01, ** p<0.05, * p<0.1.

Table 3.4: Short-run Impacts: Urbanization and industrial composition in the workforce

RHS variable: Change in log bridge distance between times:		% urban	Outcome Variable: Percentage of workforce in industry at time t					
			Manufacturing	Agriculture	Construction	Transportation	Retail & wholesale	Services
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
t-10 to t	Coeff.	-0.232	-0.593	1.487*	-0.483*	0.143	-0.371	0.116
	s.e.	0.553	0.413	0.791	0.263	0.220	0.381	0.430
t-20 to t-10	Coeff.	-1.248**	-0.854**	2.989***	-0.366	0.083	-0.799*	-0.250
	s.e.	0.595	0.420	0.896	0.290	0.191	0.452	0.377
t-30 to t-20	Coeff.	-1.693***	-0.586	2.756***	-0.443*	0.127	-0.858**	-0.297
	s.e.	0.577	0.467	0.780	0.236	0.242	0.398	0.355
t-40 to t-30	Coeff.	-1.823***	-0.253	3.095***	-0.511**	-0.238	-0.865**	-0.742**
	s.e.	0.558	0.434	0.745	0.202	0.214	0.340	0.296
	N	2713	2499	2499	2499	2499	2499	2499

Note: Coefficients from regressions of outcome variable on lagged changes in log distance to a bridge, year fixed effects and county quadratic trends. Sample consists of counties on the rivers Mississippi and Ohio, unweighted, using constant 1860 boundaries. Standard errors clustered by county and robust to spatial correlation within 200km. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.5: Long-run Results: Urbanization and industrial composition in the workforce

		RHS variable: Log distance to a bridge (2000)				
		OLS (1)	IV (2)	IV (3)	IV (4)	IV (5)
Panel a): Urbanization						
% population urban	Coeff.	-10.99***	-3.23	-2.10	-8.65	-8.50
	s.e.	2.13	4.48	4.53	6.34	6.12
	N	1284	1284	1284	745	745
Panel b): % workforce in selected industry						
Agriculture	Coeff.	0.92***	-0.06	-0.25	0.22	0.21
	s.e.	0.22	0.32	0.34	0.46	0.42
	N	1283	1283	1283	745	745
Construction	Coeff.	0.61***	-0.16	-0.23	-0.07	-0.07
	s.e.	0.17	0.47	0.52	0.63	0.62
	N	1283	1283	1283	745	745
Manufacturing	Coeff.	0.99**	0.23	0.37	0.51	0.48
	s.e.	0.47	1.33	1.32	1.59	1.58
	N	1283	1283	1283	745	745
Wholesale	Coeff.	0.05	0.61**	0.64**	0.56*	0.56*
	s.e.	0.05	0.29	0.31	0.30	0.30
	N	1283	1283	1283	745	745
Retail	Coeff.	0.08	0.25	0.29	0.51	0.51
	s.e.	0.13	0.65	0.68	0.45	0.45
	N	1283	1283	1283	745	745
Transportation	Coeff.	0.31**	1.22***	1.31***	1.48***	1.48***
	s.e.	0.15	0.28	0.27	0.47	0.47
	N	1283	1283	1283	745	745
Services	Coeff.	-3.12***	-2.12	-2.18	-3.01	-2.96
	s.e.	0.94	2.32	2.37	2.28	2.20
	N	1283	1283	1283	745	745
Sample		50km	50km	1840	1840	1840
Nearest tributary F.E.		No	Yes	Yes	Yes	Yes
Geographical controls		No	Yes	Yes	Yes	Yes
State F.E.		No	Yes	No	No	No
Controls for 1840 pop. density		No	No	No	Yes	Yes

Note: Coefficients from regressions of outcome variable listed on log distance from a bridge, instrumented by an indicator for being upstream of the nearest tributary confluence where indicated. Main sample consists of year 2000 census tracts where any part of the tract is within 10km of the Mississippi or Ohio rivers and within 50km of a tributary. Standard errors are clustered by nearest tributary and upstream status ($N_c = 56$ for sample within 50km of tributaries.).
*** p<0.01, ** p<0.05, * p<0.1.

Table 3.6: Long-run Results: Transport

		RHS variable: Log distance to a bridge (2000)			
		OLS	OLS	IV	IV
		(1)	(2)	(3)	(4)
Vehicles / capita	Coeff.	0.025***	0.033***	0.073*	0.021
	s.e.	0.004	0.007	0.038	0.024
	N	1285	1284	1285	1284
Log time to work (minutes)	Coeff.	0.031***		0.042	
	s.e.	0.007		0.039	
	N	1283		1283	
% commuting by car	Coeff.	2.45***		7.00***	
	s.e.	1.05		3.07	
	N	1283		1283	
% commuting by public transport	Coeff.	-1.73***		-3.73***	
	s.e.	0.62		2.08	
	N	1283		1283	
Nearest tributary F.E.		No	No	Yes	Yes
Geographical controls		No	No	Yes	Yes
Controls for per capita income		No	Yes	No	Yes

Note: Coefficients from regressions of outcome variable listed on log distance from a bridge, instrumented by an indicator for being upstream of the nearest tributary confluence where indicated. Sample consists of year 2000 census tracts where any part of the tract is within 10km of the Mississippi or Ohio rivers and within 50km of a tributary. Standard errors are clustered by nearest tributary and upstream status ($N_c = 56$ for sample within 50km of tributaries.). *** p<0.01, ** p<0.05, * p<0.1.

Table 3.7: Long-run Results: Housing

		RHS variable: Log distance to a bridge (2000)			
		OLS (1)	OLS (2)	IV (3)	IV (4)
Log rent (US$)	Coeff.	-0.06***	-0.03***	0.10**	-0.03
	s.e.	0.02	0.01	0.05	0.04
	N	1274	1274	1274	1274
Log housing values (US$)	Coeff.	-0.08**	-0.07**	0.32**	0.04
	s.e.	0.04	0.03	0.15	0.05
	N	1280	1280	1280	1280
Mean rooms / housing unit	Coeff.	0.09**		0.57***	
	s.e.	0.03		0.17	
	N	1283		1283	
Housing unit single household structure	Coeff.	0.04***		0.12**	
	s.e.	0.01		0.06	
	N	1275		1275	
Nearest tributary F.E.		No	No	Yes	Yes
Geographical controls		No	No	Yes	Yes
Housing characteristic controls		No	Yes	No	Yes

Note: Coefficients from regressions of outcome variable listed on log distance from a bridge, instrumented by an indicator for being upstream of the nearest tributary confluence where indicated. Sample consists of year 2000 census tracts where any part of the tract is within 10km of the Mississippi or Ohio rivers and within 50km of a tributary. Housing controls (tract level means) include number of rooms, number and type of housing units in structure, and building age. Standard errors are clustered by nearest tributary and upstream status ($N_c = 56$ for sample within 50km of tributaries.). *** p<0.01, ** p<0.05, * p<0.1.

3.7 Appendices to Chapter 3

Appendix A: Data

Census Data

Values of agricultural land I use data on values of agricultural land from ICPSR databases. After 1950, data is not available at decadal intervals (as the agricultural surveys do not coincide with census years). For 1960, I use 1959 data; for 1970, I use 1969 data; for 1980, I use 1982 data; for 1990, I use 1992 data; and for 2000, I use 2002 data. No other modifications were required to the data.

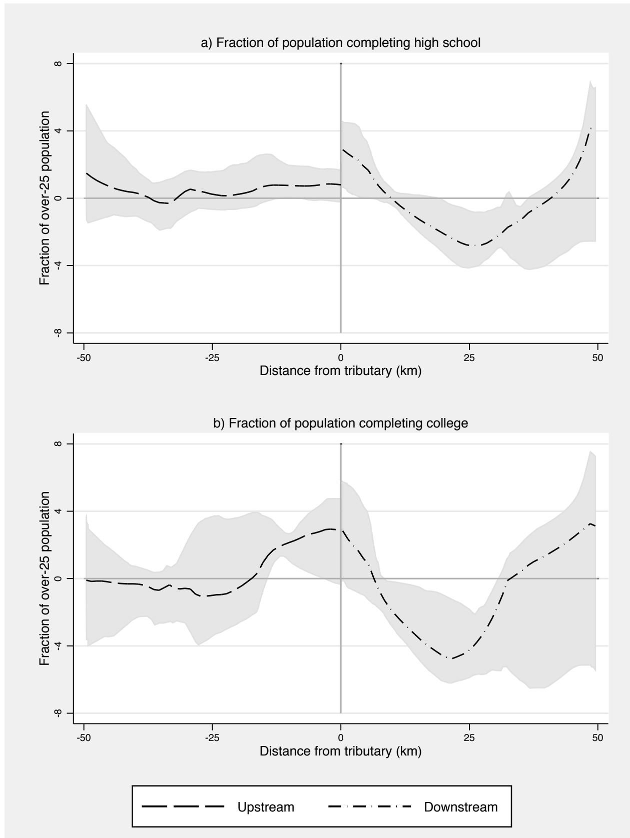
Urban and rural population Urban population is defined as living in places with 2500 or more people, following the census convention. The data on urbanization is taken from the ICPSR database for years up to and including 1980, and from the NHGIS database for years 1990 and 2000.

Industrial composition of the workforce Data is not consistently available for all sectors for the entire time period studied. For 1970 onwards, aggregate data is available from the NHGIS. For years prior to 1970, I use data from the ICPSR, when it is available, and otherwise use data from IPUMS to estimate the fraction of the workforce in each industry. In general data is missing for 1890, and for 1960 for all sectors; and IPUMS data is also missing for several counties, for which no individual is included in the IPUMS database.

Education Data For each year, I rank the counties within my sample according to an appropriate education metric. For 1860-1930, I use the fraction of the population illiterate. In 1870, 1910, 1920 and 1930 I use aggregate data from the NHGIS database, dividing by the total population, while for 1860, 1880 and 1900, I use household level data from the IPUMS database to estimate the mean fraction of households illiterate. Since the 1860 data is for the free population only, I apply a correction for the fraction of the population in slavery, assuming that 10% of the slave population is literate. Data from 1890 is missing. For 1940-2000, I use the fraction of the population with more than 8th or 9th grade education, using aggregate data from the ICPSR for 1940 and 1950 and from the NHGIS for 1970-2000. Data from 1960 are missing. Although the data is for the over 25 population, I divide by the total population. I then generate a rank within the sample for the level of education and normalize this rank by the number of counties for which data is available in a given decade. The resulting variable can be interpreted as the percentile i.e. a county with a score of 0.5 has a median level of education.

Appendix B: Figures and Tables

Figure 3.9: Variation in education levels around tributaries (2000)



Graphs show results from a local linear regression of the outcome variable on distance from a tributary, for census tracts less than 50km from a tributary confluence. Outcome variables were previously demeaned around the nearest tributary. Upstream tracts and downstream tracts are pooled separately. Bootstrapped 90% confidence intervals are shown in grey, clustered at the nearest tributary.

Table 3.8: Cumulative short-run effect on mean education levels

RHS variable: Change in log bridge distance between times:		Outcome Variable: Education percentile at time t		
		All (1)	Early (2)	Late (3)
t-10 to t	Coeff.	-0.004	0.024**	-0.032
	s.e.	0.013	0.010	0.023
t-20 to t-10	Coeff.	-0.010	0.031**	-0.064***
	s.e.	0.014	0.014	0.023
t-30 to t-20	Coeff.	-0.009	0.025**	-0.048**
	s.e.	0.012	0.012	0.020
t-40 to t-30	Coeff.	-0.015	0.018	-0.044*
	s.e.	0.013	0.015	0.026
	N	2346	2346	2346

Note: Coefficients from regressions of within-sample education percentile on lagged changes in log distance to a bridge, year fixed effects and county fixed effects and quadratic trends. Sample consists of counties on the rivers Mississippi and Ohio, unweighted, using constant 1860 boundaries. Standard errors clustered by county and robust to spatial correlation within 200km.
*** p<0.01, ** p<0.05, * p<0.1.

Table 3.9: Long-run Results: Education levels

LHS variable: % of over-25 population:		RHS variable: Log distance to a bridge (2000)			
		OLS (1)	OLS (2)	IV (3)	IV (4)
Completing high school	Coeff.	-1.02**	-0.31**	5.63**	1.72
	s.e.	0.45	0.13	2.38	1.45
	N	1284	1284	1284	1284
With a college degree	Coeff.	-3.43**	-2.48***	3.22	-2.49
	s.e.	1.42	0.89	3.02	2.84
	N	1284	1284	1284	1284
Nearest tributary F.E.		No	No	Yes	Yes
Geographical controls		No	No	Yes	Yes
Controls for per capita income		No	Yes	No	Yes

Note: Coefficients from regressions of outcome variable listed on log distance from a bridge, instrumented by an indicator for being upstream of the nearest tributary confluence where indicated. Main sample consists of year 2000 census tracts where any part of the tract is within 10km of the Mississippi or Ohio rivers and within 50km of a tributary. Standard errors are clustered by nearest tributary and upstream status ($N_c = 56$ for sample within 50km of tributaries.).
*** p<0.01, ** p<0.05, * p<0.1.

Chapter 4

Community Participation in Decision-Making: Evidence from an experiment in providing safe drinking water in Bangladesh

With Malgosia Madajewicz and Ahasan Habib[1]

[1]Grant 0624256 *DHB: Decentralization and Local Public Goods: How Does Allocation of Decision-making Authority Affect Provision? (December 2006 November 2011) Awarded amount to date:* 749926$PI : MalgosiaMadajewicz$

Abstract

The hypothesis that participation in decision-making by intended beneficiaries of social programs improves the outcomes of those programs has long been influential in the academic literature and in policy. This paper presents the first experimental evidence on the effect of transferring decision-making authority to targeted beneficiaries on the impact of a project to provide a local public good. We randomly assigned participatory and non-participatory decision-making structures to communities who received an otherwise identical intervention, a package of technical advices and subsidies to provide safe drinking water sources. Participation in decision-making resulted in larger reported increases in access to safe drinking water, but only when we imposed rules on the decision-making process that were designed to limit the appropriation of project benefits by elite or influential groups or individuals. Villages in which communities participated in decision-making under rules designed to prevent appropriation reported a significantly greater increase in access to safe drinking water (an increase of 25%) relative to villages in which project staff took decisions (14%). In villages in which the communities participated in decision-making without imposed rules, the change in access to safe drinking water was the same (14%) as in villages in which project staff took decisions. We conclude that the rules we applied to limit appropriation – minimum representation requirements and decision by unanimous consensus – were effective in accomplishing their objective.

4.1 Introduction

The hypothesis that participation in decision-making by intended beneficiaries of social programs improves the outcomes of those programs has been influential in the academic literature and in policy for some time (e.g. Stiglitz, 2002; World Bank, 2004). Advocates of the policy argue that involving communities in project decision-making has multiple benefits: improving project targeting, by drawing on information available to the community but not to outsiders; increasing 'buy–in' and generating a 'sense of ownership' of the project, thereby improving long-term management and increasing maintenance of program assets; and promoting transparency and accountability in project delivery. However, programs in which communities participate in decision-making may be more susceptible to the 'capture' of project benefits by elite or influential community members[2].

Much of the early evidence in support of this hypothesis was based on cross-sectional analyses[3], case studies[4], or was simply anecdotal. Since the choice of a decision-making structure is likely to be otherwise correlated with project, community and implementing agency characteristics, identification of causal effects is difficult and sensitive to critical assumptions. This paper presents the first experimental evidence on whether transferring decision-making authority to intended beneficiaries affects the impact of a project to provide a local public good.

We randomly assigned different decision-making structures to communities who received an otherwise identical intervention, a package of subsidies and technical advice to provide up to three sources of safe drinking water. Many rural Bangladeshi communities currently use sources of water that are susceptible to arsenic or, less commonly, bacterial contamination. Arsenic-safe drinking water sources are relatively expensive and the vast majority of households cannot afford to obtain them for themselves. As a result, the sources must generally

[2] See Mansuri and Rao (2013) for a comprehensive review.

[3] Examples include: Isham, Narayan, and Pritchett (1995), Sara and Katz (1997), A. Khwaja (2004), Fritzen (2007), A. I. Khwaja (2009).

[4] Examples include: Kleemeier (2000), Fung and Wright (2003), Rao and Ibáñez (2005).

be provided at a community level. The random assignment ensured that the communities in which we implemented the project under different decision-making structures were comparable in terms of all other characteristics, allowing us to draw causal inferences about the impacts of the decision-making structures on project outcomes.

The three decision-making structures assigned included a non-participatory decision-making structure and two participatory decision-making structures. In the non-participatory decision-making structure, project staff took all decisions, based on information provided by the community. In the first participatory decision-making structure, the community took all decisions using their own internal decision-making processes. This process was designed to approximate the way in which 'participation' is implemented by organizations which place a high value on minimizing interference with local institutions. In the second, we imposed rules on the decision-making process. Under these rules, the community took all decisions by unanimous consensus at a meeting organized by project staff, with requirements imposed for representation of women and the poor. This process was designed to approximate the way in which other organizations implement 'participation', which actively aim to broaden participation and reduce elite influence in decision-making.

Under all decision-making models, we retained an important participatory component. After decisions were taken, all treatment villages were required to contribute between 10 and 20% of the total cost of water source installation. The communities then had to decide whether or not they would contribute, and how this contribution would be raised. We therefore identify the effects of participation and decision-making over and above the effects generated by any financial contribution.

Overall, the intervention led to an increase in reported access to safe drinking water of 16% relative to a control group. The average treatment effect rises to 18%, compared to a matched control group, when we exclude a subset of villages in which the only feasible technology for providing arsenic-safe drinking water year-round was an arsenic iron removal plant (AIRP). This technology has experienced issues with reliability and effectiveness in the past (Hossain et al., 2005) and our experience suggests that communities strongly prefer tubewells to AIRPs. The treatment effect in the villages in which AIRPs are the only technically feasible option is not statistically different from zero, compared to a matched control group.

The increase in access to safe drinking water was higher in villages in which the community took decisions and in which decision-making rules were imposed (22% in all villages; 25% if we exclude the AIRP villages) compared to the villages in which project staff took decisions (13%; 14% if we exclude the AIRP villages). However, no differences were observed between the increases in access to safe drinking water when the community took decisions without the imposition of decision-making rules (14%; 15% if we exclude the AIRP villages) and when project staff took decisions. The difference between the change in reported access to safe drinking water in villages in which the community took decisions under imposed rules and the remainder of the treated villages is significant when we remove the villages in which AIRPs were the only option from the analysis. Since the treatment effect is zero in these villages regardless of the structure under which decisions were taken, including them in the analysis is not informative with regards to a comparison between decision-making structures.

We installed an average of 2.1 arsenic safe water sources in each of 127 treatment villages. We installed a slightly larger number of wells in villages in which the community was involved in decision-making (2.2 across both participatory decision-making structures) compared to those in which project staff took decisions (2.0). However, the differences are not statistically significant. Under the non-participatory structure, project staff were instructed to propose locations for water sources in public spaces wherever feasible in order to facilitate access to the sources. Under the participatory structures, communities were more likely to locate the water sources on private land. We installed 1.9 sources per village on public land when project staff made decisions, and 1.3 when communities took decisions. A significantly smaller number of individuals contributed money towards the water sources in the communities which took decisions without any imposed rules (5 individuals per village), when compared to the other two models (9 individuals).

Recent experimental studies have explored several aspects of 'participation'. One influential group of experimental studies examine the impact of a participatory or 'community-driven' development project compared to a control group which does not receive any intervention (Fearon, Humphreys, & Weinstein, 2009, 2011; Humphreys, de la Sierra, & van der Windt, 2012; Casey, Glennerster, & Miguel, 2012)[5] . The most closely related studies to this one explore variations in how participatory decision-making processes are implemented in projects to provide a local public good, conditional on implementing some kind of participatory decision-making approach: Olken (2010) compares decisions taken at representative-based meetings to those taken by direct election-based plebiscites; Beath, Christia, and Enikolopov (2013) compare decisions taken by secret ballot referenda to those taken at consultation meetings. Our study differs from these two studies in three ways. First, while these studies infer that participation in decision-making does influence decisions taken or other outcomes, since changes in the participatory process alter these outcomes, they do not directly measure the effect of introducing participation in decision-making itself. By including a treatment group in which the project is implemented, under otherwise identical conditions, without community participation in decision-making, we are able to measure the effect of introducing community participation in decision-making. Second, the two participatory decision-making processes we compare differ from those that these studies consider; neither decision-making by consensus nor decision-making without any imposed rules have previously been explored. Finally, the preceding studies have so far only reported results on how changing the participatory process alters the decisions taken, while we are also able to

[5]Another related group of experimental studies focuses on varying requirements for participation of women in a particular decision-making process (Humphreys et al., 2012; Casey et al., 2012) or institution (Chattopadhyay & Duflo, 2004). Other related experiments examine changes in incentives to participate in school monitoring committees (Banerjee, Banerji, Duflo, Glennerster, & Khemani, 2010) or changes to the institutional structure of those committees (Pradhan et al., 2014); participation in monitoring of road construction projects (Olken, 2007) and public health care providers (Björkman & Svensson, 2009); participation in project targeting for household-level interventions (Alatas et al., 2013); and dispute resolution training to improve informal institutions (Blattman, Hartman, & Blair, 2014).

report data on the project impacts[6].

Our results confirm that involving communities in decision-making can lead to greater project impacts in terms of number of projects successfully completed and changes in reported access to safe drinking water. However, the results also suggest that devolving decision-making authority to the community without measures to avoid co-option of the decision-making process by influential groups or individuals can lead to an increased incidence of elite capture. In our case, the number of safe water sources constructed increases without any reported increase in access to safe drinking water.

The paper is structured as follows. Section 4.2 describes the setting, the experimental design and the data; section 4.3 describes the results, and section 4.4 concludes.

4.2 Setting, Experimental Design and Data

4.2.1 Arsenic Pollution Problem in Bangladesh

The context for this study is the arsenic contamination problem in rural Bangladesh. In the 1970s and early 1980s, many international agencies promoted the use of groundwater — water from wells — as a safer alternative to surface water — collected from ponds or rivers — which is often contaminated by pathogens. At the time, noone had realized that groundwater in the region sometimes has naturally occurring high concentrations of arsenic. Arsenic contamination is not readily detectable in water, and symptoms of arsenic poisoning only appear after years of exposure and accumulation in the body. Information about high concentrations of arsenic in tubewells emerged only in the mid-1990s. By that time, the damage was done; the resulting epidemic of diseases associated with arsenic exposure has been described as 'the largest poisoning of a population in history' (Smith, Lingas, & Rahman, 2000). In 2008, when this project began, UNICEF estimated that 20 million people

[6]Olken (2010) reports results on decisions taken, consistency with preferences of different groups within the community, knowledge about the project, and satisfaction with the project processes; Beath et al. (2013) report results on project satisfaction, and consistency of decisions with ex-ante preferences.

were still using water from wells with arsenic concentrations above the Bangladeshi standard, which is itself five times higher than the WHO standard (UNICEF, 2008).

Creating access to safe drinking water in the presence of arsenic contamination presents a problem of providing a local public good. The great majority of drinking water sources in Bangladesh are privately owned, including almost all tubewells that have high concentrations of arsenic. Technologies to provide water with low concentrations of arsenic are considerably more expensive, and entail high fixed costs. Only the richest households can afford to purchase these sources themselves. For most households, they must be provided at the community level, at which high fixed costs can be shared among many people. As a result, communities who wish to improve access to safe drinking water must typically solve a collective action problem.

Several technologies are available to provide arsenic-safe drinking water, of which deep tubewells are the most common in rural Bangladesh. Deep tubewells draw water from deep aquifers (approximately 700-800 feet below ground level) that have low concentrations of arsenic. Standard deep tubewells are relatively expensive to install, but easy to use and maintain, and replacement parts are readily available. In some areas, arsenic safe water is available at lesser depths of approximately 300-400 feet. In these areas, shallow tubewells can provide arsenic-safe drinking water, at a lower installation than deep tubewells. Shallow tubewells are otherwise very similar to deep tubewells in terms of functionality, maintenance requirements and ease of repair.[7] In some areas, there is considerable seasonal variation in water pressure in the aquifer and standard deep tubewells may not provide year-round access to safe drinking water. An alternative design — the deep-set tubewell — can provide year-round access to safe drinking water in these areas. The pumping mechanism in the deep-set tubewell is installed below the surface of the ground, as opposed to on the surface in the standard design. This means that the deep-set tubewell is more expensive and more difficult

[7]During the study implementation period, information emerged about a problem of manganese contamination in shallow tubewells. As a result, we replaced shallow tubewells we had already installed free of charge with alternative technologies, if they tested positive for manganese.

to repair in case of failure than the standard deep tubewell, but it is equally convenient and easy to use.

In some areas, there is no accessible arsenic-safe aquifer – for example, where an intermediate layer of rock cannot be penetrated using local drilling techniques – and therefore it is not feasible to install tubewells. An alternative technology is the arsenic iron removal plant (AIRP). AIRPs remove arsenic by oxidation and filtration. They are more expensive, larger and significantly more difficult to operate and maintain than tubewells, and our experience suggested that communities strongly preferred tubewells.[8] As a result, we will throughout the paper report treatment effects by the type of feasible technology – AIRPs or tubewells – as well as the overall treatment effect.

4.2.2 Experimental Design

The project intervention consisted of a package of technical advice and subsidies for the provision of up to three safe drinking water sources per community. We carried out the interventions between 2008 and 2011, in partnership with a Bangladeshi non-governmental organization (NGO), NGO Forum for Public Health. NGO Forum for Public Health is a well-established actor in the water and sanitation sector with more than 30 years experience in the field.

We conducted our study in communities located in two *upazilas* (subdistricts): Gopalganj, about 60 miles southwest of Dhaka, and Matlab, about 30 miles southeast of Dhaka. We focused on these sites because of the severity of the arsenic contamination problem in the area more than 80% of pre-existing tubewells were arsenic contaminated and because the sites had not yet received other interventions to address the problem. We studied 250 villages, equally split between the two upazilas, and ranging in size from a minimum of 7 households

[8] Where tubewells were not feasible, we also offered communities the opportunity to install rainwater harvesting systems or a pond sand filter, but since no community selected either of these options, we do not describe them further in the paper. Both technologies have limitations with respect to tubewells or AIRPs

to a maximum of 1103, with the median size 170 households.[9] Before interventions began, we carried out an information campaign about the arsenic problem, to ensure that all villages were initially equally well informed about the arsenic problem.

Of the 250 villages studied, we assigned 100 to a control group who did not receive the intervention. 126 villages received the intervention. We initially assigned a further 24 villages to receive the intervention who eventually did not receive the intervention, due to changes in the costs of providing safe water sources over the course of the project. We originally assigned one other village to treatment, but project staff determined before the project began that there were no feasible available technologies to provide safe drinking water in the community, because no arsenic safe aquifer was accessible, and arsenic concentrations in the shallow groundwater were too high for removal with an AIRP. There was one other village in which we determined after we began the intervention that there were no feasible available technologies to provide safe drinking water.

The original protocol for selection of treated villages was random, which should have resulted in treatment and control groups which were comparable at baseline. However, we later established that the project director at the time, who was later removed from the project for unrelated reasons, did not follow the original protocol when he implemented the division of the villages into control and study villages, and he included all villages in the southern area of Matlab in the treatment group. Villages in South Matlab have much lower access to safe drinking water than the average village in the sample, meaning that overall the treated group had significantly lower access to safe drinking water at baseline than the control group.

Table 4.1 confirms that this resulted in statistically significant differences between control and treatment groups. Treated villages had reported lower access to safe drinking water, and were less likely to have changed their source of drinking water because of the arsenic contamination problem in the last five years. In Table 4.1, we show baseline summary

[9] Data on arsenic contamination of pre-existing tubewells and village size was drawn from the Bangladesh Arsenic Mitigation Water Supply Project.

statistics and randomization checks for villages by treatment status. The table shows the mean and standard errors for a selection of baseline variables which measure baseline access to safe drinking water, factors that might influence the ease of providing safe drinking water, and community-level variables that might influence the likelihood of a successful collective action. In column 2), we test whether the difference in means between treated and control villages is statistically significant. The p-values are derived from Ordinary Least Squares (OLS) regressions with the following structure:

$$Y_{i,v} = \alpha + \beta I_{treated,v} + \epsilon_{i,v} \tag{4.1}$$

where $Y_{i,v}$ is the value of a variable in household i in village v and $I_{treated,v}$ is an indicator which is one if village v was treated and zero if village v was not treated. If the treatment was randomly assigned, the coefficient β should be zero as assignment to treatment should not be correlated with any baseline characteristics of the village. The p-values test whether the coefficient is equal to zero.

Since treatment was assigned at the village level, but we collected data at the household level, it is important to account for within-village correlation in variables. Within-village correlation implies that it is more likely that differences between mean outcomes in treated and control villages arise due to chance, than if we had been able to assign treatment at the household level. In order to ensure that the statistical analyses we carry out make the correct inference about whether or not a result is likely to be due to chance or not, we follow Angrist and Pischke (2009) and cluster standard errors at the village level.

Columns 3) to 5) of Table 4.1 show that we can correct for the bias induced by the failure of randomization by three methods. First, we can drop South Matlab from the sample. Second, we can create a synthetic treatment variable generated at random in South Matlab, and equal to the treatment variable elsewhere [10]. This synthetic variable re-assigns a fraction

[10] Ideally, we would have used the original random assignment to treatment rather than this synthetic alternative

of the villages in each treatment group in South Matlab to control. Third, we can use this synthetic treatment variable to instrument for treatment. In columns 3) to 5), we report the difference in means between treated and control villages under these three approaches, after accounting for the different proportions of treated villages in Gopalganj and Matlab, because differences in treatment and control groups otherwise reflect differences between these areas. We estimate the difference in means using an equation similar to Equation 4.1, including indicators for Gopalganj and South Matlab. In each case we show that no significant differences remain between treatment and control populations.

Since the non-random selection of treatment villages in South Matlab may have introduced bias into our estimates of treatment effects, we therefore report both OLS results and results where treatment is instrumented using the synthetic treatment variable (the IV results).

Decision-making structures

Project staff implemented the intervention under one of three decision-making structures. The necessary decisions included if, how and where to install; and how to manage, each safe drinking water source. In all cases, project staff ensured that all decisions made were technically appropriate. Table 4.2 summarizes the main features of the different decision-making structures. We describe the three models in more detail in the following paragraphs.

The decision-making structures included one non-participatory structure, the Top-Down model (TD). Under this model, project staff took all project decisions, after an extended (typically 2-day) period of information gathering. The information gathering process consisted of participatory mapping of the village with members of the community, focusing on the locations of households and safe and unsafe sources of drinking water, cross-checking information with various community members. Project staff then proposed sites for safe drinking water sources, prioritizing locations with the highest density of households not al-

but we have not been able to recover the initial, randomly assigned treatment lists.

ready served by safe drinking water sources, choosing public locations wherever possible, and convenient locations where no suitable public land was available. Staff then organized and publicized a community meeting at which they presented the proposed locations. This model was designed to approximate the 'traditional' approach to decision-making about local public goods in which decisions are taken by a centralized organization, such as local government or an NGO.

The decision-making structures also included two participatory structures, in which decision-making authority was devolved to the community. Under the 'pure' Community Participation (CP) model, project staff visited the community to arrange a meeting at a site and time of the community's choosing. At the meeting, project staff explained the project rules and announced that they would return to the village after a few days to find out whether they wanted to participate in the project, and if so, which sites they had chosen. Sites that were not technically appropriate were rejected, but otherwise the community's decisions were final, conditional on raising the community contribution. We did not directly observe the decision-making process used, but communities reported to us that they took these decisions in a variety of ways including open meetings (sometimes but not always including women), meetings at a mosque, or closed-door meetings of village elites. This model was designed to approximate the way in which some organizations implement community participation in practice, avoiding interference with a community's internal hierarchies and decision-making processes .

Under the second participatory decision-making structure, the NGO-Facilitated Community Participation model (NGO), we imposed rules about how decisions should be taken. Project staff initially organized a series of separate small group meetings with men and women who the community identified as poor and non-poor. At these small group meetings, project staff explained the project rules and emphasised the right all individuals would have to participate in the decision-making process and benefit from the interventions. These meetings were followed by a community meeting, at which both men and women, and poor and non-poor, had to be represented. The community proposed and selected water source locations by unanimous consensus at the meeting, in the presence of project staff and with their active facilitation. If the community could not reach a consensus at the first meeting, a second and in some cases subsequent meetings were organized. This model was designed to approximate the way in which other organizations implement community participation, with project staff playing a strong facilitatory role, and rules imposed that are intended to reduce the likelihood that influential groups or individuals co-opt the decision-making process.

Before installing a safe drinking water source, we required the community to contribute between 10% and 20% of its cost, depending on the technology installed. Table 4.3 shows the cost of installing each of these technologies and the community contribution that we required. The difference in required community contributions reflects the difference in cost of the selected technology. We also scaled the community contribution so that the subsidy could be either concentrated on one water source or spread between up to three water sources. The price per water source therefore increased as more water sources were installed in the village. Budget constraints meant that when the best feasible technology was one of the more expensive alternatives, we were only able to offer up to two water sources.

After the initial decision-making process, project staff gave the communities up to twelve weeks to raise the funds for the community contribution. Construction of the safe drinking water sources began as soon as the community had raised their contribution. If after twelve weeks the community had not raised their contribution, construction of the safe drinking water sources did not go ahead. We initially intended the decision-making structures to apply to decisions about who contributed to the community contribution, but this proved impossible to enforce. However, project staff did propose a list of contributors at the Top Down model meetings, and communities did agree a list of contributors at the NGO-Facilitated Community Participation meeting.

We randomly assigned the decision-making structures to the communities who received the intervention. Of the 126 treated villages, we initially assigned 42 to each decision-making model. We replaced the village in which we determined before beginning the project that there was no feasible safe drinking water technology with another village, randomly drawn from the villages which we had initially assigned to treatment but in which we had not carried out the intervention due to budget constraints. As a result 43 villages were assigned to the Top-Down model.

Table 4.4 shows that the villages assigned to each decision-making model were comparable at baseline to the villages assigned to the other decision-making models. We test whether the difference in variable means between villages in which the project was implemented under a given decision-making structure and the remainder of the treated villages is statistically different from zero. The p-values in the table are therefore derived from OLS regressions similar in structure to Equation 4.1 but the indicator $I_{m,v}$ is one if village v received treatment under decision-making structure m, and zero otherwise:

$$Y_{i,v} = \alpha + \Sigma\beta_m I_{m,v} + \epsilon_{i,v} \tag{4.2}$$

Only the treated villages are included in the regressions in Table 4.4. We do not use the control group for comparison in this case because the results in Table 4.1 already confirm that the treated villages are not directly comparable to the control villages.

We compare 15 variables across the 3 decision-making structures, resulting in a total of 45 tests. In 43 of these tests we fail to reject at the 10% level the null hypothesis that there is no difference in means between groups treated under one decision-making structure and the other treated villages. In 2 tests we find statistically significant differences between the mean of a variable in villages treated under one decision-making structure and in the remaining treated villages. One test rejects this hypothesis at the 5% level, and one at the 10% level. This is consistent with what we would expect due to chance. From these checks we conclude that there is no evidence to suspect that assignment to model, conditional on treatment, was not random, as required by the project protocol.

The same project staff – one team in Gopalganj and one team in Matlab – implemented the project under all three decision-making structures. We implemented the intervention in cycles during which project staff would complete the entire process from meeting organization to water source installation for a group of villages, where the villages were grouped geographically for ease of logistics. The project was initially implemented in 114 villages in 6 cycles across both upazilas. We later added an additional 12 villages in Gopalganj when funds became available, in a 7th cycle.

Government policy had changed by the time we carried out the 7th cycle, and community expectations that the government would provide free tubewells may have increased. We installed fewer safe water sources under the 7th cycle, but the number installed is not significantly less than under the first 6 cycles in Gopalganj, once we account for the feasible technology.

4.2.3 Data Description

We carried out a baseline survey in 2007 in 40 households in each of the 250 villages, sampled randomly from census lists . We surveyed a total of 9797 households, as in some very small villages there were fewer than 40 households. The baseline questionnaire included standard components of a household survey with a special focus on social networks and social capital, and full details on water use behavior. We also collected village-level information from focus groups.

We encountered significant problems with the data entry process after the baseline survey. First, some of the individuals employed to enter the data in spreadsheets copied and pasted entire villages of data, changing names and other identifiers to conceal what they had done. Data checking revealed this problem by chance several months after data collection and entry had been carried out. When we discovered this problem, we checked extensively for additional incidences and had the missing data re-entered. Second, by the time we discovered this problem, termites had unfortunately attacked the stored questionnaires, and destroyed a small percentage of the questionnaires. As result, we are missing baseline data from 140 households from control and treated villages, since enumerators did not initially enter the data correctly and termites then destroyed the hard copy of the questionnaires. We do not however have any reason to think that there was any systematic pattern to either the false data entry or the losses to termites, so the remaining baseline data should still represent a randomly selected sample of the baseline population.

We carried out follow-up surveys in control and treated villages in 2010 and 2011 after we carried out the safe water intervention, interviewing the same households that we interviewed for the baseline survey. We did not carry out follow-up surveys in the 24 villages which were initially assigned to treatment but in which we did not carry out the intervention. We therefore attempted to resurvey 8,890 households from the original panel, of which we successfully re-surveyed 8630 households, representing an average attrition rate of 2.9%. The attrition rates broken down by treatment group are as follows: 2.7% in control villages; 3.1% in treated villages. Among the treated villages, attrition rates were 2.6% in NGO-Facilitated Community Participation villages; 3.2% in Community Participation villages; and 3.4% in Top-Down villages. The attrition rates in treated groups and sub-groups were not statistically different from the control group, or from each other.

We also carried out follow-up surveys in 1424 additional households in treated villages, to bring the minimum survey coverage up to 15% of households in all treated villages (based on census data). The additional households were randomly selected from the remaining households on the census lists who had not been surveyed at baseline. Extending the survey coverage in this way was intended to ensure that the survey captured the effects of the intervention in larger villages, where the three safe drinking water sources constructed were unlikely to serve the entire community. However, the data from these additional households is inconsistent with the data collected from the panel households, and we have established that there were violations of the sampling protocol; in particular, in some villages, some of these additional households were sampled from the neighbourhood of the installed water sources, rather than from the census lists. As a result, we do not use this data in this analysis.

We also collated data on the numbers and types of safe drinking water sources installed, and project staff kept detailed records of the implementation process, including the number of contributors in each community and the time taken to raise the community contribution. We also carried out focus group discussions in treatment villages to obtain qualitative information about why the project was successful in some communities and not in others.

4.3 Results

We first show how attendance at decision-making meetings varied by decision-making model to demonstrate that the NGO-Facilitated Community Participation was marginally more successful in including more people in the decision-making process, and in including a more diverse range of people. We then report the project outcomes in the study villages. Projects implemented under the participatory decision-making models were more successful in terms of installing safe water sources, but the differences between models are not statistically significant. We were far more successful in installing safe water sources in villages where tubewells were feasible, than in villages where only AIRPs were feasible.

Projects implemented under the Top Down decision-making model were however much more successful in installing safe water sources in public places, although whether an installation site is public or not is not necessarily a good predictor of how well-used the water source will be. Finally, we show that in villages implemented under the Community Participation model, fewer households contributed towards the cost of installation.

We then report the average treatment effect in terms of changes in reported access to safe drinking water for all villages. We show that overall, the project led to a 16% increase in reported access to safe drinking water. The treatment effect was higher in villages where tubewells were feasible (18%), and zero in villages where only AIRPs were feasible. We then show that the treatment effect was substantially higher in villages where the project was implemented under the NGO-Facilitated Community Participation model, and that the differences are significant when we exclude the villages in which only AIRPs were feasible; given that the average treatment effect is zero in these villages, including these villages is not informative with respect to a comparison between decision-making models. Finally, we present some robustness checks on these main results.

4.3.1 Participation

Table 4.5 shows the recorded numbers and characteristics of individuals attending the main community meetings. Overall, the mean number of participants was 30.7, with 27% of meeting attendees female, 42% of low socioeconomic status (as recorded by project staff) and 60% with less than secondary education.

More people, from a more diverse range of groups, attended meetings held under the NGO-Facilitated Community Participation decision-making model than meetings held under the other two decision-making structures. The number of participants was somewhat higher in NGO-Facilitated Community Participation model meetings (33.6), and lowest in the Top Down model meetings (28.5), but the differences between models are not statistically significant. The NGO-Facilitated Community Participation model meetings were the most diverse in representing different groups, with the highest percentage of female participants and the highest percentage of participants with less than secondary education. The Top Down meetings were the least diverse, with the lowest percentage of female participants, and the lowest participation of participants with low socio-economic status. However, again, not all the differences between models are statistically significant. Nonetheless, to the extent that attendance at meetings really reflects *de facto* participation in decision-making, the NGO-Facilitated Community Participation model was more successful in involving a larger number and wider range of community members in decision-making.

4.3.2 Project Outcomes

Table 4.6 shows how the decision-making model influenced project outcomes in the treated villages. On average, we installed 2.14 safe water sources in the treated villages. If we had installed all technically feasible water sources given our project rules, we would have installed an average of 2.75 safe water sources, based on installing three sources per village in most cases, and two sources per village where only a more expensive technology was feasible.

We offered communities the choice between all technically appropriate technologies to provide safe drinking water, given local hydrogeological conditions. In Gopalganj, we carried out the intervention in 70 villages. In 16 villages, AIRPs were the only feasible technology. In two villages, no treatment was feasible, as there was a layer of impenetrable rock, and shallow groundwater was too strongly contaminated with arsenic and iron for removal with an AIRP. In Matlab, tubewells were feasible in all villages.

A clear preference gradient between the available technologies emerged. People chose shallow tubewells wherever possible, followed by standard deep tubewells and deepset tubewells. The three types of tubewell are comparable in ease of use and maintenance, but increase in cost with depth and design complexity. AIRPs were the least preferred option by a wide margin. There were 16 villages, all in Gopalganj, where the only type of water source that could be installed was an AIRP, meaning that we could have installed a total of 32 AIRPs. We were only successful in installing 5 AIRPs during the course of the project, a success rate of approximately 16%. In comparison, in the remaining villages in Gopalganj — in which tubewells were feasible — we installed 79% of the maximum number of wells we could have installed under our project rules. The reasons given by the communities for rejection of the AIRPs were that they took up too much space, required too much work to operate and maintain, and were not perceived to be reliable or trustworthy. When we consider only the villages in which tubewells were feasible, the average number of water sources constructed rises to 2.45 out of a maximum possible 2.85.

The rejection of AIRPs did not seem to be a direct function of the price of the technology.

However, in Matlab, in the 10 villages where only deep-set tubewells could be installed (which are comparable in price to AIRPs, and for which we required the same level of community contribution), we installed on average 90% of the maximum feasible number of water sources, compared to an average of 89% in all other villages in Matlab (where either deep tubewells or shallow tubewells were feasible). However, we cannot comment on what would have happened if we had offered AIRPs at a lower price.

We installed 10% more water sources in the villages in which communities participated in decision-making than in the villages in which project staff took decisions, as shown in column 1). Installing more water sources is one measure of success of the project, but it may not translate into increased access to safe drinking water if the sources are not fully accessible to the community. However, the differences are not statistically significant. In Table 4.6, we assess whether differences in project outcomes across models are statistically significant using OLS regression for the following equation:

$$Y_v = \beta_{NGO} I_{NGO,v} + \beta_{CP} I_{CP,v} + \beta_{TD} I_{TD,v} + \epsilon_v \tag{4.3}$$

We then test pairwise equality of the coefficients β_{NGO}, β_{CP} and β_{TD}. The differences between the number of water sources installed under the different decision-making models are attenuated further in both magnitude and significance when we consider only the villages in which tubewells were feasible.

We installed more water sources in public spaces, as recorded by our project staff, under the non-participatory Top-Down model. Public spaces were defined to include communal land, open spaces, areas beside roads, and institutions such as mosques or schools, as opposed to privately owned land. Under the Top-Down model, project staff had a specific mandate to install water sources in public places. The differences are strongly significant with respect to both the participatory decision-making models. Water sources installed in public places may be accessible to a larger number of people. However, space that is appropriate for water source construction is quite strongly constrained in villages in this region, and the most convenient location for a water source may not necessarily be located on public land. This is primarily because land that is not vulnerable to flooding is relatively scarce, and safe water sources cannot be installed on land that is vulnerable to flooding because of potential contamination.

Fewer people contributed to raising the community contribution in the unregulated Community Participatory model than under the other two models. The difference is significant with respect to both the other two decision-making models. A small number of contributors may be efficient, as some community members will have a greater ability to contribute than others. However, it may also be indicative of a high degree of influence over the decision-making procedure, which may not be efficient if used to co-opt project benefits for private use.

Overall, the number of contributors was relatively low in all cases, considering that the median village size was 170 households. In villages where we successfully installed at least one safe water source, the mean number of contributors per water source installed was 5.1 in NGO villages, 2.3 in CP villages and 4.0 in TD villages. There was only one contributor per safe water source installed in 34% of the NGO villages, 56% of the CP villages and 45 % of the TD villages.

4.3.3 Reported Project Impact

We primarily measure access to safe drinking water based on an outcome variable which measures whether or not the household reports using safe drinking water. The indicator is based on the source of water that the household identifies as being its most important source of water for drinking and cooking. The indicator for reporting use of safe drinking water is constructed as being equal to one where the household reports using a source of drinking water that is safe from both bacterial and arsenic contamination, and zero when they report that the source is unsafe, if they don't know whether it is safe or not, or if it is a source that is vulnerable to bacterial contamination e.g. a dug well or surface water. Further details regarding the construction of this variable is included in Appendix A.

We report the average overall treatment, relative to the full control group, but we also break down the treatment effect by whether tubewells or only AIRPs were feasible. There is strong spatial correlation between locations where only AIRPs are feasible, reflecting the extent of the rock layer overlaying the deep aquifer. Since other village level characteristics are also spatially correlated, there are as a result some differences on baseline characteristics between villages in which tubewells were feasible and villages in which only AIRPs were feasible in Gopalganj.

When we report effects for villages in which a specific technology was feasible, we use a matched control group, because we do not observe which technologies are feasible in the control villages. Using a matched control group removes these differences (see Appendix Table 4.10). However, in robustness checks, we show that the results for tubewell villages are not sensitive to using a different matched control group, or to simply using the full set of control villages. The construction of the matched control group exploits spatial correlation in the location of villages in which AIRPS were the only feasible technology. Details of the construction of the matched control group are given in Appendix A.

Average treatment effect

In Table 4.7 we show reported access to safe drinking water at baseline (Panel A) and follow-up (Panel B), and the resultant change in access (Panel C). We show results for all villages in columns 1) to 3); in all villages in which tubewells were feasible in columns 4) to 6); and in those villages where only AIRPs were feasible in column 7).

Columns 1) and 4) show the OLS results, which as previously discussed show that treated villages have worse access to safe drinking water at baseline. Columns 2) and 5) show the results dropping South Matlab, where the randomized assignment to treatment was not correctly implemented. Columns 3) and 6) show IV results, using a synthetic assignment to treatment variable in South Matlab. EIther of these approaches removes the baseline differences in reported access to safe drinking water. In Gopalganj, there were no problems with random assignment to treatment, so we only report one set of results.

To estimate the results in Panels A) and B), we use data from all panel households and estimate the following equation:

$$ {}_{i,t} = \alpha + \beta I_{treated,v} + \epsilon_{i,t} \tag{4.4} $$

where i is a household in village v and $Y_{i,t}$ is the access to safe drinking water at baseline or followup.

Panel A) shows baseline access to safe drinking water. This panel repeat the comparisons shown in Table 4.1, and demonstrates the validity of the approaches used for compensating for the failure of random assignment. Panel B) shows the follow-up comparisons. As a result of the initial differences between treatment and control villages, the differences between follow-up and control villages are not statistically significant at follow-up in the OLS analysis, but are clearly evident when we correct for the failure of random assignment.

To estimate the change in access to safe drinking water between baseline and followup (Panel C), we estimate a first difference equation for all households for which we have both baseline and follow-up data, as follows:

$$ \Delta Y_i = Y_{if} - Y_{ib} = \alpha + \beta I_{treated,v} + \epsilon_i \tag{4.5} $$

where ΔY_i is the change in access to safe drinking water between baseline and followup. With two time periods, the first difference analysis is directly equivalent to including household fixed effects. As before, we cluster standard errors at the village level to account for within-village correlation in outcomes.

Panel C) shows the change in reported access to safe drinking water. The OLS and IV results are almost identical, suggesting that although assignment to treatment was not random in all areas, it was not correlated with trends in reported access to safe drinking water. The estimated average treatment effect is 16% overall, and 18% in villages in which tubewells were feasible. There is no treatment effect in the AIRP villages.

For these results and the remainder of the results in this section, we use survey weights which ensure that each village counts equally in the analysis. Where part of the data for a village was lost through the baseline data entry problems, the baseline weights compensate for these losses, as there is no reason to think that the lost data introduces any bias to the estimates of a variable in the village. We do not introduce compensatory weights for migration, but attrition rates were low overall, so this is unlikely to influence the results.

Treatment effect by decision-making model

Table 4.8 repeats the analyses shown in Table 4.7, with the estimated effects broken down by the decision-making model under which we implemented the project. Once again, Panel A) shows baseline access to safe drinking water, Panel B) shows access at follow-up, and Panel C) shows the change in access. Columns 1) to 3) show the results in all villages; columns 4) to 6) show the results in villages where tubewells were feasible; and column 7) shows the results where only AIRPs were feasible. In columns 2) and 5), we correct for failure of random assignment to treatment by dropping South Matlab; and in columns 3) and 6) we correct by instrumenting for the model assigned by the interaction between an indicator for the implemented model, and the synthetic treatment variable.

In Panel A) and B), we use equations with the following structure:

$$Y_{i,t} = \alpha + \beta_{NGO} I_{NGO,v} + \beta_{CP} I_{CP,v} + \beta_{TD} I_{TD,v} + \epsilon_{i,t} \tag{4.6}$$

where $Y_{i,t}$ is reported access to safe drinking water at baseline or follow-up in household i in village v, and I is an indicator whether the village was treated under a given decision-making structure.

Panel A) shows that villages treated under each decision-making model differ from the control villages (as a result of the failure of random assignment to treatment), but that the villages treated under each model are comparable to each other. Columns 2), 3), 5) and 6) also confirm that the strategies for correcting for the failure for random assignment to treatment are also effective for removing significant baseline differences between villages treated under a given model and the control villages as a whole. This is consistent with the results shown in Table 4.4, in which we showed that conditional on treatment, villages assigned to different decision-making models were comparable on baseline statistics. However, note that the magnitude of differences between treatment groups is quite large with respect to the treatment effects we estimate. In particular, 36% of households in NGO Facilitated Community Participation villages report having access to safe drinking water at baseline in comparison to 41% in Community Participation villages and 44% in Top Down villages.

Panel B) shows that, in the specifications when we correct for the failure of random assignment to treatment, there are significant treatment effects under all decision-making models. The estimated effects using only follow-up data are largest for the NGO Facilitated Community Participation, but the differences across models are generally small. However, note that the pattern of access to safe drinking water is reversed: villages in which the project was implemented under NGO Facilitated Community Participation model have the lowest access to safe drinking water at baseline, and the highest access to safe drinking water at follow-up.

In Panel C), we show the results for the change in reported access to safe drinking water. We find that the reported increase in safe drinking water is greatest in NGO model villages, and the reported increase in safe drinking water is almost exactly equivalent in TD and CP model villages. We estimate a first difference regression of the change in reported access to safe drinking water using the following equation:

$$\Delta Y_i = Y_{i,f} - Y_{i,b} = \alpha + \beta_{NGO} I_{NGO,v} + \beta_{CP} I_{CP,v} + \beta_{TD} I_{TD,v} + \epsilon_i \tag{4.7}$$

The estimated coefficients are extremely consistent, with the estimated increase in access to safe drinking water between 21% and 22% in NGO-Facilitated Community Participation model villages (24% and 25% in tubewell villages); between 13% and 14% in Community Participation model villages (between 13% and 15% in tubewell villages) and between 11% and 13% in Top Down model villages (between 12% and 15% in tubewell villages). None of the models shows any significant treatment effect in AIRP villages.

The size of the effect is economically quite important, as the treatment effect almost doubles. However, the differences between models are just below the threshold of statistical significance when all villages are considered. The difference in size of the treatment effect between the NGO-Facilitated Community Participation model villages and the other treated villages is statistically significant when we exclude the villages in which only AIRPs were feasible. Since the treatment effect was zero in these villages, including these villages reduces all the estimated model-specific treatment effects, making it more difficult to distinguish between them, and introduces noise that is not informative with respect to a comparison between the decision-making models.

We do not include data from the additional households that we surveyed at follow-up, because of inconsistencies between the additional households and the panel households, and because of concerns that the sampling protocol may not have been implemented consistently. Including data from the additional households surveyed at followup increases the magnitude and the statistical significance of the difference between the NGO model villages and the other treated villages.

4.3.3.1 Robustness Checks

In Table 4.9, we show the effects of changing the main specification on the estimates of the size of the treatment effects under the different decision-making models. Unless specified, we focus only on the tubewell villages, where the differences between models are statistically significant. For brevity, we only report the p-values of the tests of interest, given the main results: whether the results from the NGO-Facilitated Community Participation model villages are equivalent to the results from the other villages.

In Column 1), we show that the results are similar when we consider only the treated villages, and do not include the control villages in the analysis. The results are given as OLS as within the treated villages, assignment to model was random.

In columns 2) and 3) we show the results by upazila. The treatment effects are larger in Gopalganj, and the differences between models are more pronounced. However, the pattern of results is also consistent for Matlab in that the increase in access to safe drinking water is largest in the villages treated under the NGO-Facilitated Community Participation model.

In columns 4) and 5), we show that the results are not sensitive to how the matched control group is constructed. Column 4) shows that an alternative matched control group, generated by assigning villages in Gopalganj to AIRP and tubewell-matched groups at random, given the probability of AIRP/tubewell feasibility in their neighbourhood, yields almost identical results. Column 5) shows that using the full control group also yields similar estimates.

In columns 6) to 8), we include additional controls. We include the interactions between the control variables and the treatment variables, since otherwise including control variables results in bias in the estimated treatment effect (Freedman, 2008)[11]. In column 6), we use the full set of treated villages and control for the best available technology. We do not include the control villages as we do not directly observe the best available technology in

[11] We include interactions with the raw treatment indicator, rather than the synthetic treatment indicator used for the first stage, as we are not clear which is the correct procedure in the context of an instrumental variables regression. However, the change in the results is minor if we instead control for interactions with the synthetic treatment variables.

these villages. The coefficients are very similar.

In column 7), we include a quadratic function of village size, and its interactions with the treatment indicators. Allowing for heterogeneity in the treatment effect by village size increases the size of the estimated effect in NGO-Facilitated Community Participation. This reflects substantial heterogeneity in the overall effect, and the effect by model, across village size. Figure 4.1 shows the heterogeneity in the overall treatment effect. The effect size decreases with village size, reflecting the fact that the intervention was limited to install a maximum of three safe water sources, regardless of village size, and possibly, the increasing difficulty of solving a collective action problem with a group of increasing size, as theory predicts (Olson, 1971). No treatment effect is detectable in villages with more than 500 households, although the number of villages in this group is small (8 treated villages and 10 control villages). Figure 4.2 shows the heterogeneity in the treatment size by decision-making model:, the effect sizes decreases for each decision-making model, with the effect size greatest under the NGO Facilitated Community-Participation model over the entire range of sizes at which treatment effects are observed. We use the full set of treated and control villages in this column, and column 8.

In column 8), we include a quadratic function in total household assets and its interactions with the treatment indicator. The results are similar to the main specification.

In columns 9) and 10) we report results using alternative measures of access to safe drinking water, described in detail in Appendix 4.6. In column 9, we report results using a measure of whether the source the household was using could be verified to be safe. To verify safety of a tubewell, enumerators inspected the tubewell that the household reported using if it was less than 5 minutes walk away, and recorded whether it was marked red (unsafe), green (safe) or unmarked. The increased treatment effects are larger across all models, partly because our intervention also increased the fraction of tubewells that were verifiably safe, and also because we did not collect this data from all villages at baseline. In column 10), we report results using a measure than uses the verified data when it is available, and the reported data when enumerators were not able to verify the safety of a source. The results are broadly consistent with the results which use only the reported measures.

4.4 Conclusions

This study has provided the first experimental evidence to support the claim that delegating decision-making authorities to communities in projects to provide local public goods can improve outcomes and increase reported impact. In villages where we implemented a project to provide safe sources of drinking water under a participatory decision-making structure (the NGO Facilitated Community Participation model), we installed a slightly larger number of safe drinking water sources (0.2 more sources) but obtained a 9% higher increase in access to safe drinking water, than under a non-participatory decision-making structure (the Top-Down model). Under this model, community members took key decisions by unanimous consensus at a community meeting with minimum representation requirements. These results are broadly consistent with evidence accumulated in the past through practitioner's experience and cross-sectional analysis, but this is the first time that experimental evidence has been available to test the hypothesis that participation in decision-making has a positive impact on the result of social programs.

However, the study also suggests that these benefits may not be realised if protective measures are not put in place to prevent the decision-making process from being co-opted by influential groups or individuals. Under the 'pure' Community Participation model, under which communities took decisions without imposed rules, we installed the same number of sources as under the NGO-Facilitated Community Participation model, but we obtained a 9% smaller increase in access to safe drinking water.

Since we did not test alternative strategies for preventing the decision-making process from co-option, we cannot comment as to whether the method used here (imposing the requirement that decisions be taken by unanimous consensus at a community meeting where all groups were represented and conducting small group meetings beforehand to raise awareness about the project objectives and the rights of all individuals to participate) was the most effective possible in the context. We also did not delegate technical decision-making authority to the community (our project staff determined the feasibility of any given technology and location) and therefore cannot determine whether the results would be the same or different if decision-making authority is delegated to the community over other types of decisions.

A potential weakness of our results is that we rely on reported data, and it is possible that participation in project decision-making may influence the way in which intended beneficiaries report project outcomes. We have also collected data on actual use of the installed water sources by monitoring their use directly using enumerator observations. This data is currently being analysed.

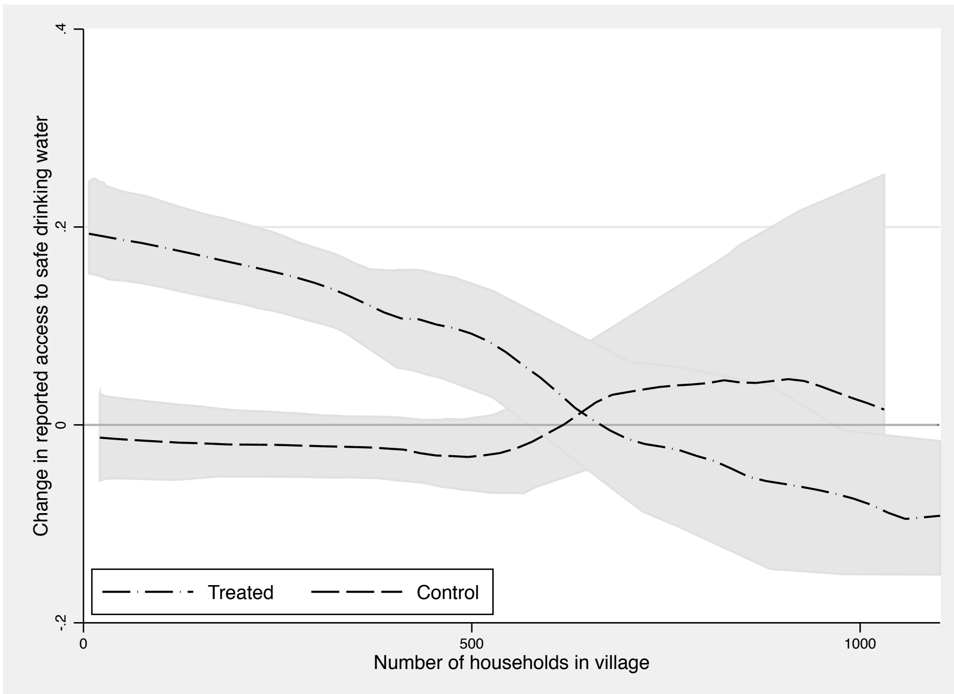
The role of the community contribution appears key in determining outcomes. The number of contributors is low over all. Those that can contribute towards the cost of the water source may have significant influence over the decision-making process. The number of contributors is lowest in the pure Community Participation villages, where we find suggestive evidence of a higher degree of elite capture. Anecdotally, project staff reported to us that in some Top-Down model villages where community groups failed to raise the community contribution, individuals volunteered to pay the community contribution, but only if the water source was installed on their private land.

The result of delegating decision-making authority to the community may vary substantially depending on the local context, for example depending on existing inequalities within the community or on the size and homogeneity of the group to which authority is delegated. We cannot determine whether the results of this study would be applicable in other contexts. The study would benefit from replication in different social and cultural contexts.

Bearing these caveats in mind, the results provide important experimental evidence regarding an influential policy recommendation, and suggest that careful consideration should be given to the structure of a participatory decision-making process, if the potential benefits are to be realized.

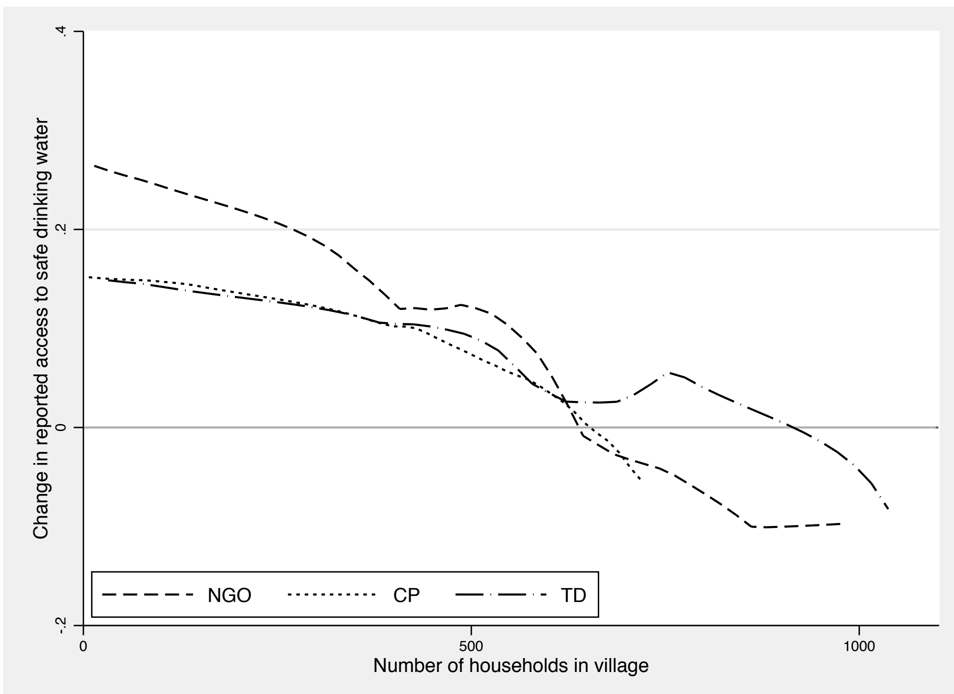
4.5 Tables and Figures

Figure 4.1: Heterogeneity in average treatment effect with village size



Graph shows results from a local linear regression of the change in access to safe drinking water on village size for treated and control villages. 90% confidence intervals are cluster bootstapped at the village level.

Figure 4.2: Heterogeneity in treatment effect by decision-making model with village size



Graph shows results from a local linear regression of the change in access to safe drinking water on village size for treated and control villages.

Table 4.1: Treated vs Control
Baseline Summary Statistics and Randomization Checks

		Control	Treated	Treatment - Control		
		(1)	(2)	(3)	(4)	(5)
Proportion of villages in Gopalganj	Mean	0.51	0.55			
	s.e.	(0.05)	(0.04)			
Proportion of villages in South Matlab	Mean	0.00	0.23***			
	s.e.	-	(0.04)			
No of households in village	Mean	245	221	-26	-29	-34
	s.e.	(21)	(17)	(29)	(27)	(31)
% of water sources arsenic contaminated	Mean	0.96	0.95	0.01	0.00	0.00
	s.e.	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Reports using arsenic safe water	Mean	0.55	0.41***	-0.01	-0.03	-0.03
	s.e.	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)
Changed source of drinking water due to arsenic in last 5 years?	Mean	0.49	0.35***	0.00	-0.02	-0.03
	s.e.	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)
Anyone in household has symptoms of arsenic poisoning?	Mean	0.009	0.009	-0.001	0.000	0.000
	s.e.	(0.002)	(0.001)	(0.003)	(0.002)	(0.003)
Total value of household assets	Mean	570284	540165	-14573	-5071	-5865
	s.e.	(30362)	(21720)	(41233)	(37175)	(42881)
Access to electricity?	Mean	0.46	0.39	-0.05	-0.02	-0.03
	s.e.	(0.03)	(0.03)	(0.05)	(0.04)	(0.05)
Household head literate	Mean	0.61	0.60	0.01	0.00	0.00
	s.e.	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Household head Muslim	Mean	0.70	0.70	0.04	0.05	0.05
	s.e.	(0.04)	(0.04)	(0.05)	(0.05)	(0.06)
Household head farmer	Mean	0.42	0.45	0.03	0.02	0.03
	s.e.	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
Number of associations in community	Mean	6.25	6.30	-0.18	-0.18	-0.20
	s.e.	(0.14)	(0.15)	(0.22)	(0.19)	(0.22)
Number of collective actions in community	Mean	0.89	0.96	0.05	0.07	0.08
	s.e.	(0.08)	(0.09)	(0.05)	(0.06)	(0.06)
	Number of villages	99	127	197	226	226
	Number of households	3914	4976	7755	8890	8890

Note: P-values test significance of differences between treated and control villages, controlling for the different treatment proportions in Gopalganj, North and South Matlab in columns 3-5). Data in rows 1) and 2) come from project records. Data in rows 3) and 4) comes from data from the Bangladesh Arsenic Mitigation Water Supply Project. All other data is from baseline household surveys. Two villages are missing all baseline data. Standard errors (in parentheses) are robust, and clustered at the village level in rows 5) onwards.
*** p<0.01, ** p<0.05, * p<0.1.

Table 4.2: Decision-making structures

Non-participatory	Top Down (TD)	Project staff took all project decisions, after an extended (typically 2-day) period of information gathering, using the following criteria to decide water source location: • public/convenient location • population density • existing safe water options
Participatory	Community Participation (CP)	The community took all project decisions using their own (unobserved) decision-making structures, following a community-wide information meeting led by project staff.
	NGO-Facilitated Community Participation (CP)	The community took all project decisions at a community- wide meeting, following smaller information meetings for different groups. We imposed two decision-making rules. If decisions made did not satisfy these rules, project staff did not implement the decisions: • Attendance at the community meeting had to include: at least 10 men, of which 5 had to qualify as poor; and at least 10 women, of which 5 had to qualify as poor. • Decisions had to be unanimous.

Table 4.3: Technologies to provide arsenic-safe drinking water

		Required community contribution per safe water source installed		
Technology	Cost	1	2	3
Deep tubewell (DTW)	50000	4500	6000	7500
Shallow tubewell (STW)	20000	3000	3500	4000
Arsenic-Iron Removal Plant (AIRP)	60000	6000	7500	N/A
Deep-set tubewell (DSTW)	60000	6000	7500	N/A

Note: All prices in Bangladeshi Taka. 1 US$≈ 80BDT.

Table 4.4: Assignment to decision-making structure
Baseline Summary Statistics and Randomization Checks

		NGO (1)	CP (2)	TD (3)
Proportion of villages in Gopalganj	Mean	0.55	0.55	0.56
	s.e.	(0.08)	(0.08)	(0.08)
Proportion of villages in South Matlab	Mean	0.24	0.21	0.23
	s.e.	(0.07)	(0.06)	(0.06)
No of households in village	Mean	213	213	236
	s.e.	(33)	(24)	(32)
% of water sources arsenic contaminated	Mean	0.96	0.95	0.95
	s.e.	(0.01)	(0.01)	(0.01)
AIRPs only feasible technology	Mean	0.10	0.14	0.14
	s.e.	(0.05)	(0.05)	(0.05)
Reports using arsenic safe water	Mean	0.36	0.42	0.44
	s.e.	(0.05)	(0.05)	(0.05)
Changed source of drinking water due to arsenic in last 5 years?	Mean	0.32	0.35	0.37
	s.e.	(0.05)	(0.05)	(0.05)
Anyone in household has symptoms of arsenic poisoning?	Mean	0.004**	0.009	0.012*
	s.e.	(0.002)	(0.003)	(0.003)
Total value of household assets	Mean	543364	548174	529180
	s.e.	(39597)	(42208)	(30413)
Access to electricity?	Mean	0.37	0.39	0.42
	s.e.	(0.05)	(0.05)	(0.05)
Household head literate	Mean	0.60	0.58	0.62
	s.e.	(0.03)	(0.03)	(0.02)
Household head Muslim	Mean	0.68	0.70	0.73
	s.e.	(0.07)	(0.06)	(0.06)
Household head farmer	Mean	0.44	0.46	0.44
	s.e.	(0.03)	(0.02)	(0.02)
Number of associations in community	Mean	6.36	6.05	6.47
	s.e.	(0.25)	(0.18)	(0.31)
Number of collective actions in community	Mean	0.91	1.00	0.98
	s.e.	(0.14)	(0.16)	(0.15)
	Number of villages	42	42	43
	Number of households	1638	1635	1703

Note: P-values test significance of the difference between model and other treated villages. Data from household surveys except rows 1), 2) and 5) which come from project records and rows 3) and 4) which come from the Bangladesh Arsenic Mitigation Water Supply Project. Baseline data for one CP village is missing. Standard errors (in parentheses) are robust, and clustered at the village level in rows 5) onwards. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4.5: Participation in project decision-making

		No. of participants	Fraction of meeting participants: Female	Low s.e. status	No 2° education
		(1)	(2)	(3)	(4)
All treated	Mean	30.7	0.27	0.42	0.60
	s.e.	(1.1)	(0.02)	(0.02)	(0.02)
NGO	Mean	33.6	0.31	0.44	0.64
	s.e.	(2.5)	(0.03)	(0.04)	(0.03)
CP	Mean	29.9	0.29	0.46	0.55
	s.e.	(1.4)	(0.03)	(0.02)	(0.03)
TD	Mean	28.5	0.21	0.36	0.61
	s.e.	(1.8)	(0.03)	(0.03)	(0.03)
NGO = CP	p-value	0.200	0.615	0.570	0.031
CP = TD	p-value	0.561	0.073	0.006	0.154
TD = NGO	p-value	0.111	0.023	0.085	0.490
	N	126	123	119	122

Note: P-values test pairwise significance of the difference between the means across models indicated, from a regression of the outcome variable on indicators for the three types of treatment (with no constant). Robust standard errors shown in parentheses. In two villages, meetings were not held.

Table 4.6: Project Outcomes

		Outcome Variable		
		Water sources installed	Installed in public places	Number of contributors
		(1)	(2)	(3)
Panel A: All villages				
All treated	Mean	2.14	1.38	7.81
	s.e.	(0.10)	(0.10)	(0.83)
NGO	Mean	2.21	1.19	9.37
	s.e.	(0.16)	(0.14)	(1.62)
CP	Mean	2.21	1.00	5.40
	s.e.	(0.17)	(0.14)	(0.93)
TD	Mean	2.00	1.93	8.67
	s.e.	(0.18)	(0.18)	(1.63)
NGO = CP	p-value	1.000	0.334	0.036
CP = TD	p-value	0.394	0.000	0.084
TD = NGO	p-value	0.376	0.002	0.764
	N	127	127	126
Panel B: Villages where tubewells feasible				
All treated	Mean	2.45	1.58	9.02
	s.e.	(0.08)	(0.10)	(0.92)
NGO	Mean	2.46	1.30	10.61
	s.e.	(0.13)	(0.14)	(1.74)
CP	Mean	2.53	1.14	6.11
	s.e.	(0.14)	(0.15)	(1.03)
TD	Mean	2.36	2.31	10.33
	s.e.	(0.16)	(0.15)	(1.82)
NGO = CP	p-value	0.718	0.440	0.028
CP = TD	p-value	0.428	0.000	0.046
TD = NGO	p-value	0.624	0.000	0.912
	N	109	109	108

Note: P-values test pairwise significance of the difference between the means across models indicated, from a regression of the outcome variable on indicators for the three types of treatment (with no constant). Robust tandard errors shown in parentheses. In villages where no water sources were installed, the number of contributors is coded as zero. In one village, the number of contributors was not recorded.

Table 4.7: Estimates of average treatment effect

		OLS (1)	OLS (2)	IV (3)	OLS (4)	OLS (5)	IV (6)	OLS (7)
Panel A: Reported access to safe drinking water at baseline								
Treated	Coefficient	-0.141***	-0.011	-0.031	-0.181***	-0.017	-0.042	-0.024
	s.e.	(0.05)	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.10)
Control	Coefficient	0.547	0.830	0.837	0.609	0.833	0.842	0.253
	s.e.	(0.04)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.07)
	First-stage F-test			1340			896	
	N	8695	7608	8695	7375	6288	7375	1242
Panel B: Reported access to safe drinking water at follow-up								
Treated	Coefficient	0.022	0.141***	0.132***	0.001	0.148***	0.137***	-0.007
	s.e.	(0.05)	(0.03)	(0.04)	(0.05)	(0.04)	(0.04)	(0.08)
Control	Coefficient	0.533	0.820	0.823	0.607	0.817	0.821	0.201
	s.e.	(0.04)	(0.02)	(0.03)	(0.04)	(0.03)	(0.03)	(0.06)
	First-stage F-test			1335			896	
	N	8442	7387	8442	7168	6113	7168	1201
Panel C: Change in reported access to safe drinking water								
Treated	Coefficient	0.164***	0.148***	0.157***	0.182***	0.164***	0.175***	0.004
	s.e.	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.09)
Control	Coefficient	-0.014	-0.012	-0.015	-0.003	-0.017	-0.021	-0.042
	s.e.	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)
	First-stage F-test			1337			897	
	N	8427	7375	8427	7154	6102	7154	1200
	Feasible technology	All	All	All	Tubewell	Tubewell	Tubewell	AIRP
	Controls for upazila	No	Yes	Yes	No	Yes	Yes	-
	Includes S. Matlab	Yes	No	Yes	Yes	No	Yes	-
	Control villages	All	All	All	Matched	Matched	Matched	Matched

Note: Treatment is instrumented using synthetic assignment to treatment in South Matlab in columns 3) and 6). In columns 4) to 7) the control group is matched to the subset of treated villages using baseline propensity score matching. Survey weights are applied so that each village counts equally in the analysis. Controls for upazila include an indicator for Gopalganj, and an indicator for South Matlab, where included. Standard errors (in parentheses) are robust and clustered at the village level.
*** p<0.01, ** p<0.05, * p<0.1.

Table 4.8: Estimates of treatment effect by decision-making model

		OLS (1)	OLS (2)	IV (3)	OLS (4)	OLS (5)	IV (6)	OLS (7)
Panel A: Reported access to safe drinking water at baseline								
NGO	Coefficient	-0.19***	-0.06	-0.07	-0.24***	-0.08	-0.09	-0.10
	s.e.	(0.06)	(0.05)	(0.05)	(0.07)	(0.06)	(0.06)	(0.09)
CP	Coefficient	-0.13**	0.01	-0.02	-0.17**	0.00	-0.04	0.04
	s.e.	(0.07)	(0.06)	(0.05)	(0.07)	(0.06)	(0.06)	(0.17)
TD	Coefficient	-0.10	0.02	0.00	-0.13*	0.03	0.01	-0.04
	s.e.	(0.06)	(0.05)	(0.05)	(0.07)	(0.05)	(0.06)	(0.10)
Constant	Coefficient	0.55	0.83	0.84	0.61	0.83	0.84	0.25
	s.e.	(0.04)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.07)
	NGO = CP	0.437	0.295	0.460	0.395	0.257	0.443	0.389
	CP = TD	0.708	0.916	0.673	0.563	0.683	0.456	0.609
	TD = NGO	0.239	0.199	0.208	0.151	0.106	0.119	0.540
	NGO = pooled	0.258	0.180	0.252	0.185	0.114	0.180	0.327
	CP = pooled	0.819	0.587	0.851	0.878	0.673	0.988	0.480
	TD = pooled	0.372	0.423	0.329	0.246	0.241	0.183	0.873
	N	8695	7608	8695	7375	6288	7375	1242
	Feasible technology	All	All	All	Tubewell	Tubewell	Tubewell	AIRP
	Controls for upazila	No	Yes	Yes	No	Yes	Yes	-
	Includes S. Matlab	Yes	No	Yes	Yes	No	Yes	-
	Control villages	All	All	All	Matched	Matched	Matched	Matched

Note: Model is instrumented using model interacted with synthetic assignment to treatment in South Matlab in columns 3) and 6). In columns 4) to 7) the control group is matched to the subset of treated villages using baseline propensity score matching. Survey weights are applied so that each village counts equally in the analysis. Controls for upazila include an indicator for Gopalganj, and an indicator for South Matlab, where included. Standard errors (in parentheses) are robust and clustered at the village level.
*** p<0.01, ** p<0.05, * p<0.1.

Table 4.8, continued: Estimates of treatment effect by decision-making model

		OLS (1)	OLS (2)	IV (3)	OLS (4)	OLS (5)	IV (6)	OLS (7)
Panel B: Reported access to safe drinking water at follow-up								
NGO	Coefficient	0.03	0.15***	0.16***	0.01	0.16***	0.16***	-0.13*
	s.e.	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.07)
CP	Coefficient	0.00	0.14***	0.11**	-0.03	0.14**	0.10*	0.09
	s.e.	(0.06)	(0.05)	(0.05)	(0.07)	(0.05)	(0.05)	(0.13)
TD	Coefficient	0.03	0.13***	0.13***	0.02	0.15***	0.15***	-0.02
	s.e.	(0.06)	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.09)
Constant	Coefficient	0.53	0.82	0.82	0.61	0.82	0.82	0.20
	s.e.	(0.04)	(0.02)	(0.03)	(0.04)	(0.03)	(0.03)	(0.06)
	NGO = CP	0.702	0.840	0.430	0.657	0.725	0.308	0.091
	CP = TD	0.697	0.860	0.634	0.499	0.875	0.372	0.473
	TD = NGO	0.998	0.708	0.699	0.809	0.823	0.811	0.178
	NGO = pooled	0.823	0.748	0.504	0.893	0.745	0.467	0.042
	CP = pooled	0.660	0.978	0.461	0.528	0.764	0.272	0.217
	TD = pooled	0.817	0.740	0.968	0.584	0.963	0.704	0.958
	N	8442	7387	8442	7168	6113	7168	1201
	Feasible technology	All	All	All	Tubewell	Tubewell	Tubewell	AIRP
	Controls for upazila	No	Yes	Yes	No	Yes	Yes	-
	Includes S. Matlab	Yes	No	Yes	Yes	No	Yes	-
	Control villages	All	All	All	Matched	Matched	Matched	Matched

Note: Model is instrumented using model interacted with synthetic assignment to treatment in South Matlab in columns 3) and 6). In columns 4) to 7) the control group is matched to the subset of treated villages using baseline propensity score matching. Survey weights are applied so that each village counts equally in the analysis. Controls for upazila include an indicator for Gopalganj, and an indicator for South Matlab, where included. Standard errors (in parentheses) are robust and clustered at the village level.
*** p<0.01, ** p<0.05, * p<0.1.

Table 4.8, continued: Estimates of treatment effect by decision-making model

		OLS (1)	OLS (2)	IV (3)	OLS (4)	OLS (5)	IV (6)	OLS (7)
Panel C: Change in reported access to safe drinking water								
NGO	Coefficient	0.22***	0.21***	0.22***	0.25***	0.24***	0.25***	-0.05
	s.e.	(0.05)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)	(0.09)
CP	Coefficient	0.14***	0.13***	0.13***	0.15***	0.13***	0.14***	0.03
	s.e.	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.15)
TD	Coefficient	0.13***	0.11**	0.13***	0.15***	0.12**	0.14***	0.02
	s.e.	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.12)
Constant	Coefficient	-0.01	-0.01	-0.02	0.00	-0.02	-0.02	-0.04
	s.e.	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)
	NGO = CP	0.178	0.238	0.149	0.116	0.143	0.083	0.653
	CP = TD	0.946	0.783	0.969	0.954	0.765	0.930	0.947
	TD = NGO	0.151	0.160	0.155	0.109	0.104	0.106	0.636
	NGO = pooled	0.119	0.150	0.110	0.078	0.088	0.064	0.568
	CP = pooled	0.429	0.556	0.358	0.338	0.432	0.250	0.786
	TD = pooled	0.356	0.302	0.388	0.298	0.231	0.329	0.850
	N	8427	7375	8427	7154	6102	7154	1200
	Feasible technology	All	All	All	Tubewell	Tubewell	Tubewell	AIRP
	Controls for upazila	No	Yes	Yes	No	Yes	Yes	-
	Includes S. Matlab	Yes	No	Yes	Yes	No	Yes	-
	Control villages	All	All	All	Matched	Matched	Matched	Matched

Note: Model is instrumented using model interacted with synthetic assignment to treatment in South Matlab in columns 3) and 6). In columns 4) to 7) the control group is matched to the subset of treated villages using baseline propensity score matching. Survey weights are applied so that each village counts equally in the analysis. Controls for upazila include an indicator for Gopalganj, and an indicator for South Matlab, where included. Standard errors (in parentheses) are robust and clustered at the village level.
*** p<0.01, ** p<0.05, * p<0.1.

Table 4.9: Robustness checks

		OLS (1)	OLS (2)	IV (3)	IV (4)	IV (5)	OLS (6)	IV (7)	IV (8)	IV (9)	IV (10)
NGO	Coefficient	0.24***	0.37***	0.10	0.25***	0.27***	0.24***	0.46***	0.28***	0.28***	0.30***
	s.e.	(0.05)	(0.09)	(0.07)	(0.06)	(0.06)	(0.06)	(0.13)	(0.07)	(0.07)	(0.07)
CP	Coefficient	0.15***	0.19***	0.08	0.14***	0.15***	0.14***	0.18	0.15***	0.16**	0.13***
	s.e.	(0.04)	(0.07)	(0.07)	(0.05)	(0.05)	(0.04)	(0.16)	(0.06)	(0.07)	(0.05)
TD	Coefficient	0.14***	0.21***	0.06	0.14***	0.16***	0.12**	0.21*	0.16***	0.20***	0.14***
	s.e.	(0.04)	(0.07)	(0.06)	(0.05)	(0.05)	(0.05)	(0.12)	(0.06)	(0.07)	(0.05)
Constant	Coefficient		-0.04	0.01	-0.02	-0.03		-0.03	-0.03	0.20	0.02
	s.e.		(0.04)	(0.02)	(0.02)	(0.02)		(0.06)	(0.02)	(0.04)	(0.03)
	NGO = CP	0.118	0.055	0.784	0.083	0.084	0.165	0.133	0.098	0.117	0.020
	NGO = TD	0.111	0.107	0.637	0.106	0.108	0.122	0.117	0.151	0.280	0.027
	NGO = pooled	0.080	0.052	0.674	0.064	0.065	0.106	0.077	0.086	0.130	0.013
	N	4017	3201	3953	7342	7784	4660	8347	8357	7154	7154
	Measure of access	Reported	Reported	Reported	Reported	Reported	Reported	Reported	Reported	Verified	Combined
	Feasible technology	Tubewell	Tubewell	Tubewell	Tubewell	Tubewell	All	All	All	Tubewell	Tubewell
	Sample	All	Gopalganj	Matlab	All	All	All	All	All	All	All
	Upazila Controls	No	-	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
	Additional controls	No	No	No	No	No	Technology	Vill. size	Assets	No	No
	Control villages	None	Matched	All	Alternative	All	None	All	All	Matched	Matched

Note: Model is instrumented using model interacted with synthetic assignment to treatment in South Matlab where indicated. Survey weights are applied so that each village counts equally in the analysis. Controls for upazila include an indicator for Gopalganj, and an indicator for South Matlab, where included. Where additional controls are included, the interactions between the controls and the treatment indicators are also

Standard errors (in parentheses) are robust and clustered at the village level.

*** p<0.01, ** p<0.05, * p<0.1.

4.6 Appendices to Chapter 4

Appendix A: Data

Variable Construction

We asked the households to list all the sources of water they used for drinking and cooking. In the analysis, we focus on the most important source of water for drinking and cooking, which we asked households to list first. We also asked households to report the percentage of water for drinking and cooking that they obtained from each source, but results based on the source from which households report drawing the largest percentage of water are unstable between baseline and followup, whereas the results based on the first-listed water source are more consistent. This may be attributable to slight differences in the way in which the question was asked as to whether the question referred to water used for drinking only or drinking and cooking.

Reports using safe drinking water If the household reports using a tubewell, we code the household as reporting using safe water if they report that the source is arsenic-safe, and reporting unsafe water if it is unsafe or if they dont know the source's safety. If the household reports using an unsafe source with respect to bacterial contamination (i.e. a dug well or surface water), we code the household as reporting using unsafe water. Some sources can be presumed to be safe from both bacterial and arsenic contamination (e.g. AIRPs, PSF, rainwater, deep-set tubewells). In these cases, we code the household as reporting using safe water unless the household reports that the water is unsafe. The numbers of households using these sources is small. If the household reports using any other source, we code the household as reporting using safe water if they report that the source is safe, and reporting unsafe water if it is unsafe or if they dont know the source's safety status.

Reports using verified safe drinking water Many tubewells in Bangladesh have been tested for arsenic safety in the past and marked with green (safe) or red (unsafe). If households reported using a tubewell, enumerators visited the tubewell to confirm whether it was marked safe, unsafe, or not marked, as long as the tubewell was less than 5 minutes walk away. However, an early version of the baseline survey used did not include this question, so this information is missing for some villages at baseline. We code this variable as follows. We always code households reporting an unsafe source as not using a verified source of safe drinking water. If the household reports using a source that can otherwise be presumed to be safe (i.e. rainwater), we code the household as using a verified source of safe drinking water, unless they report it to be unsafe. If they report using a tubewell, we code the house as using a verified source of safe drinking water if the source is verifiably safe, and not using a verified source of safe drinking water if the tubewell is either marked unsafe or can't be verified safe (either because it is too far, or because it is unmarked, or because the question wasn't asked in this village at baseline). If the household reports using any other source, we code the household as reporting using a verified source safe water if the enumerators recorded that the source could be verified safe, and reporting using an unverified or unsafe source if the source is unsafe or if it cannot be verified.

Combined measure We combine these measures by using the verified data, when the safety of the source could be verified, and the reported data, when the safety of the source could not be verified.

Construction of matched control groups in Gopalganj

Tubewells are not feasible where there is a rocky layer separating the surface from the arsenic-safe deep aquifer, which cannot be penetrated with local drilling technologies. This was only a problem in this study in Gopalgnaj, as tubewells of varying kinds were feasible in all villages in Matlab. The following discussion is therefore limited to Gopalganj. There is substantial spatial correlation in the location of these rocky layers.

In Gopalganj, there was no overall problem with random assignment to treatment. Appendix Table 4.10, columns 1) and 2) confirm this; of 12 tests comparing treated to control villages in Gopalganj, only one shows statistically significant differences at the 10% level, which is approximately what we would expect due to chance. However, when we compare either the AIRP villages (column 3) or the tubewell villages (column 6), to the full group of control villages, there is some evidence that feasible technology is correlated with other village level characteristics. In column 3), only 1 of 12 tests shows statistically significant differences at the 10% level, but in column 6), 1 test shows statistically significant differences at the 5% level, and additionally another shows statistically significant differences at the 10% level. To be conservatve, we construct matched control groups for the AIRP villages and tubewell villages, primarily exploiting the spatial correlation in location of the rocky layer.

In Gopalganj, there are 18 unions (the smallest rural administrative and local government units in Bangladesh). In three of these unions, only AIRPs were feasible in all treated villages. In 4 unions, tubewells were feasible in all treated villages. First, we assign the five control villages in unions where only AIRPs were feasible to the AIRP-matched control group. Second, we assign the 18 control villages in unions where tubewells were always feasible to the tubewell-matched control group. For the remainder of the unions — for which tubewells were feasible in a fraction of the villages — we use two strategies to identify matched control villages.

First, we construct a propensity score index by running a logit regression of an indicator for whether tubewells (or only AIRPs) were feasible on a set of baseline characteristics which we observe for both treated and control groups and which are correlated feasible technology in the treatment group. These baseline characteristics are: union fixed effects, access to electricity, number of tubewells operative, number of tubewells arsenic contaminated, reported access to safe drinking water at baseline, verifiability of arsenic safety of primary reported source, verified access to safe drinking water at baseline, fraction of village changing source of safe drinking water in the preceding five years due to arsenic contamination; whether any member of the household has arsenic poisoning; mean literacy of household heads; and number of coordinated actions in the village. We then assign the remaining villages to the relevant matched control group where their propensity score is greater than 0.5 (although in reality the propensity scores are strongly clustered around 0 and 1). This process assigns a total of 16 villages to the AIRP-matched control group, and 33 villages to the tubewell-matched control group. 3 villages are not assigned to either control group, and 2 villages are assigned to both groups. Columns 4) and 7) show the characteristics of the groups constructed in this way. No tests show statistically significant differences between these groups and the AIRP or tubewell villages from the treated groups.

The second method we use is to assign the villages in each union to the tubewell-matched and AIRP-matched control groups according to the observed, union-level probability of AIRP and tubewell feasibility. We do this by generating a random number for each control village and assigning them to the relevant control group if the random number is less than the proportion of treated villages in that union in which the technology was feasible. This process assigns 9 villages to the AIRP-matched control villages and 40 villages to the tubewell-matched control villages. One village is not assigned to either control group, and by construction none are assigned to both groups. Columns 5) and 8) show the characteristics of the groups constructed in this way. 3 tests show statistically significant differences between the AIRP-matched control villages and the AIRP villages from the treated groups, but no tests show statistically significant differences between the tubewell-matched control villages constructed in this way and the tubewell villages from the treated groups.

For the main specification, we therefore use the matched control group created using propensity score matching, but we will compare these results to results using the full control group and the second matching strategy in robustness checks.

Appendix B: Figures and Tables

Table 4.10: Baseline Summary Statistics for AIRP and non-AIRP villages in Gopalganj

	Treated (1)	Control (2)	AIRP (3)	AIRP Control 1 (4)	AIRP Control 2 (5)	Tubewell (6)	Tubewell Control 1 (7)	Tubewell Control 2 (8)
No of households in village	262	257	338	268	271	242	255	263
	(26)	(33)	(71)	(44)	(57)	(27)	(45)	(41)
% of water sources arsenic contaminated	0.97	0.96	0.97	0.97	0.99*	0.96	0.96	0.96
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Reports using arsenic safe water	0.27	0.26	0.23	0.25	0.26	0.27	0.27	0.26
	(0.03)	(0.04)	(0.07)	(0.07)	(0.11)	(0.04)	(0.04)	(0.04)
Changed source of drinking water due to arsenic in last 5 years?	0.21	0.20	0.22	0.21	0.19*	0.19	0.18	0.19
	(0.03)	(0.03)	(0.06)	(0.07)	(0.10)	(0.03)	(0.04)	(0.04)
Anyone in household has symptoms of arsenic poisoning?	0.006	0.009	0.005	0.008	0.006	0.006	0.003	0.008
	(0.002)	(0.003)	(0.002)	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)
Total value of household assets	469230	512028	568106	531573	527437	436844	493045	514595
	(21450)	(40937)	(47317)	(82727)	(146437)	(23279)	(47148)	(41289)
Access to electricity?	0.37	0.45	0.60*	0.68	0.63	0.31**	0.35	0.41
	(0.04)	(0.04)	(0.07)	(0.03)	(0.05)	(0.04)	(0.05)	(0.05)
Household head literate	0.58	0.56	0.54	0.49	0.45	0.59	0.62	0.58
	(0.03)	(0.04)	(0.06)	(0.07)	(0.08)	(0.03)	(0.04)	(0.04)
Household head Muslim	0.55	0.48	0.66	0.46	0.52	0.52	0.46	0.48
	(0.05)	(0.06)	(0.11)	(0.10)	(0.14)	(0.06)	(0.08)	(0.07)
Household head farmer	0.50	0.46	0.45	0.39	0.39	0.51*	0.49	0.47
	(0.02)	(0.02)	(0.04)	(0.03)	(0.05)	(0.02)	(0.03)	(0.02)
Number of associations in community	6.76	6.86	6.65	7.18	7.62	6.77	6.79	6.82
	(0.23)	(0.22)	(0.36)	(0.50)	(0.70)	(0.29)	(0.23)	(0.22)
Number of collective actions in community	0.23	0.14*	0.26	0.12	0.10*	0.22	0.15	0.14
	(0.04)	(0.02)	(0.08)	(0.04)	(0.04)	(0.05)	(0.03)	(0.03)
Number of villages	70	50	16	16	9	52	33	38
Number of households	3859	1969	1102	627	347	2677	1304	1504

Note: Stars indicate significance of tests of difference of means: between treated and control villages in Gopalganj in column 2); between AIRP and all control villages in column 3); between AIRP and matched control villages in columns 4) and 5); between tubewell and all control villages in column 6; and between tubewell and matched control villages in columns 7) and 8). Data in rows 1) and 2) come from the Bangladesh Arsenic Mitigation Water Supply Project. The remaining data are from household surveys. Standard errors (shown in parentheses) are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1.

Bibliography

Åkerman, A. (2009). *Trade, reallocations and productivity: A bridge between theory and data in Öresund*. IFN Working Paper No. 795.

Alatas, V., Banerjee, A., Hanna, R., Olken, B. A., Purnamasari, R., & Wai-Poi, M. (2013, January). *Does elite capture matter? Local elites and targeted welfare programs in Indonesia.*

Allen, T. & Arkolakis, C. (2013, June). *Trade and topography of the spatial economy.*

Angrist, J. D. & Pischke, J. (2009). *Mostly harmless econometrics*. Princeton University Press.

Armenter, R., Koren, M., & Nagy, D. K. (2013, May). *Bridges*. Working Paper.

Atack, J., Bateman, F., Haines, M., & Margo, R. A. (2010). Did railroads induce or follow economic growth? Urbanization and population growth in the American Midwest, 1850-1860. *Social Science History*, *34*(2).

Banerjee, A., Banerji, R., Duflo, E., Glennerster, R., & Khemani, S. (2010). Pitfalls of participatory programs: Evidence from a randomized evaluation in education in India. *American Economic Journal: Economic Policy*.

Banerjee, A., Duflo, E., & Qian, N. (2012, February). *On the road: Access to transportation infrastructure and economic growth in China*. Working Paper.

Baum-Snow, N. (2007). Did highways cause suburbanization? *Quarterly Journal of Economics*, *122*(2), 775–805.

Beath, A., Christia, F., & Enikolopov, R. (2013, January). *Direct democracy and resource allocation: Experimental evidence from Afghanistan*. Centre for Economic and Financial Research at New Economic School Working Paper No 192.

Berger, T. & Enflo, K. (2013, July). *Locomotives of local growth: The short- and long-term impact of railroads in Sweden*. EHES Working papers in Economic History No. 42.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, *119*(1), 249–275.

Björkman, M. & Svensson, J. (2009). Power to the people: Evidence from a randomized field experiment on community-based monitoring in Uganda. *Quarterly Journal of Economics*, *124*(2).

Black, D. A., Muszynska, M., Sanders, S. G., Taylor, E. J., & Taylor, L. J. (2011, April). *The great migration and African-American mortality: Evidence from Mississippi.*

Blattman, C., Hartman, A. C., & Blair, R. A. (2014). How to promote order and property rights under weak rule of law? an experiment in changing dispute resolution behavior through community education. *American Political Science Review*, *108*(01).

Bleakley, H. & Lin, J. (2012). Portage and path dependence. *Quarterly Journal of Economics*, *127*, 587–644.

Casey, K., Glennerster, R., & Miguel, E. (2012). Reshaping institutions: Evidence on aid impacts using a preanalysis plan. *Quarterly Journal of Economics*, *127*(4).

Chandra, A. & Thompson, E. (2000). Does public infrastructure affect economic activity? Evidence from the rural interstate highway system. *Regional Science and Urban Economics*, *30*, 457–490.

Chattopadhyay, R. & Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in India. *Econometrica*, *72*(5).

Conley, T. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, *92*, 1–45.

Conley, T. (2008). New Palgrave Dictionary of Economics. (2nd Edition, Chap. Spatial Econometrics). Macmillan.

Cordell, G. (2011, March). *Mississippi river bridges*. Retrieved from http://tennesseeencyclopedia.net/entry.php?rec=1564

Coşar, A. K. & Fajgelbaum, P. D. (2012, December). *Internal geography, international trade, and regional outcomes*.

Costello, M. C. A. (1995). *Climbing the Mississippi River bridge by bridge*. Mary C. Costello.

Costello, M. C. A. (2002). *Climbing the Mississippi River bridge by bridge: Minnesota*. Mary C. Costello.

Davis, D. R. & Weinstein, D. E. (2002). Bones, bombs, and break points: The geography of economic activity. *American Economic Review*, *92*(5).

Donaldson, D. (2010, October). *Railroads of the Raj: estimating the impact of transportation infrastructure*. Revise and resubmit, American Economic Review.

Donaldson, D. (2012, September). *Railroads of the Raj: Estimating the impact of transportation infrastructure*.

Donaldson, D. & Hornbeck, R. (2013, July). *Railroads and American economic growth: A "market access" approach*. Working Paper.

Duranton, G., Morrow, P. M., & Turner, M. A. (2014). Roads and trade: Evidence from the US. *Review of Economic Studies, forthcoming*.

Duranton, G. & Turner, M. A. (2011, June). *Urban growth and transportation*. Review of Economic Studies, forthcoming.

Duranton, G. & Turner, M. A. (2012). Urban growth and transportation. *Review of Economic Studies*, *79*(4).

Faber, B. (2013, July). *Trade integration, market size, and industrialization: evidence from china's national trunk highway system*. Working Paper.

Fearon, J. D., Humphreys, M., & Weinstein, J. M. (2009). Can development aid contribute to social cohesion after civil war? Evidence from a field experiment in post-conflict Liberia. *American Economic Review: Papers and Proceedings*, *99*(2).

Fearon, J. D., Humphreys, M., & Weinstein, J. M. (2011, December). *Democratic institutions and collective action capacity: Results from a field experiment in post-conflict Liberia.*

Fishlow, A. (1965). *American railroads and the tranformation of the antebellum economy*. Cambridge (Mass.)

Fogel, R. (1964). *Railroads and American economic growth : Essays in econometric history*. Baltimore: John Hopkins Press.

Freedman, D. A. (2008). On regression adjustments to experimental data. *Applied Mathematics*, *40*, 180–193.

Fritzen, S. (2007). Can the design of community-driven development reduce the risk of elite capture? Evidence from Indonesia. *World Development*, *35*(8), 1359–1375.

Fujita, M., Krugman, P., & Venables, A. J. (2001). *The spatial economy: Cities, regions and international trade*. MIT Press.

Fung, A. & Wright, E. O. (2003). *Deepening democracy: Institutional innovations in empowered participatory governance*. Verso.

Glaeser, E. L., Kahn, M. E., & Rappaport, J. (2008). Why do the poor live in cities? The role of public transportation. *Journal of Urban Economics*, *63*, 1–24.

Gonzalez-Navarro, M. & Quintana-Domeque, C. (2012, September). *Paving streets for the poor: Experimental analysis of infrastructure effects*. Working Paper.

Hossain, M. A., Sengupta, M. K., Ahamed, S., Rahman, M. M., Mondal, D., Lodh, D., … Chakraborti, D. (2005). Ineffectiveness and poor reliability of arsenic removal plants in West Bengal, India. *Environmnetal Science Technology*, *39*(11), 4300–4306.

Hsiang, S. M. (2010). Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America. *PNAS*, *107*(35).

Humphreys, M., de la Sierra, R. S., & van der Windt, P. (2012, June). *Social and economic impacts of Tuungane: Final report on the effects of a community driven reconstruction program in Eastern Democratic Republic of Congo.*

International Energy Agency. (2013). *Global land transport infrastructure requirements: Estimating road and railway infrastructure capacity and costs to 2050.*

Isham, J., Narayan, D., & Pritchett, L. (1995). Does participation improve performance? Establishing causality with subjective data. *World Bank Economic Review*, *9*(2), 175–2000.

Jedwab, R. & Moradi, A. (2013, October). *Transportation technology and economic change: The impact of colonial railroads on city growth in Africa.*

Khwaja, A. (2004). Is increasing community participation always a good thing? *Journal of the European Economic Association: Papers and Proceedings of the Eighteenth Annual Congress of the European Economic Association*, *2*(2/3), 427–436.

Khwaja, A. I. (2009). Can good projects succeed in bad communities? *Journal of Public Economics*, *93*.

Kleemeier, E. (2000). The impact of participation on sustainability: An analysis of the Malawi rural piped scheme program. *World Development*, *28*(5), 929–944.

Mansuri, G. & Rao, V. (2013). *Localizing development: Does participation work?* Washington, DC: World Bank.

Michaels, G. (2008). The effect of trade on the demand for skill: Evidence from the Interstate Highway system. *Review of Economics and Statistics*, *90*(4).

Michaels, G., Rauch, F., & Redding, S. (2012). Urbanization and structural transformation. *Quarterly Journal of Economics*, *127*(2), pages.

Office of the Chief of Engineers, United States Army. (1948). *List of bridges over the navigable waters of the United States* (Revised to July 1, 1941). US Government Printing Office, Washington.

Olken, B. A. (2007). Monitoring corruption: Evidence from a field experiment in Indonesia. *Journal of Political Economy*, *115*(2).

Olken, B. A. (2010). Direct democracy and local public goods: Evidence from a field experiment in Indonesia. *American Political Science Review*, *104*(2).

Olson, M. (1971). *The logic of collective action*. Harvard University Press.

Plowden, D. (1974). *Bridges: The spans of North America*. W Norton & Company, New York & London.

Pradhan, M., Suryadarma, D., Beatty, A., Wong, M., Gaduh, A., Alisjahbana, A., & Artha, R. P. (2014). Improving educational quality through enhancing community participation: Results from a randomized field experiment in Indonesia. *American Economic Journal: Applied Economics*, *6*(2).

Rao, V. & Ibáñez, A. M. (2005). The social impact of social funds in Jamaica: A 'participatory econometric' analysis of targeting, collective action, and participation in community-driven development. *Journal of Development Studies*, *41*(5), 788–838.

Sara, J. & Katz, T. (1997). *Making rural water sustainable: Report on the impact of project rules.* UNDP – World Bank Water and Sanitation Program.

Smith, A., Lingas, E., & Rahman, M. (2000). Contamination of drinking water by arsenic in Bangladesh: A public health emergency. *Bulletin World Health Organization*, *78*, 1093–1103.

Stiglitz, J. (2002). Participation and development: Perspectives from the comprehensive development paradigm. *Review of Development Economics*, *6*(2), 163–182.

Tompsett, A. (2014). *Bridges: Transport infrastructure and economic geography on the Mississippi and Ohio 1860-2000.*

UNICEF. (2008, October). *Arsenic mitigation in bangladesh.*

Volpe Martincus, C., Carballo, J., & Garcia, P. (2011). *Bridges and trade: Firm-level evidence based on a natural experiment*. Inter-American Development Bank, mimeo.

Weeks, J. (n.d.). *Benton and Stearns Co. Highway 2*. Retrieved from http://www.johnweeks.com/river_mississippi/pages03/rice125th.html

Wooldridge, J. M. (2001). *Econometric analysis of cross section and panel data*. MIT Press.

World Bank. (2004). *World Development Report 2004: Making services work for poor people*. Washington, DC: The World Bank.