ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Severnini, Edson R.

Working Paper

The Power of Hydroelectric Dams: Agglomeration Spillovers

IZA Discussion Papers, No. 8082

Provided in Cooperation with:
IZA Network @ LISER, Luxembourg Institute of Socio-Economic Research (LISER)

Suggested Citation: Severnini, Edson R. (2014) : The Power of Hydroelectric Dams: Agglomeration Spillovers, IZA Discussion Papers, No. 8082, Institute for the Study of Labor (IZA), Bonn

This Version is available at:
https://hdl.handle.net/10419/96760

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

DISCUSSION PAPER SERIES

IZA DP No. 8082

The Power of Hydroelectric Dams: Agglomeration Spillovers

Edson R. Severnini

March 2014

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

The Power of Hydroelectric Dams: Agglomeration Spillovers

Edson R. Severnini
Carnegie Mellon University
and IZA

Discussion Paper No. 8082
March 2014

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA Discussion Paper No. 8082
March 2014

ABSTRACT

The Power of Hydroelectric Dams: Agglomeration Spillovers*

How much of the geographic clustering of economic activity is attributable to agglomeration spillovers as opposed to natural advantages? I present evidence on this question using data on the long-run effects of large scale hydroelectric dams built in the U.S. over the 20th century, obtained through a unique comparison between counties with or without dams but with similar hydropower potential. Until mid-century, the availability of cheap local power from hydroelectric dams conveyed an important advantage that attracted industry and population. By the 1950s, however, these advantages were attenuated by improvements in the efficiency of thermal power generation and the advent of high tension transmission lines. Using a novel combination of synthetic control methods and event-study techniques, I show that, on average, dams built before 1950 had substantial short run effects on local population and employment growth, whereas those built after 1950 had no such effects. Moreover, the impact of pre-1950 dams persisted and continued to grow after the advantages of cheap local hydroelectricity were attenuated, suggesting the presence of important agglomeration spillovers. Over a 50 year horizon, I estimate that at least one half of the long run effect of pre-1950 dams is due to spillovers. The estimated short and long run effects are highly robust to alternative procedures for selecting synthetic controls, to controls for confounding factors such as proximity to transportation networks, and to alternative sample restrictions, such as dropping dams built by the Tennessee Valley Authority or removing control counties with environmental regulations. I also find small local agglomeration effects from smaller dam projects, and small spillovers to nearby locations from large dams.

JEL Classification: N92, R11, R12, Q42

Keywords: hydroelectric dams, agglomeration spillovers, employment growth, event-study analysis with synthetic control methods

Corresponding author:

Edson R. Severnini
Carnegie Mellon University
4800 Forbes Ave
Pittsburgh, PA 15213
USA
E-mail: ersevernini@gmail.com

* I am extremely grateful to David Card, Patrick Kline and Enrico Moretti for invaluable guidance and support throughout the duration of this project. I also thank Ana Rute Cardoso, Mayssa Dabaghi, Monica Deza, Frederico Finan, Willa Friedman, Francois Gerard, Tadeja Gracner, Hedvig Horvath, Brian Kovak, Attila Lindner, Tarso Mori Madeira, Joana Naritomi, Valentina Paredes, Markus Pelger, Michaela Pagel, Issi Romem, Michel Serafinelli, Lowell Taylor, Reed Walker, and seminar participants at UC Berkeley, Carnegie Mellon University, London School of Economics, Toulouse School of Economics, NBER Summer Institute 2013 - Development of the American Economy (DAE), AEA Annual Meeting 2014, 8th Meeting of the Urban Economics Association, George Washington University, Case Western Reserve University, University of Calgary, World Bank, Inter-American Development Bank, Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Getulio Vargas Foundation (EPGE), Brazilian Econometric Society 35th Annual Meeting, and Cornerstone Research for useful comments. I acknowledge generous support from the Center for Labor Economics at UC Berkeley.

1 Introduction

Economic activity is geographically concentrated (e.g., Ellison and Glaeser, 1997, 1999; Duranton and Overman 2005, 2008; Ellison, Glaeser and Kerr, 2010; and Moretti, 2011). However, the mere occurrence of agglomeration does not imply the existence of agglomeration spillovers. The location of economic agents in space might be due to agglomeration economies (increasing returns, knowledge spillovers, and pooling of specialized skills) and/or natural advantages (topography, climate, and resource endowments). Although there is recent evidence that locational choices are not uniquely determined by fundamentals (e.g., Redding, Sturm, and Wolf, 2011; Bleakley and Lin, 2012; and Kline and Moretti, 2014), influential studies have found a major role for natural advantages by examining growth in the aftermath of war bombing (e.g., Davis and Weinstein, 2002, 2008; Brakman, Garretsen and Schramm, 2004; and Miguel and Roland; 2011). Despite significant losses during war, bombed cities almost all returned to their prewar growth paths. In this paper, I present new evidence on the importance of agglomeration spillovers in population density by keeping natural advantages constant[1], and evaluate whether such spillovers are strong enough to generate long-run effects. Instead of bombing, I use installation of large hydroelectric dams in the U.S. in the first half of the twentieth century.

Throughout my analysis, I use a unique database of U.S. counties with similar natural endowments associated with comparable suitability for hydroelectric projects. In the 1990s, a team of engineers determined the hydropower potential across the nation at the request of the U.S. Department of Energy. Thus, natural advantages are arguably held constant. I define counties with hydroelectric dams as "treated" counties, and combinations of counties with no dams but with hydropower potential as counterfactuals, or "synthetic control" counties. To identify agglomeration spillovers, I build on Bleakley and Lin (2012) and rely on the expansion of the electrical grid around 1950, and the consequent attenuation of the cheap local power (CLP) advantage brought about by the development of hydro projects in the first half of the twentieth century[2].

Figure I depicts my identification strategy more clearly. Time is represented in the horizontal axis. My analysis spans the entire twentieth century. Log of population density, my main outcome variable capturing economic activity, is represented in the vertical axis. Suppose that two counties - treated and counterfactual - have the same population density up to the construction of a hydroelectric dam. From the completion of the dam until mid century, time period displayed between the two vertical dotted lines, the treated county experiences a temporary CLP advantage. I argue that counties hosting dams in the first half of the century have a temporary advantage because of the *local* availability of cheap power. Indeed, I provide suggestive evidence that electricity prices were approximately 45 percent lower in treated counties. Electricity transmission networks used to transmit power efficiently only around dam sites before the advent of modern high-tension transmission lines. Also, the competing technology of thermal power generation was not advanced enough to face the low-generating costs of hydroelectric turbines. That

[1]Rosenthal and Strange (2004) argue that the agglomeration effects found in the literature should be interpreted as upper bound estimates of agglomeration economies. They point out that those estimates most likely include the influence of natural advantages.

[2]Bleakley and Lin (2012) use a fading locational advantage - obsolescence of portage sites - to examine the role of path dependence in the spatial distribution of economic activity the U.S.

CLP advantage gives rise to a strong concentration of economic activity around dam sites, represented by the CLP agglomeration jump between dam completion and 1950 in the Figure. In the second half of the century, technological improvements in thermal power generation make fossil fuels (mostly coal and natural gas) and nuclear energy almost as competitive as hydropower in producing electricity. In addition, the construction of high-voltage transmission lines weakens the need of power generation next to consumers. Therefore, the appeal of local cheap hydroelectricity reduces considerably after 1950.

The dynamics of population density in the second half of the twentieth century, represented on the right-hand side of the second vertical dotted line in Figure I, sheds light on the presence and magnitude of agglomeration spillovers. If the jump in population density that happens between the completion of the dam and the middle of century reverts, then there would be no agglomeration spillovers, and the evidence would be similar to that of most cities bombed during wars. If population density stabilizes after the jump, then there would be persistence in economic activity probably due to sunk investments and/or durability of goods such as housing. Now, if population density continues growing after the initial CLP-driven jump, then there would be agglomeration spillovers, and the magnitude would be measured by the difference between the post-1950 level and the 1950 level of population density. My empirical analysis shows evidence of reversion, stabilization and agglomeration spillovers for individual U.S. counties with hydroelectric dams, but the average result indicates the presence of large agglomeration spillovers.

There is some evidence that infrastructure projects such as hydroelectric dams may affect local economic activity for a long time (e.g., Kline and Moretti, 2014). Besides raising the stock of capital, public infrastructure may increase local productivity directly. Also, hydro projects might promote the development of certain areas permanently because of the amenities that they provide. Recreational actitivies around reservoirs are a good example. In order to eliminate these confounding factors from my estimates of agglomeration spillovers, I subtract any effects of hydro dams built after 1950. Therefore, my ultimate measure of spillovers is the growth in population density of counties with pre-1950 dams over and above the growth of counties with post-1950 dams, in post-1950 years, when the advantage of cheap local electricity is attenuated[3]. Although this procedure refines my identification of spillovers, it removes any potential agglomeration effects of post-1950 dams. In this sense, my estimate represents a lower bound of agglomeration spillovers.

To illustrate my approach, Figure II plots the log population density from 1900 to 2000 for two counties: one with a hydroelectric dam built before 1950, and the other with a dam built after 1950. The solid line with solid circles displays the observed time series of log population density for each county, and the dashed line with hollow squares depicts the predicted time series of a counterfactual county. I explain the procedure to find a counterfactual in the next paragraph. The left chart, for Blount County, Tennessee, shows a high degree of agglomeration after the construction of the Calderwood Dam, in the 1930s. There is also clear evidence of agglomeration spillovers, over and above the agglomeration accumulated until 1950 (dotted line with hollow triangles), which one could see as arising solely from the advantage of cheap local hydroelectricity in the first half of the twentieth century. In addition, there seems to be strong

[3]This is related to Greenstone, Hornbeck and Moretti's (2010) strategy to identify agglomeration economies. They examine whether attracting a Million Dollar Plant (MDP) leads to increases in total factor productivity (TFP) of incumbent manufacturing plants, over and above the mechanical effect of the new plant on the TFP of the winning county.

evidence of path dependence. While the local economic advantage of nearby hydroelectricity is decreasing in the second half of the century, population density is still increasing. It is fair to say that the trend flattens out after 1950, but the upward slope is still unequivocally positive. The right chart, for Lincoln County, North Carolina, depicts a low degree of agglomeration, even right after the installation of the Cowans Ford Dam, in the 1960s, when increases in the stock of capital were still recent. As mentioned above, this feature of the data helps to strengthen the interpretation of the effects of pre-1950 dams in post-1950 years as agglomeration spillovers. Though less extreme, I find similar patterns for the impact of a broad sample of large hydroelectric dams on the economic activity of U.S. counties.

To obtain average estimates of agglomeration spillovers, I need credible impact estimates of hydroelectric dams. Despite the number of large hydro dams in the U.S. - 185 dams with a capacity of 100 megawatts or more - there is surprisingly little research on their local economic impacts. Hence, in this paper, I also present the first evaluation of the short- and long-run effects of new hydropower facilities on local economies[4]. To estimate the causal effects of hydro dams, I employ a novel empirical strategy. I combine synthetic control methods (Abadie and Gardeazabal, 2003; and Abadie, Diamond and Hainmueller, 2010) with event-study techniques (Jacobson, LaLonde, and Sullivan, 1993). First, I use the synthetic control estimator to generate a counterfactual for each treated county. Each counterfactual, which I refer to as a "synthetic control county", is a weighted average of the originally defined control counties that reproduces more closely the outcome trajectory that the affected county would have experienced in the absence of dams. Then, I pool all pairs of treated and synthetic control counties, and proceed with my event-study analysis. A byproduct of the synthetic control approach is the estimation of the impact of hydro dams for each treated county separately. This allows me to examine heterogeneity in dam effects by plotting the distribution of effects across treated counties for each decade after dam installation.

My impact estimates of hydro dams show that counties with dams built before 1950 have population density increased by approximately 51 percent after 30 years, and 139 percent after 60 years, indicating substantially different short- and long-term effects. Kline's (2010) theoretical argument that assessing place-based policies requires understanding the long-run effects of temporary interventions finds clear empirical support here. On the other hand, counties with dams built after 1950 have no statistically significant effects. I argue that the large difference in the impact of pre- and post-1950 hydro dams can be accounted for by the attenuation of the advantage of cheap local hydroelectricity in the second half of the twentieth century, as mentioned above.

Regarding agglomeration spillovers, the causal effects of hydro dams imply an average lower bound of up to 45 percent five decades after dam construction (three decades after spillovers kick in). This long-run estimate represents almost half of the full impact of hydro dams over the same time span. Interestingly, my short-run estimate of agglomeration externalities is very close to that of Greenstone, Hornbeck and Moretti (2010). My lower bound nearly a decade after spillovers kick in is around 11.5 percent, while their estimate five years after the opening of a Million Dollar Plant is 12 percent. Taken together, my short- and long-run estimates point to an amplification of spillovers over time. This suggests that spillovers may

[4]Other studies that examine long-term adjustments to temporary shocks are, for example, Blanchard and Katz (1992), Davis and Weinstein (2002, 2008), Redding and Sturm (2008), Miguel and Roland (2011), Hornbeck (2012), and Kline and Moretti (2014).

sustain high levels of local development in the long-run.

I also find that the estimated short- and long-run effects are highly robust to alternative procedures for selecting synthetic controls, to controls for confounding factors such as proximity to transportation networks, and to alternative sample restrictions, such as dropping dams built by the Tennessee Valley Authority (TVA) or removing control counties with environmental regulations. I also provide evidence of small local agglomeration effects from smaller dam projects, and small spillovers to nearby locations from large dams.

There is a growing empirical literature investigating path dependence. Much of it is concerned with persistence of the spatial distribution of economic activity following the attenuation or elimination of a locational advantage. The seminal contribution is Bleakley and Lin (2012). They document the continuing importance of historical portage sites after their original advantages had become obsolete. Only a small number of recent papers emphasize the mechanisms driving those persistent effects. In what is probably the most closely related paper to my own, Kline and Moretti (2014) provide evidence that the long-run effects of one of the most ambitious regional development programs in U.S. history - the TVA -, are consistent with the presence of agglomeration economies in manufacturing. A structural approach is then used to suggest that those agglomeration gains in the TVA region are offset by losses in the rest of the country. My paper not only reinforces the role of agglomeration spillovers in understanding the long-term impact of large infrastructure projects, but crucially shows that those spillovers are prominent even when natural advantages are directly and flexibly taken into account in the analysis. This is an important point because Rosenthal and Strange (2004) argue that most agglomeration effects found in the literature suffer from biases due to the inability to control for natural advantages.

Finally, this work also contributes to the literature dealing with the effects of dams and electrification. Duflo and Pande (2007)'s influential study of the impact of large irrigation dams on agricultural production and poverty rates in India pioneers the use of geographical suitability to identify the effects of dams. Instead of just comparing outcomes of districts with and without irrigation dams, they use variation in dam construction induced by differences in river gradient across districts within Indian states to obtain instrumental variable estimates. Lipscomb, Mobarak and Barham (2013) follow suit by using hypothetical electricity grids associated with only geography-based cost considerations to estimate the development effects of electrification across Brazil over the period 1960-2000. My approach looks more like a randomized control trial. My analysis is restricted to U.S. counties with similar hydropower potential, some of which received hydro dams and some did not.

The remainder of the paper is organized as follows. Section 2 provides a historical discussion of the process of electrification in the U.S., focusing on the local importance of hydroelectricity in the first half of the twentieth century. Section 3 presents a simple theoretical framework to identify agglomeration spillovers. Section 4 discusses the research design and related issues. Section 5 describes the databases used in the study. Section 6 outlines the methodology for the empirical analysis. Section 7 reports and discusses results regarding the impact estimates of hydro dams, the estimate of agglomeration spillovers, and the persistence of dam effects. Finally, Section 8 provides some concluding remarks.

2 Historical Background

My strategy to identify agglomeration spillovers relies on the strength of the cheap local power (CLP) advantage arising from the construction of hydroelectric dams in the first half of the twentieth century, and the attenuation of that advantage in the second half of the century. In this section, I provide suggestive evidence on the CLP advantage using cross-county variation in prices of electricity purchased by manufacturing, and discuss the historical context. Then, I present a case study illustrating key points of my identification strategy.

2.1 Cheap Local Power (CLP) Advantage

In the first half of the twentieth century, hydroelectric dams seem to provide cheap and abundant electricity only for surrounding areas. In fact, power appears to be much cheaper in counties hosting pre-1950 dams than in counties with hydropower potential but no hydroelectric facilities. By regressing the log of the unit value of electricity purchased by manufacturing in 1947 (Bureau of the Census, 1949) on a dummy for pre-1950 treatment, controlling for state fixed effects, a cubic function on latitude and longitude, and 50-year average rainfall and 50-year average temperature for each season of the year, I find a coefficient of approximately -0.60 (s.e. 0.22). This means that electricity price was roughly 45 percent lower in pre-1950 treated counties relative to controls[5]. This is not the case, however, for counties that had hydropower plants constructed only in the second half of the century. The estimated coefficient for the post-1950 treatment in a regression similar to the one described above was not statistically significant (-0.28, with s.e. 0.34).

In the second half of the century, construction of high-tension transmission lines and technological improvements in the competing thermal power generation may have attenuated differences in electricity prices across the country. To test this hypothesis, I redo my analysis above with electricity prices in 2000, constructed from data available from the Energy Information Administration (EIA) form 861, as in Kahn and Mansur (2012). Relative to the control group, neither pre-1950 nor post-1950 treated counties have lower prices that are statistically significant. For pre-1950, the coefficient of the dummy of treatment is approximately -0.10 (s.e. 0.06), and for post-1950, it is nearly -0.07 (s.e. 0.05). Therefore, my assumption of attenuation of the CLP advantage after 1950 seems reasonable.

Another piece of evidence supporting cheap electricity as the force driving my results later on comes from the dispersion of prices across counties. If electricity could be transmitted costlessly from suppliers to consumers, and electricity markets were nationally competitive, then we would observe no difference in prices over the country. Because there was no major change in regulation around the middle of the century, variations in dispersion in that period might be associated with transmission costs. The standard deviation of the log of the unit value of electricity purchased by manufacturing decreased from nearly 0.46,

[5]It is important to point out that Kitchens (2012) finds that the price of electricity for residential consumers was not lower between 1929 and 1955 in the TVA area, where many pre-1950 hydro dams were built. However, he does provide evidence that TVA electric rates for large industrial consumers and commerical firms were lower than comparable electric service. According to him, rates for large industrial firms were between 3 and 34 percent lower depending on the demand capacity and usage level.

in 1947, to 0.38 (Davis et al., 2013), in 1972. Because most of the high-voltage transmission lines were built in the 1950s and 1960s, this 17 percent drop in the standard deviation may reflect a higher degree of spatial interconnection of the electrical grid.

In order to better understand the assumption of CLP advantage, let us go back to the historical context around the discovery and emergence of electricity. The invention of the continuous direct current (D.C.) dynamo in 1870 was a major breakthrough. Although lighting and traction (i.e., electric streetcars) were among the first benefits of the new technology, the most important consequence of the dynamo revolution was the process of industrial electrification following the invention of the polyphase alternating current (A.C.) motor in 1888 (see chronology in Table I). Between 1880 and 1930, the production and distribution of mechanical power rapidly evolved from water and steam-driven line shafts connected by belts to electric motors that drove individual machines.

During this rapid transition in energy use, the U.S. power industry became highly specialized and electric utilities gained prominence. In early 1900s, around 60 percent of the electricity used in manufacturing was generated by the establishments themselves. By 1917, however, eletric utilities were already generating more power than industrial plants. By the late 1950s, their production had reached almost 90 percent of total generation, as shown in Table II. This shift of electricity generation to the power sector led manufacturing to become sensitive to local availability of power.

In the beginning, hydroelectricity prevailed as a source of motive power. This was probably due to familiarity with water power and the high cost of coal to drive steam turbines. The availability of (cheap) hydroelectricity significantly affected the locational decisions of industrial plants. After the first hydroelectric generating station was built near Niagara Falls in 1881, manufacturing flourished around hydro dam sites. Although "*its supply is limited and plants have to locate where favourable sites exist*" (Schramm, 1969, p.220), hydroelectricity continued having a large comparative advantage in the U.S. until the 1950s. The costless energy content of falling water, and the high mechanical efficiency of hydraulic turbines, were among the key factors maintaining that advantage.

Important innovations, however, took place in the field of thermal power generation. Technological improvements increased boiler temperatures and operating pressures substantially, attaining greater thermal efficiencies. As depicted in Figure III, thermal efficiencies increased gradually until the early 1940s, and rapidly in the post-World War II years, reaching their current level around 1960. At the same time, electric utilities started constructing larger generating plants to capture economies of scale. As a result, steam power gave location flexibility to firms. Thermal power plants could be built almost anywhere, provided that fuels and a minimum amount of boiler and cooling water were available.

Summarizing the transformation of the electricity industry around the middle of the twentieth century, Schramm (1969) states that "*the historical cost advantage of hydropower vis-à-vis other generating methods has been (...) drastically reduced. (...) Fifteen to twenty years ago [1949-1954], power-intensive industries found it advantageous to move their plants to sites where cheap hydropower was still available. Today, with generating cost differentials reduced so drastically, these advantages have all but disappeared.*" (p.225-226). The tremendous post-1950 growth of steampower relative to hydropower, presented in Figure IV, corroborates his conclusions.

Yet another major change happened in the electricity sector in the 1950s: the emergence of higher voltage transmission lines. The nineteenth century inventors who first began to harness electricity to useful purposes did so by putting their small generators right next to the machines that used electricity. The development of the A.C. system allowed power lines to transmit electricity over much longer distances. In 1896, for instance, an eleven-kilovolt A.C. line was built to connect a hydroelectric generating station at Niagara Falls to Buffalo, twenty miles away. From then on, the voltage of typical transmission lines grew rapidly.

Nevertheless, until 1950, the number of circuit miles of high-voltage transmission lines - 230 kilovolts and above - was extremely small in the U.S. That number more than tripled to over 60,000 circuit miles in the 1960s (Brown and Sedano, 2004). This was a huge expansion: approximately 40 percent of all high-voltage transmission lines installed in the U.S. at the end of the twentieth century had been constructed in the 1950s and 1960s. Such developments gave utilities access to ever more distant power sources, further reducing the appeal of cheap local hydroelectricity.

2.2 Case Study - Bonnevile Dam

To illustrate the key ideas advanced in this study with historical anecdotal evidence, I discuss the case of the Bonneville Dam, a hydropower plant built in the 1930s. In an effort to prevent extortion against the public by giant electric utilities, and to provide employment during the Great Depression, the U.S. government started large hydroelectric projects in the West. One of them was the Bonneville Dam, on the Columbia River between the states of Washington and Oregon. Construction began in June 1934, and commercial operation was achieved in 1938.

In 1937, Congress created the Bonneville Power Administration (BPA) to deliver and sell the power from Bonneville Dam (see important historical facts in Table I). The first line connected the dam to Cascade Locks, a small town just three miles away. Major construction from the 1940s through the 1960s created networks and loops of high-voltage wire touching most parts of BPA's current service territory, which includes Idaho, Oregon and Washington, and small portions of California, Montana, Nevada, Utah and Wyoming (BPA, n.d.). During that time, Congress authorized BPA to commercialize power from other federal dams on the Columbia and its tributaries. By 1940, however, the federal pricing policy was set: all federal power was marketed at the lowest possible price while still covering costs.

The initial wholesale cost of power from Bonneville Dam was \$17.50/kW year (0.2 cents/kWh), a rate that was maintained for 28 years. To take advantage of this cheap and abundant electricity, the Aluminum Company of America (ALCOA) and Reynolds Metals started to mobilize to build aluminum smelters in the Northwest. ALCOA purchased property in Vancouver, Washington, in December 1939 and poured the first ingot on September 23, 1940 (Voller, 2010). Reynolds Metals purchased the property for a smelter in Longview, Washington, in 1940. By this time, with war raging in Europe, the U.S. government saw a strategic need to increase aluminum production for the impending defense effort and agreed to underwrite the construction of the Longview facility. The smelter opened in September 1941, just in time to meet the aircraft industry's increased need for aluminum (McClary, 2008).

Hydroelectric dams brought prosperity to their hosting locations, attracting hundreds of workers.

Vancouver, for instance, saw an industrial boom in the 1940s, including the Kaiser shipyard and the Boise Cascade paper mill, besides ALCOA (Jollata, 2004). Over the years, though, as many heavy industries left the U.S., Vancouver's economy has largely changed to high tech and service industry jobs. The city contains the corporate headquarters for Nautilus, Inc. and The Holland (parent company of the Burgerville, U.S. restaurant chain). It seems that nowadays other forces attract people to Vancouver. They might be agglomeration spillovers.

3 Theoretical Framework

In this section, I enrich Greenstone, Hornbeck and Moretti's (2010) framework with insights from Duflo and Pande (2007) to illustrate how the installation of hydroelectric dams could affect the attractiveness of hosting counties through advantage of local cheap electricity and agglomeration spillovers. I focus on the profitability of firms in hosting counties, and I assume factor-neutral spillovers related to the impact of hydro dams.

Suppose that all firms have a production technology that uses labor, capital, land and electricity to produce a nationally traded good whose price is fixed and normalized to one. Firms choose their amount of labor, L, capital, K, land, T, and electricity, E, to maximize profits:

$$\max_{L,K,T,E} f(A, L, K, E) - wL - rK - qT - sE,$$

where w, r, q and s are input prices, and A is a productivity shifter (TFP).

More specifically, A includes all factors that affect the productivity of labor, capital, land and electricity equally, such as technology and agglomeration spillovers of hydro dams, if they exist. To explicitly allow for such agglomeration externalities, I let A depend on the population density in a county, N:

$$A = A(N).$$

Factor-neutral agglomeration spillovers exist if A increases in N: $\partial A/\partial N > 0$.

Let $L^*(w, r, q, s)$ be the optimal level of labor inputs, given the prevailing wage, cost of capital, rent, electricity rate, and population density. Similarly, let $K^*(w, r, q, s)$, $T^*(w, r, q, s)$ and $E^*(w, r, q, s)$ be the optimal level of capital, land and electricity, respectively. In equilibrium, L^*, K^*, T^* and E^* are set so that the marginal product of each of these four factors is equal to its price. I assume that capital is internationally traded, so its price r does not depend on local demand or supply conditions. On the other hand, I allow wage and rent to depend on local economic conditions: $w(N)$ and $q(N)$. In particular, I allow the supply of labor and land to be less than infinitely elastic at the county level. Hence, $w(N)$ represents the inverse of the reduced-form labor supply function that links population density in a county, N, to the local wage level, w. Similarly, $q(N)$ represents the inverse of the reduced-form land supply function (Greenstone, Hornbeck and Moretti, 2010).

The equilibrium level of profits, Π^*, can be written as

$$\begin{aligned}\Pi^* &= f[A(N),\, L^*(w(N),r,q(N),s),\, K^*(w(N),r,q(N),s),\, T^*(w(N),r,q(N),s),\, E^*(w(N),r,q(N),s)] \\ &\quad -w(N)L^*(w(N),r,q(N),s) - rK^*(w(N),r,q(N),s) \\ &\quad -q(N)T^*(w(N),r,q(N),s) - sE^*(w(N),r,q(N),s),\end{aligned}$$

where now I make explicit the fact that TFP, wages and rents depend on the population density in a county.

Consider the total derivative of profits with respect to a change in electricity price and in population density:

$$\frac{d\Pi^*}{ds} = \frac{\partial L^*}{\partial s}\left(\frac{\partial f}{\partial L} - w\right) + \frac{\partial K^*}{\partial s}\left(\frac{\partial f}{\partial K} - r\right) + \frac{\partial T^*}{\partial s}\left(\frac{\partial f}{\partial T} - q\right) + \frac{\partial E^*}{\partial s}\left(\frac{\partial f}{\partial E} - s\right) - E^*, \tag{1}$$

$$\begin{aligned}\frac{d\Pi^*}{dN} &= \left(\frac{\partial f}{\partial A} \times \frac{\partial A}{\partial N}\right) - \left(\frac{\partial w}{\partial N}L^*\right) - \left(\frac{\partial q}{\partial N}T*\right) \qquad (2)\\ &\quad + \frac{\partial w}{\partial N}\left\{\frac{\partial L^*}{\partial w}\left(\frac{\partial f}{\partial L} - w\right) + \frac{\partial K^*}{\partial w}\left(\frac{\partial f}{\partial K} - r\right) + \frac{\partial T*}{\partial w}\left(\frac{\partial f}{\partial T} - q\right) + \frac{\partial E^*}{\partial w}\left(\frac{\partial f}{\partial E} - s\right)\right\} \\ &\quad + \frac{\partial q}{\partial N}\left\{\frac{\partial L^*}{\partial q}\left(\frac{\partial f}{\partial L} - w\right) + \frac{\partial K^*}{\partial q}\left(\frac{\partial f}{\partial K} - r\right) + \frac{\partial T*}{\partial q}\left(\frac{\partial f}{\partial T} - q\right) + \frac{\partial E^*}{\partial q}\left(\frac{\partial f}{\partial E} - s\right)\right\}. \qquad (3)\end{aligned}$$

If all firms are price takers and all factors are paid their marginal product, equations (1) and (2) simplify to

$$\frac{d\Pi^*}{ds} = -E^* < 0, \tag{4}$$

$$\frac{d\Pi^*}{dN} = \left(\frac{\partial f}{\partial A} \times \frac{\partial A}{\partial N}\right) - \left\{\left(\frac{\partial w}{\partial N}L^*\right) + \left(\frac{\partial q}{\partial N}T*\right)\right\}. \tag{5}$$

Equation (4) is just the Hotelling's lemma, implying that profits are higher when local electricity prices are lower. Equation (5) points out that the effect of an increase in N is the sum of two opposing forces. The first term, $\left(\frac{\partial f}{\partial A} \times \frac{\partial A}{\partial N}\right)$, represents the effect of positive agglomeration spillovers, if they exist. Agglomeration spillovers of hydro dams allow firms to produce more output using the same amount of inputs. On the other hand, the second term, $\left\{\left(\frac{\partial w}{\partial N}L^*\right) + \left(\frac{\partial q}{\partial N}T*\right)\right\}$, represents negative effects from increases in the cost of production, specifically, wages and rents. Intuitively, an increase in N is an increase in the level of economic activity in the county and therefore an increase in the local demand for labor ($\partial w/\partial N > 0$) and land ($\partial q/\partial N > 0$) . Unlike the beneficial effects of agglomeration spillovers, the increase in factor prices is costly for firms because they now have to compete more for locally scarce resources[6].

[6]As shown in Table C.I, in the Appendix, I find no strong evidence that dams affect land prices: $\partial q/\partial N$ appears small. Kline and Moretti (2014) do not find effects of TVA on wages either. In that case, it may be reasonable to assume that $\partial w/\partial N$ is also small or even zero.

Equations (4) and (5) provide useful guidance on what would happen after installation of hydroelectric dams. Assuming free entry, firms would move to counties hosting hydro dams to exploit rents arising from cheaper electricity ($d\Pi^*/ds < 0$). Supposing imperfect substitution between labor and electricity, firms would hire lots of workers. This would potentially pull individuals from other locations, leading to an increase in population density. In the presence of strong agglomeration spillovers, this increase in population density would trigger a positive feedback mechanism that would continue attracting both firms and workers to those places ($d\Pi^*/dN > 0$). Strong spillovers would keep attracting businesses even when cheap electricity had become available everywhere.

4 Location Decisions and Research Design

In testing for the presence of agglomeration spillovers, a key econometric challenge is that concentration of economic activity also can be generated by natural advantages. Specific geographical features that bring people and businesses to an area, such as detailed topography, resource endowments, and climate, are often unobserved, but they are problematically correlated to those location decisions.

Therefore, a naive comparison of population growth across counties with distinct natural endowments is likely to yield biased estimates of agglomeration spillovers. Some places might attract people because of such externalities, but some others might pull people in just because of an unobserved geographical attribute. Credible estimates of spillovers require the identification of locations which are similar in natural advantages.

The first appealing characteristic of my research design is the narrowing of my sample to encompass counties that are comparable in some natural features. Indeed, I compare only counties that have similar potential to generate hydroelectricity, as determined by an engineering team in the beginning of the 1990s, at the request of the U.S. Department of Energy. As is well-known in the engineering literature, suitability of sites for hydro dams depends on parameters associated with topography and inflow in the catchment area, morphology of the river valley, geological and geotechnical conditions, and climate and flood regime. Therefore, I believe my approach controls for important geographical peculiarities of counties in my sample.

Locations with similar geography tend to have similar economic activity. It would then be desirable to have temporary shocks that would make some of these areas more attractive than others. In that case, workers and firms would concentrate in certain areas even if they could enjoy the same geographical features in other places. After the interruption of the shocks, affected areas would have higher population density. If they continued having higher population growth after the cessation of the shocks, this would be an indication of the presence of agglomeration spillovers.

The second valuable feature of my research design is the use of the appeal of cheap local hydroelectricity in the first half of the twentieth century as such a shock. Not every county with suitable dam sites ever had hydropower plants constructed. Therefore, not every county with appropriate dam sites had access to the cheapest source of electricity until 1950. By the middle of the century, however, such advantage was reduced considerably, as argued in previous sections. As thermal power generation was enjoying major

technological improvements, and high-voltage transmission lines were being constructed, cheap electricity was becoming available to most counties across the nation.

Obviously, the validity of my research design depends on the assumption that places with hydro potential where dams were not built provide a valid counterfactual for similar places where dams were built. This in turn requires a clearer understanding of why dams were not built in some counties with hydroelectric potential. Prior to World War I, hydroelectric power development was mostly a private venture. However, with private hydro plants increasingly interfering with navigation in the East and Midwest, government regulation evolved to become stricter. Congress initially attempted to regulate dam construction through the Rivers and Harbors Acts of 1890 and 1899, requiring that dam sites and plans for dams on navigable rivers be approved by the U.S. Army Corps of Engineers and the Secretary of War. Then, the Right-of-Way Act of 1901 gave the Secretary of the Interior the authority to grant rights-of-way over public lands for dams, reservoirs, waterpower plants, and transmission lines (see Table I for key historical facts related to hydroelectricity in the U.S.).

Although "*between 1894 and 1906 Congress issued 30 permits for private dams, mostly along the Mississippi River*" (Billington, Jackson and Melosi, 2005, p.37), the federal government began to reserve waterpower sites for conservation and wise use, and to enter the business of hydroelectricity. Indeed, "*in the 1903 veto of private construction of a dam and power stations on the Tennessee River at Muscle Shoals, Alabama, Roosevelt protected the site for later government development, but he also helped to establish the principle of national ownership of resources previously considered only of local value.*" (Billington, Jackson and Melosi, 2005, p.37).

The General Dam Act of 1906 was, perhaps, the legislation that most discouraged the entrance of private enterprise into the hydroelectric sector. It "*standardized regulations concerning private power development, requiring dam owners to maintain and operate navigation facilities - without compensation - when necessary at hydroelectric power sites.*" (Billington, Jackson and Melosi, 2005, p.38). Private companies fought for more favorable legislation, but ended up accepting the permit system. At the same time, the federal government started to link hydropower to plans for waterway improvements. A 1910 amendment to the 1906 act, for instance, underscored hydropower as a mechanism for navigation and flood control projects. The connection between hydropower and local development, and the participation of the federal government in the power sector, would only increase afterwards.

The organization of the electricity markets in the beginning of the century also attracted governmental intervention. At first, private electric utility companies dominated the market. However, the proliferation, consolidation, and complexity of such companies coincided with a number of financial and securities abuses, sometimes inflating costs that were passed through to the retail customers. As a response, "*Georgia, New York and Wisconsin established State public service commissions in 1907, followed quickly by more than 20 other states. Basic state powers included the authority to franchise the utilities, to regulate their rates, financing, and service, and to establish utility accounting systems*" (EIA, n.d.).

The foundations for strong federal involvement in the electricity industry were also established between 1900 and 1930 (EIA, n.d.). First, the electric power industry became recognized as a natural monopoly in interstate commerce subject to federal regulation (Supreme Court Ruling of 1927 - see chronology). Sec-

ond, the federal government owned most of the nation's hydroelectric resources. Third, federal economic development programs accelerated, including electricity generation. After 1930, the federal government became both a regulator of private utilities and a major producer of electricity. Both regulation and production were aimed at generating less expensive electricity for customers.

Federal participation also increased because of national efforts to overcome the Great Depression in the 1930s, and to meet the massive electricity requirements for wartime production in the 1940s. Considerable funding was provided for the construction of large federal dams and hydroelectric facilities. This is the period known as the Big Dam Era (Billington, Jackson and Melosi, 2005). Bonneville Dam, completed in 1938, was a public works project to help relieve regional unemployment during the Great Depression. Grand Coulee Dam, opened in 1942, supplied the electricity needed to produce planes and other war material in support of World War II efforts. Later on, to meet escalating electricity needs in response to the dramatic expansion of consumer demand and industrial production throughout the decades of the 1950s, 1960s, and 1970s, many new electric generating facilities, including hydroelectric developments, were constructed.

As we can see, the decision of where to construct hydroelectric dams basically changed from the private sector to the federal government in the first quarter of the twentieth century. Regional development concerns - and environmental issues in the second half of the century - frequently guided the allocation of dams throughout the nation. Therefore, most hydro projects can be characterized as shocks of federal investment to local economies, creating rents to be exploited by local consumers and firms. Moreover, counties with higher economic growth potential have not always attracted more investments than counties with lower growth potential. Instead, Hansen et al. (2011) argue that politics might have shifted the eventual location of major water infrastructure projects away from what otherwise might have been the optimal location. They provide evidence suggesting that membership in congressional committees for water resources, agriculture, and appropriations, often times unrelated to population pressures, generally has a positive and significant impact on the number of dams and the proportion of dams constructed in a state.

As the discussion above attests, my research design may be valid under relatively mild assumptions. In my main estimation, I end up using the group of counties with hydropower potential that was not developed as my potential control group. Nevertheless, I also control flexibly for climate variables and geographic coordinates, allowing them to vary with time, and use synthetic control methods to match treated and control counties more closely (details are presented in section 6)[7]. In my robustness checks, I include other sets of controls such as proximity to transportation networks, and consider alternative definitions for my potential control group, such as counties with hydropower potential but no environmental regulations.

[7]As discussed in the next sections, the synthetic control method allows me to match the trajectory of population or employment density in treated and control counties before the construction of the dams. If by any chance private firms influenced the selection of dam sites due to economic activity considerations, such as ALCOA suggesting the location of the Gran Coulee Dam to officials in Washington D.C., then that match takes such concerns into account in the empirical analysis. In that case, individual counties in the potential control group might not be suitable controls, but the constructed "synthetic control" counties, combinations of potential control counties which will be introduced later on, may be arguably valid controls for the treated ones.

A third and last feature of my research design is its ability to control for the effect of direct investment in the estimation of agglomeration externalities. Counties that hosted hydroelectric dams in the first half of the twentieth century might have had another important initial stimulus in their local economies, on top of the appeal of cheap local hydroelectricity. The construction of hydropower plants might have increased the local stock of capital substantially, and might have attracted workers and firms as well. Because not all counties with hydro dams had their facilities built at the same time, I use the counties with dams built after 1950 to refine my identification of spillovers. Since these counties experienced the same infusion of capital as the counties with pre-1950 dams, but did not enjoy the advantage of cheap local power, they can be used to purge the estimates of agglomeration economies of this source of confoundedness.

5 Data Description

My basic dataset is a balanced panel of 154 U.S. counties covering the period from 1900 to 2000. It includes all counties with a hydropower potential of 100 megawatts or more. This choice of sample ensures that counties are similar with respect to natural endowments. County-level data on population and employment are drawn from the U.S. census of population (Haines and ICPSR, 2010; Minnesota Population Center - NHGIS, 2011). Variables of interest include population density and employment density, which have usually been used as proxies for economic activity. Other data sources include climate data from Schlenker and Roberts (2009), market access data from Donaldson and Hornbeck (2012), and hydropower data from INL (1998) and eGRID (2010), as presented below. To account for county border changes, data are adjusted in later periods to maintain the 1900 county definitions (Hornbeck, 2010).

A natural measure of hydropower potential is capacity installed in hydro plants plus undeveloped capacity. A unique feature of my database is the inclusion of a measure for the undeveloped capacity. It comes from the 1998 U.S. Hydropower Resource Assessment, prepared by the Idaho National Engineering and Environmental Laboratory (INL) for the U.S. Department of Energy (DOE) (Conner, Francfort, and Rinehart, 1998; INL, 1998). A measure of installed capacity comes from the U.S. Environmental Protection Agency's Emissions and Generation Resource Integrated Database (eGRID) for 2007 (eGRID, 2010).

The INL report presents DOE's efforts to produce a more definitive assessment of undeveloped hydropower resources within the U.S. No agency had previously estimated the undeveloped hydropower capacity based on site characteristics, stream flow data, and available hydraulic heads. Initial efforts began in 1989 and information from the last state was received in 1998. State agencies such as departments of dam safety, water resources, environmental quality, fish and game, history, and commerce, contributed information about hydropower resources within their states. The report summarizes and discusses the undeveloped *conventional* hydropower capacity for the 5,677 sites within the country. It does not include the capacity produced by pumped storage sites. However, for conventional hydropower, the resource assessment contains site identification information, geographic coordinates, and crucially the estimated nameplate capacity[8].

[8]Nameplate capacity refers to the intended technical full-load sustained output of a facility.

The eGRID is a comprehensive inventory of the generation and environmental attributes of all power plants in the U.S. Much of the information in this database, including plant opening years, comes from DOE's Annual Electric Generator Report compiled from responses to the EIA-860, a form completed annually by all electric-generating plants. In addition, eGRID includes plant identification information, geographic coordinates, number of generators, primary fuel, plant nameplate capacity, plant annual net generation, and whether the plant is a cogeneration facility.

My sample consists of counties that have either (i) non-cogeneration plants with installed capacity of 100 megawatts or more, generating electricity only through conventional hydropower, or (ii) undeveloped sites with estimated nameplate capacity of 100 megawatts or more. Because often capacity builds up gradually, I assume that a county has hydroelectric dams only when it reaches the 100-megawatt nameplate. I use the same cut-off to determine the year in which a dam is completed. Counties with hydroelectric facilities are my "treated" counties, and those with undeveloped sites are my "control" counties. Here, "undeveloped" means with no dams at all, or with dams for purposes other than power generation (e.g., flood control, irrigation, and navigation).

Figure V displays the sample counties. As we can see clearly, most of them are located in two regions of the country: South (44.8 percent) and West (38.3 percent). Because they have similar hydropower potential, they likely have comparable topography. However, because they are somewhat spread within regions, climate variables (50-year average rainfall and 50-year average temperature for each season of the year) and geographic coordinates (latitude and longitude) are included in the empirical analysis to control for other possible natural advantages.

Table III reports county statistics for my main outcomes (population and employment density) and some hydroelectric-related variables. Among the reported statistics, notice that pre-1950 treated counties have the highest levels of outcomes throughout the twentieth century, followed by post-1950 treated counties, and then by control counties. (Figures A1 and A2, in the Appendix, display outcome trajectories from decade to decade.) Also, observe that most of the hydroelectric dams were constructed from the 1920s to the 1980s, with a boom around the 1950s. In the beginning of the century, they were small, then became larger to tap economies of scale in electricity generation, and finally came to be small again because of environmental concerns. Last, note the increase in hydroelectricity capacity after installation of the first plants. In some cases, hydropower facilities were upgraded; in others, new plants were constructed. Because these changes in installed capacity may affect outcomes directly, I control for them in my empirical analysis.

6 Empirical Framework

In this section, I present my novel empirical approach to obtain impact estimates of hydroelectric dams in the short and long run and my strategy to estimate agglomeration spillovers. My new approach is a two-step procedure that combines synthetic control methods and event-study techniques. In the first step, I use synthetic control analysis to uncover the effect of dams for each treated county separately and, more importantly, to construct counterfactuals, which I refer to as "synthetic control counties". In the second

step, I pool all pairs of treated and synthetic control counties, and run an event-study analysis to find the average treatment effect of dams across all treated counties. My strategy to estimate agglomeration spillovers follows the intuition provided by Figure I in the introduction.

6.1 Estimation by County - Synthetic Control Analysis

Initially, I estimate the impact of hydroelectric dams on population density for each treated county separately. I use synthetic control methods (Abadie and Gardeazabal, 2003; and Abadie, Diamond and Hainmueller, 2010), which basically compare the evolution of population density for a treated county to the evolution of the same outcome for a synthetic control county. The synthetic control county is a weighted average of potential control counties chosen to approximate the treated county in terms of the outcome predictors. The evolution of the outcome for the synthetic control county is an estimate of the counterfactual of what would have been observed for the affected county in the absence of dam installation. Once treated and synthetic control counties have similar outcome behavior over extended periods of time before dam installation, a discrepancy in the outcome variable following installation is interpreted as produced by the dam itself.

To provide a more formal summary of this approach, suppose that there is a sample of $C+1$ counties indexed by c, among which unit $c = 1$ is the treated county and units $c = 2$ to $c = C + 1$ are potential controls. Also, assume a balanced panel which includes a positive number of pre-intervention periods, T_0, as well as a positive number of post-intervention periods, T_1, with $T_0 + T_1 = T$.

Let Y_{ct} be the outcome of unit c at time t. For a post-intervention period t (with $t \geq T_0$), the synthetic control estimator of the effect of dam installation is given by the comparison between the outcome for the treated county and the outcome for the synthetic control at that period:

$$Y_{1t} - \sum_{c=2}^{C+1} w_c^* Y_{ct}.$$

Let $W = (w_2, ..., w_{C+1})'$ be a $(C \times 1)$ vector of positive weights that sum to one. Also, let X_1 be a $(k \times 1)$ vector containing the values of the pre-intervention characteristics of the treated county, and let X_0 be the $(k \times C)$ matrix collecting the values of the same variables for the counties in the "donor pool"[9]. Abadie and Gardeazabal (2003) and Abadie, Diamond and Hainmueller (2010) choose W^* as the value of W that minimizes

$$\|X_1 - X_0 W\|_V = \sqrt{(X_1 - X_0 W)' V (X_1 - X_0 W)},$$

where an optimal choice of V assigns weights to linear combinations of the variables in X_0 and X_1 to minimize the mean squared error (MSE) of the synthetic control estimator.

The matching variables in X_0 and X_1 are meant to be predictors of post-intervention outcomes, which are not themselves affected by the intervention. Notwithstanding, using a linear factor model, Abadie, Diamond, and Hainmueller (2010) argue that if the number of pre-intervention periods in the data is large, the inclusion of pre-intervention outcomes in X_0 and X_1 helps control for unobserved factors affecting the

[9] "Donor pool" is defined as the set of potential control counties out of which the synthetic control unit is constructed.

outcome of interest as well as for heterogeneity of the effect of the observed and unobserved factors. This approach ends up extending the traditional difference-in-differences framework, allowing the effects of unobserved variables on the outcome to vary with time. In my analysis, I use the following matching variables in X_0 and X_1 : (i) pre-dam log of population density up to the year before installation, (ii) dummies for the four regions of the country (Northeast, Midwest, South, and West), (iii) cubic function in latitude and longitude, and (iv) 50-year average rainfall and 50-year average temperature for each season of the year.

A byproduct of the synthetic control estimation is the construction of synthetic control counties. Using the weights that minimize the MSE of the synthetic control estimator, I generate a counterfactual for each treated county in my sample. Each counterfactual, or synthetic control county, represents a weighted average of the counties contained in the donor pool. Hence, it has outcomes and characteristics representing weighted averages of outcomes and characteristics, respectively, of the originally defined control counties. In the end, I obtain a pair of treated and synthetic control counties for each county hosting hydroelectric dams.

6.2 Pooled Estimation - Event-Study Analysis

To provide an average estimate of the impact of hydroelectric dams on population density, with dams built in different decades, I pool all pairs of treated and synthetic control counties, and use an event-study research design (e.g., Jacobson, LaLonde, and Sullivan, 1993; McCrary, 2007; and Kline, 2012). An event-study analysis can recover the dynamics of the impact of those dams in the short and long run, and test whether hydro dams were constructed in response of county-specific trends in population density. I follow Kline's (2012) exposition of such an approach here.

Consider the following econometric model of population density:

$$Y_{ct} = \sum_{y} \beta_y D_{ct}^{y} + \alpha_c + \gamma_{rt} + Z_c^{'}\phi_t + X_{ct}^{'}\lambda + \varepsilon_{ct}, \tag{6}$$

where Y_{ct} is the log of population density in county c in calendar year t, α_c is a county effect, γ_{rt} is a region-by-year fixed effect, Z_c is a vector of time-invariant county characteristics (cubic function in latitude and longitude, and 50-year average rainfall and 50-year average temperature for each season of the year) whose coefficients are allowed to vary in each year, X_{ct} is a vector of time-varying county attributes (cubic function in dam size and in capacity of thermal power plants), and ε_{ct} is an error term that may exhibit arbitrary dependence within a county but is uncorrelated with the other right-hand side variables[10].

The D_{ct}^{y} are a series of event-time dummies that equal one when dam installation is y years away in a county. Formally, we may write

$$D_{ct}^{y} \equiv I[t - e_c = y],$$

where $I[.]$ is an indicator function for the expression in brackets being true, and e_c is the year a dam is installed in county c.

[10]The reason for this exact specification will be clear in section 7, subsection "Specification Issues".

Thus, the β_y coefficients represent the time path of population density relative to the date of dam installation, conditional on observed and unobserved heterogeneity. If dams are randomly assigned to counties, the restriction $\beta_y = 0$ should hold for all $y < 0$. That is, dam installation should not, on average, be preceded by trends in county-specific population density. Because not all of the β's can be identified due to the collinearity of D's and county effects, I normalize $\beta_0 = 0$, so that all post-installation coefficients can be thought of as treatment effects. Lastly, I impose the following endpoint restrictions:

$$\beta_y = \begin{cases} \overline{\beta}, & \text{if } y \geq 80 \\ \underline{\beta}, & \text{if } y \leq -80, \end{cases}$$

which simply state that any dynamics wears off after eighty years. This restriction helps to reduce some of the collinearity between the region-by-year and event-time dummies. As explained in Kline (2012), because the sample is unbalanced in event time, these endpoint coefficients give unequal weight to counties installing hydro dams early or late in the sample. For this reason, I focus the analysis on the event-time coefficients falling within an eighty-year window that are identified off of a nearly balanced panel of counties.

Reweighting/Matching Empirical Strategy

My two-step procedure to obtain impact estimates of hydroelectric dams can be seen as a reweighting/matching strategy to estimate treatment effects that accounts for time-varying unobserved heterogeneity. First, I find a synthetic control unit for each treated county using synthetic control methods. As discussed above, a synthetic control is a weighted average of potential control counties that replicates the counterfactual outcome that the treated county would have experienced in the absence of the treatment (dam installation). Recalling that time-varying unobserved heterogeneity is taken into account in the estimation of the synthetic-control optimal weights (Abadie, Diamond and Hainmueller, 2010), synthetic control counties represent reweighted aggregations of the originally defined control counties that account for time-varying unobserved heterogeneity.

Second, I match each treated county with its corresponding synthetic control to generate the sample with which I run the event-study analysis. Because synthetic controls are objects intrinsically associated with their treated counterparts, I conduct hypothesis testing using standard errors clustered at the case level, where a case is a pair of a treated and its corresponding synthetic control county[11].

6.3 Agglomeration spillovers

Having described my methodology to estimate the effects of hydroelectric dams on population density over a long period of time, I present my strategy to uncover lower bound estimates of agglomeration spillovers. As exemplified in the introduction, my measure of spillovers is the growth in population density in pre-1950 treated counties over and above (i) the growth experienced by them until 1950, which is mostly due to the

[11]In Appendix B, I discuss an alternative approach to this reweighting-matching strategy. I consider the "synthetic propensity score reweighting".

advantage of cheap local hydroelectricity, and (ii) the growth experienced by post-1950 treated counties, which probably results from changes in stock of capital, given the attenuation of the appeal of cheap local hydropower in the second half of the twentieth century. Hence, this measure reflects the dynamics of population growth that might arise when the effects of cheap electricity and direct investment fade away. It represents a lower bound for the agglomeration spillovers of pre-1950 dams because the subtraction of the impact of post-1950 dams eliminates not only the effects of changes in stock of capital, but also any potential agglomeration effects of those dams.

Although my measure of agglomeration economies is easily illustrated in Figure I, it can be less clear when I average it across pre-1950 treated counties because of different dam completion dates. For any number of years y after dam installation, I can express it as

$$\widehat{AS}_y = \widehat{\beta}_y^{CTB1950} - \widehat{G}^{CTB1950until1950} - \widehat{\beta}_y^{CTA1950}, \tag{7}$$

where $\widehat{\beta}_y^{CTB1950}$ is the coefficient of an event-time dummy for Counties Treated Before 1950 ($CTB1950$), $\widehat{G}$ is the estimate of the average growth of population density from the time of dam installation until 1950 for $CTB1950$, and $\widehat{\beta}_y^{CTA1950}$ is the coefficient of an event-time dummy for Counties Treated After 1950 ($CTA1950$).

To accommodate treated counties with hydroelectric plants built in different decades, $\widehat{G}$ is a weighted average of the impact of dams up to forty years after installation of the facilities, depending on the county-specific completion date. That is,

$$\widehat{G}^{CTB1950until1950} = (d_{1910s} * \widehat{\beta}_{40}) + (d_{1920s} * \widehat{\beta}_{30}) + (d_{1930s} * \widehat{\beta}_{20}) + (d_{1940s} * \widehat{\beta}_{10}), \tag{8}$$

where d_{19_0s} is the number of counties with dams built in a specific decade.

7 Results

In this section, I present two sets of results. First, I discuss the effects of hydro dams case by case for a representative group of treated counties. These are my county-specific estimates, found through synthetic control methods. Then, I discuss the average impact of hydroelectric facilities for all treated counties in my sample. These are my pooled estimates, arising from event-study analyses. The evidence of agglomeration spillovers is examined within this last subsection.

7.1 Synthetic Control Approach: County-Specific Estimates

Pre-1950 Dams

Emblematic Case. In Figure I, I have introduced the synthetic control estimator for Blount County, Tennessee, where Calderwood Dam was installed in the 1930s. In that Figure, population density grows rapidly from dam completion until 1950, and then slows down afterwards. Had the dam not been constructed, there would have been just slight growth after 1930s, probably due to the Great Depression.

Therefore, the impact of the hydropower plant until 1950 was approximately 0.40 log points (49 percent). From 1950 to 2000, the effect was roughly 0.50 log points (65 percent). This second number is also an approximate measure of agglomeration spillovers for Blount County, since the advantage of cheap local hydroelectricity weakened around 1950.

No Agglomeration Economies. Figure VI, panel A, displays a case of strong short-run effect of hydroelectric facilities, but almost no agglomeration economies. From completion in the 1930s until 1950, Norris Dam induced growth of approximately 1.17 log points (223 percent) in population density of Anderson County, Tennessee, relative to the counterfactual. Despite this enormous short-run impact, Anderson County grew only 0.12 log points (13 percent) afterward. Thus, the dam generated virtually no agglomeration externalities once cheap electricity spread around the country.

Reversion. A disturbing result of a public investment is illustrated by Hawks Nest Dam, installed in the 1930s in Fayette County, West Virginia. Following dam completion, the county experienced a growth of almost 0.15 log points (16 percent) in population density relative to its counterfactual, as shown in Figure VI, panel B. Nevertheless, in the second half of the twentieth century, when the appeal of cheap local hydroelectricity became attenuated, that trend reversed and the county had a drop of 0.45 log points (36 percent) in population density. This is an emblematic case of lack of path dependence. Once the advantage of cheap local power reduces, and capital depreciates, people fly away.

Constant Growth. An interesting outcome is the one exemplified by Travis County, Texas, which had Mansfield Dam (formerly Marshall Ford Dam) constructed in the 1940s. As displayed in Figure VII, panel A, once the dam was built, the county embarked on a stable path of population density growth, with a rate that remained constant until the end of the twentieth century. In the first decade, Travis County expanded nearly 0.39 log points (48 percent) relative to its counterfactual. From 1950 to 2000, that trend did not become flatter, and the county grew approximately 1.07 log points (191 percent). From a local perspective, this is what every policymaker would like to witness. After the initial push, agglomeration spillovers kicked in vigorously, producing a sustainable dynamic of growth.

Indifference. An unattractive situation from a policymaking point of view is the one portrayed by Fort Loudoun Dam, built in the 1940s in Loudoun County, Tennessee. The trajectories of population density of treated and synthetic control counties, shown in Figure VII, panel B, do not differ significantly after the installation of the plant. Apparently, the county would have grown steadily even without the hydropower facilities.

Timing of Attenuation of Cheap Local Power Advantage

In the emblematic example of Figure I, the path of population density flattens out in 1950. As discussed in the historical section, this might be a good estimate of the period in which the appeal of cheap local hydroelectricity started to fade away. However, most of the high-voltage transmission lines were constructed in the 1950s and 1960s, and thermal efficiency increased gradually from the 1940s to the 1960s, as displayed in Figure III. So it is possible that some counties experienced the compression in growth rates earlier or later than 1950. Indeed, Figure VIII presents two emblematic cases of such possibilities. Haywood County, North Carolina, for instance, witnesses the flattening in 1940, just a

decade after Walters Dams was completed. San Bernardino County, California, on the other hand, sees its growth rates in population density reduce only in 1960, two decades after Parker Dam had been completed. Therefore, the use of 1950 as the turning point of the attenuation of the advantage of cheap local hydroelectricity must be seen as an approximation only.

Post-1950 Dams

Figures IX and X display the dynamics of population density for four counties where hydroelectric plants were installed in the second half of the twentieth century. These cases illustrate the typical effects that I find in my analysis with post-1950 dams: moderately positive, small but positive, nonexistent, and somewhat negative.

First, consider the case of Lewis County, Washington, which had both Mayfield Dam and Mosyrock Dam constructed in the 1960s (Figure IX, Panel A). Notice that the initial jump in population density, representative of the impact of pre-1950 dams, was relatively small here: around 0.16 log points (17 percent) relative to the counterfactual, after a decade. This small effect seems to reinforce hydroelectricity as an advantageous local attribute in the first half of the century. Moreover, it indicates that direct investment might also play a role in generating growth following dam completion. The impact grew gradually to nearly 0.35 log points (42 percent) in 2000. It is quite possible that some agglomeration spillovers are present here, over and above the effect of changes in stock of capital, but I cannot separate them out.

In the second case (Figure IX, Panel B), the initial impact is even smaller, and potential agglomeration externalities look negligible. Indeed, after installation of Keowee Dam in the 1970s, Pickens County, South Carolina, grew only approximately 0.09 log points (10 percent) in the first decade. Subsequently, the growth was minimal. In 2000, three decades after dam completion, the impact was just under 0.17 log points (18 percent). Therefore, agglomeration economies seem insignificant.

The third case, of Detroit Dam (Figure X, Panel A), concluded in the 1950s in Marion County, Oregon, shows no effect at all. It appears that the county would have grown as much as it did without the hydro plant. Lastly, the fourth case (Figure X, Panel B) depicts a decrease in population density. After completion of Roanoke Rapids Dam, in the 1950s, and Gaston Dam, in the 1960s, people started to leave Halifax County, North Carolina. Part of such population decline might be due to displacement, but it seems odd that the county had not recovered even its pre-dam level of population density as of 2000.

Distribution and Heterogeneity of County-Specific Dam Impacts

The cases mentioned above illustrate reasonably well typical dynamics of population density growth in my sample. To provide a summary of all counties, I plot the distribution of pre-1950 dam effects in Figure XI, and of post-1950 dams in Figure XII. When we examine each column of Figure XI, we can see clearly that the distribution of impacts of hydroelectric plants is shifting to the right. This movement happens despite the attenuation of the appeal of cheap local hydroelectricity and the fading of the direct investment effect. It is then quite plausible that agglomeration spillovers kick in at some point, and give rise to such a path dependence.

Figure XII, on the other hand, portrays a rather different story. Each column shows a distribution of dam effects somewhat inert around zero even after fifty years. This observation reinforces the idea of cheap local hydroelectricity as a driving force of concentration of economic activity and subsequent agglomeration externalities.

Besides allowing me to plot the distribution of effects, county-specific estimates give me the possibility of analyzing the heterogeneity of dam impact in a very simple and direct way. All that is necessary is to run panel data regressions of dam effects on characteristics of dams or locations hosting them, for example, controlling for years relative to dam completion. Doing so, I find that pre-1950 dam effects are, on average, nearly 6 percent stronger (coefficient: 0.0575; s.e.: 0.0210) when dam size at completion is a hundred megawatts larger. In my sample, dam size at completion ranges from one to approximately nine hundred megawatts for pre-1950 dams. On the other hand, I do not find any statistically significant heterogeneity regarding population density a decade before dam completion (coefficient: -0.0499; s.e.: 0.0385).

7.2 Event-Study Analysis: Pooled Estimates

Short- and Long-Run Impact of Hydro Dams

When I pool all pairs of treated and synthetic control counties and estimate equation (6), I find quite interesting results[12]. First, the timing of dam installation appears to be crucial to the pattern of observed effects. Hosting hydropower facilities in the first half versus the second half of the twentieth century means enjoying a period of great prosperity versus no detectable changes, as depicted in Figure XIII and Table IV. Indeed, counties with pre-1950 dams experience average increases in population density of approximately 0.41 log points (51 percent) after 30 years[13], and 0.87 log points (139 percent) after 60 years, as shown in Table IV. On the other hand, counties with post-1950 dams experience no statistically significant effects[14].

[12]I discuss my results using the group of synthetic control counties. As explained in the methodology section, because both observed and unobserved heterogeneity are taken into account in the estimation of synthetic controls, and are allowed to vary flexibly over time, the comparison between treated counties and these controls might provide a more accurate estimate of the treatment effect. Nevertheless, for comparison purpose, I present results using the originally defined control counties as well. Both sets of estimates are shown side-by-side in Tables and Figures. Notice that the estimates using the original control counties are smaller than those using synthetic controls. This might indicate that hydroelectric dams are allocated to counties partially in response to negative population shocks. A similar issue associated with targetting poor places for infrastructure projects is discussed in Duranton and Turner (2012) for the case of roads.

[13]Duflo and Pande's (2007) pioneer study of the impact of large irrigation dams on agricultural production and poverty rates in India provides estimates over a time span of only three decades. Notwithstanding, their work produces credible estimates of the impact of irrigation dams. Instead of just comparing outcomes of districts with and without irrigation dams, they use variation in dam construction induced by differences in river gradient across districts within Indian states to obtain instrumental variable estimates.

[14]One may argue that focusing the analysis on counties hosting hydropower facilities instead of places accessing the power grid might underestimate the impact of hydro dams. Through the grid, electricity might have been available in counties originally defined as controls. Although this is a possibility, and an example would be Hoover Dam supplying power more than 200 miles away to Los Angeles starting in 1936, a visual comparison between the map of transmission lines in 1923 and 1962 (available upon request) show that the transmission network was very local in the first half of the twentieth century. Furthermore, the estimated effects of hydro dams in neighboring counties, presented below, suggest a much smaller role there.

Second, the enormous magnitude of the impact of hydroelectric plants seems somewhat remarkable. On average, county sizes more than double five decades after dam completion. Actually, population density grows nearly 0.79 log points (120 percent) after 50 years of pre-1950 dam installation. Third, the difference in short- and long-term effects is revealing. The 30-year estimate presented above is less than half its 60-year counterpart. This suggests that the assessment of large infrastructure projects does require understanding of long-run effects, as advocated by Kline (2010). The comprehensive report of the World Comission on Dams, in 2000, does provide some evidence of long-term effects of a few dams around the world, but my study seems to provide the first systematic evaluation of impact of dams in the long run. Fourth, and last, the strength of economic growth in the long run looks rather surprising. The annualized growth rate of population density is roughly 1.6 percent in the first 40 years, and still 1 percent in the following 40 years. This might indicate the presence of either slow adjustment of capital stock to shocks or strong agglomeration economies. Below, I provide evidence that agglomeration spillovers may explain a great part of the late growth in population density.

Specification Issues

The first specification that I estimate in this study is equation (6), but without X_{ct}. Such specification, which I refer to as the "basic specification", includes event-time dummies, county effects, region-by-year fixed effects, and time-invariant county characteristics (cubic function in latitude and longitude, and 50-year average rainfall and 50-year average temperature for each season of the year) interacted with year effects. Because this set of covariates does not seem enough to eliminate all pre-treatment trends, as evident in Table V and in Figure XIV, I add controls for dam size to the basic specification. As explained in section 5, dams above the 100-megawatt cutoff still differ in size, and might expand over time. When I include a cubic function in dam capacity, the post-treatment estimates remain unchanged, and the coefficients of such controls are statistically insignificant. What seems to drive the effects of hydro dams is timing, not size. However, the inclusion of controls for dam size does remove any pre-treatment trends.

Next, I add controls for the size of thermal power plants present in my sample counties. My concern is that counties with lower hydropower potential might respond to this natural constraint by building fossil fuel and/or nuclear power plants. In that case, the impact of hydro dams would be underestimated. My estimates do increase when I include such controls, but just a little bit.

Also, I add controls for the interaction of year effects with three county-specific measures of market access in 1890: mileage of railroad tracks, distance to closest waterway, and log of market access as estimated by Donaldson and Hornbeck (2012). One could argue that my control group, while appealing in some respects, does not necessarily have the same natural advantages simply because they have the same hydropower potential. A concern will be that if a county has good hydro potential and is near a river or a rail line, the value of installing hydroelectric facilities there would be higher because then local manufacturers could easily ship out their products. That could lead to bribes and political deals that would get the dams built in those places, and not in isolated places where, even if a dam was built, it would be hard to get manufacturers to move in. In such case, the impact of hydro dams would be overestimated. After controlling for pre-dam proxies of market access, my post-dam estimates do decrease,

but not significatively. However, pre-treatment trends seem to be slightly more pronounced than in the previous specification.

Lastly, I check how state-by-year fixed effects versus region-by-year fixed effects would change my findings. One can argue that most energy policies are made at the state level, so the impact of hydro dams may reflect only state-specific effects. As evident in the chart with original control counties, this is not the case here. (I could not run this specification with the synthetic control group because there were fewer observations than parameters to be estimated.) In the end, the specification that I use to report my findings throughout the paper - my main specification: equation (6) in the text - is the one with region-by-year fixed effects, controls for dam size, and controls for capacity of thermal power plants.

Placebo Tests

In order to check whether my research design is capturing the impact of hydroelectric dams rather than the effect of some unobserved intervention, I run placebo tests. The estimated impact of dams would be undermined if I obtained effects of similar or even greater magnitudes in cases where/when the intervention did not take place (Heckman and Hotz, 1989; Abadie and Gardeazabal, 2003; Bertrand, Duflo, and Mullainathan, 2004; and Abadie, Diamond and Hainmueller, 2010). Therefore, I pool all counties and select randomly treated before 1950, treated after 1950, and controls, and redo my analysis fifty times. For artificially treated counties, I also set the year of dam completion at random.

The estimated effects of artificial treatments are shown in Figures XV and XVI. For comparison purposes, I plot the impact of the real treatment as well. As we can see clearly, the effect of pre-1950 hydro plants in counties that actually hosted them are distinguished from the placebo effects. This effect is higher than any other effect in Figure XV. On the other hand, the impact of post-1950 dams is completely mixed with placebo impacts. These pieces of evidence reinforce the importance of pre-1950 hydroelectric dams in driving my results.

Agglomeration Spillovers

The effects of hydroelectric dams long after dam completion suggest that economic growth may be amplified by additional forces such as agglomeration spillovers. The installation of large hydropower facilities combined with the local appeal of hydroelectricity in the first half of the twentieth century might have induced a process of cumulative causation in counties hosting pre-1950 dams. The initial investments and the resulting cheap electricity might have set in motion a chain reaction with multiple changes in population density.

As illustrated by Figure I, and described in the methodology, in order to provide evidence of agglomeration economies I make two simplifying assumptions. First, I attribute any growth in population density until 1950 to the advantage of cheap local hydroelectricity. Second, I suppose that the effects of post-1950 dams are driven mostly by changes in the stock of capital associated with dam installation, and, potentially, small agglomeration externalities. Hence, a lower bound of agglomeration economies can be found by subtracting any impact of post-1950 dams from the post-1950 population density growth of counties hosting pre-1950 dams.

Figure XVII displays the impact of pre- and post-1950 dams in the top charts, and the difference of those effects in the bottom charts[15]. The dotted line in the bottom charts represents the growth of population density from the completion of pre-1950 dams until 1950. Therefore, the lower bound of agglomeration spillovers (or the magnitude of "pure" agglomeration externalities) is up to 45 percent (0.37 log points) five decades after dam installation. The advantage of cheap local hydroelectricity might explain most of the concentration of economic activity in the first 20 years after dam completion. Subsequently, spillovers seem to kick in, and may account for roughly 47 percent of the 50-year impact of hydroelectric dams.

My lower bound for agglomeration economies seems to be very large compared to other estimates in the literature. For instance, Greenstone, Hornbeck and Moretti (2010) find that, five years after the opening of a Million Dollar Plant, incumbent plants' TFP is 12 percent higher in hosting counties. Thus, my estimate appears to be almost four times larger than theirs. Most of the difference, however, is due to the time horizons of our analyses. Theirs is a short-run estimate, mine is a long-run one. When I consider the lower bound nearly a decade after spillovers kick in, my estimate is around 11.5 percent (0.11 log points), which is close enough to theirs. (My estimate then becomes 34 percent (0.29 log points) after two decades of prominence of spillovers, and 45 percent (0.37 log points) after three decades.)

Sectoral Decomposition

Although I have described my results in terms of population density, I find similar patterns for employment density. As depicted in Figure XVIII, and reported in Tables IV and VI, the effects of hydroelectric dams on population and employment are closely related, but there is a lot of heterogeneity in the impact on different sectors of the economy. In fact, the sectoral decomposition in Table VII and Figure XVIII shows construction and trade (trade defined as wholesale plus retail) with the weakest effects, manufacturing and agriculture with intermediate effects, and other sectors with the strongest effects. "Other sectors" consists basically of the nontradable sector of local economies. It includes employment in industries that provide local goods and services such as real estate, cleaning services, legal services, medical services, and personal services.

The patterns of sectoral employment after dam installation seem to make sense. In many hydro projects, dams serve multiple purposes such as navigation, flood control, irrigation and hydropower. So, it is expected that agriculture benefits from them, especially in the West, where agricultural production depends heavily on irrigation. In fact, the development of hydro projects may lead to the adoption of central pivot irrigation, which should increase productivity/crop yields and, as a consequence, land values[16]. It is also not surprising that employment in construction goes up when dams are being built, and down afterwards. This suggests that the agglomeration spillovers presented above do not seem to arise from construction sites. The dynamics of employment in trade is the only puzzling result. In general, it follows the pattern of other industries in the nontradable sector.

The strong long-run impact on the nontradable sector might reflect a local multiplier induced by

[15]Differences are shown in columns 5 and 6 of Table IV.

[16]Table C.I presents suggestive evidence of farmland value appreciation after the construction of hydroelectric dams.

agriculture and manufacturing growth, which in turn might have been driven by cheap hydroelectricity. Moretti (2010) finds that, for each additional job in the tradable sector in a given city, 1.6 jobs are created in the nontradable sector in the same city. My results indicate a local multiplier of 1.5 five decades after dam completion, very close to Moretti's estimate. Therefore, it is likely that the main mechanism generating agglomeration spillovers is the concentration of agricultural activity and manufacturing next to hydropower plants. The importance of agriculture in this setting is somewhat remarkable. Indeed, Hornbeck and Keskin (2012) provide evidence that agricultural production does not appear to generate local economic spillovers.

Neighboring Counties

Because hydroelectricity seems to be the main driving force of agglomeration, and power could be transmitted to neighboring areas at low cost, even in the first half of the twentieth century, we should see some impact of hydroelectric dams in contiguous counties as well. Figure XIX displays the short- and long-term effects of dams in contiguous counties. As expected, they look strong enough: they represent more than half of the effects in the treated counties 10 years after dam completion, nearly a third 50 years after, and roughly a fifth 80 years after. Detailed estimates are presented in Table VIII.

Notice, however, that there is a steep upward trend before dam completion. This might be due to two things. First, my definition of completion refers to the point at which a dam reaches 100 megawatts of capacity. So, it is possible to observe some effects in decades before "completion", especially if hydropower facilities are expanded step-by-step. Indeed, it is common to see a generator being installed in a year, and the next ones years later. Second, during dam construction population might be temporarily or permanently displaced, and then economic activity might flourish in neighboring locations. Upon dam completion, though, people might return to their original counties, flattening the trend.

Smaller Hydro Dams (30-100 Megawatts)

The focus of my analysis is large hydroelectric dams: 100 megawatts or more of capacity. The goal of my study is to estimate the impact of new, independent hydropower plants on the economic activity of local economies. Plants smaller than 100 megawatts were excluded because they tend to be built in connection with existing industrial facilities. It is interesting, however, to check how the results change when we use these smaller dams as treatment. It can give us some insights about potential nonlinearities of the agglomeration effects. If larger dams induce proportionately larger effects than smaller dams, one might infer that larger interventions are more efficient than smaller ones from the point of view of a social planner.

Table IX reports and Figure XX plots the results of my main analysis along with the effects of smaller dams. I estimate the impact of smaller plants using two different groups of controls: (i) original control counties, and (ii) counties with hydropower potential between 30 and 100 megawatts but with no hydroelectric facilities. The analysis with the first control group reveals effects proportionally larger than the difference in average dam size. The average size of smaller dams is a quarter of the average size of larger dams, but the impact of smaller dams is up to half of the impact of larger dams. This suggests the

presence of nonlinear agglomeration effects but with direction opposite to that expected. Taken seriously, this would imply diminishing agglomeration returns to investments in hydropower plants. Nevertheless, when I use the second control group, which makes treated and control counties more comparable in the estimation, I find effects proportional to the difference in average dam size. Therefore, larger dams appear not to induce larger agglomeration effects. If anything, potential nonlinearities indicate dispersion forces prevailing against agglomeration.

Tennessee Valley Authority (TVA) versus Non-TVA Counties

As we can see in Figure V, many counties treated before 1950 are located in the TVA region. The TVA was one of the most ambitious place-based economic development policies in the history of the U.S. It was created in the 1930s to boost economic activity in the region by investing in large scale infrastructure programs, particularly hydroelectric dams, and in an extensive network of new roads, canals, and flood control systems. Kline and Moretti (2014) conduct an evaluation of the dynamic effects of the TVA on local economies. They find that it led to short-run gains in agricultural employment that were eventually reversed, but impacts on manufacturing employment that continued to intensify well after the program's subsidies had lapsed.

Given that my results mirror the amplifying effects of the TVA, and that in my sample a third of counties with pre-1950 dams are situated in the TVA region, one can argue that my findings are driven by the TVA. One can attribute most of my results to infrastructure projects not related to hydroelectricity. However, as shown in Table X and in Figure XXI, when I estimate the effects of hydro dams without TVA counties, the pattern is very similar to the one with the counties all together. Hence, hydropower may indeed be the main force behind my findings.

All versus No Control Counties with Environmental Regulations

One concern with my impact estimates of hydro dams is that my originally defined control counties may be subject to environmental regulations. If a county has sites suitable for hydroelectric dams, but construction is not allowed by law or by pressure of civil society, population changes in that county may be severely constrained. As a consequence, my estimates would be overestimated[17].

It is true that most of my originally defined control counties - 78 percent - have some kind of land regulation aimed at preserving wilderness and wildlife. However, only two out of 55 would have hydro projects completely forbidden under hydroelectric licensing rules. Being extremely conservative, and eliminating all control counties with some environmental regulations from my analysis, I find no significative changes in my estimates, as shown by Figure XXII and Table XI. The number of originally defined control counties, however, decreases considerably.

To make sure that my results are not driven by environmental constraints in my originally defined control counties, I redefine my control group to include every county with hydropower potential above ten megawatts but with no land regulations. Although geographical features of treated and control counties

[17]Even more troublesome would be the fact that regulations have become stricter over time. Nevertheless, this has already been considered in the analysis: my empirical strategy takes into account time-varying unobserved heterogeneity.

might be less comparable than before, when both groups had 100 megawatts or more of capacity, this procedure increases my sample size substantially - from 55 to 192 control counties. Nevertheless, the new estimates are still within the 95 percent confidence of my main analysis.

Hydroelectric Dams as Big Push?

In order to shed light on the sustainability of the higher steady states reached with dam installation, I construct an empirical agglomeration function. Basically, I compute decadal increments of log population density after treatment for counties treated before 1950, and plot a smoothed version of them against time relative to dam completion. The increments are derived from my short- and long-run estimates, and the smoothing is done with locally weighted regressions (lowess). Figure XXIII displays the shape of the resulting function: concave until approximately 60 years, and positive flat afterwards. This implies that the rate of population growth decreases as investments depreciate, but not indefinitely. At a certain point, the declines in that rate stop, and population continues growing at a smaller constant rate. Given the high degree of agglomeration spillovers found previously, it is not surprising that local economies tend to keep their higher/better steady states reached with the help/push of large hydro dams.

8 Concluding Remarks

How much of the geographic clustering of economic activity is attributable to agglomeration spillovers as opposed to natural advantages? I present evidence on this question using data on the long-run effects of large scale hydroelectric dams built in the U.S. over the twentieth century, obtained through a unique comparison between counties with or without dams but with similar hydropower potential. Using a novel combination of synthetic control methods and event-study techniques, I show that, on average, counties with dams built before 1950 have population density increased by approximately 51 percent after 30 years, and 139 percent after 60 years, indicating substantially different short- and long-term effects. This suggests that the assessment of large infrastructure projects does require understanding of long-run effects, as advocated by Kline (2010). On the other hand, counties with dams built after 1950 have no statistically significant effects. I argue that the large difference in the impact of pre- and post-1950 hydro dams can be accounted for by the attenuation of the advantage of cheap electricity in the second half of the twentieth century. Until mid-century, the availability of cheap local power from hydroelectric dams conveyed an important advantage that attracted industry and population. By the 1950s, however, these advantages were weakened by improvements in the efficiency of thermal power generation and the advent of high tension transmission lines.

By using a unique sample of counties with or without hydroelectric dams but with similar hydropower potential, as determined by a team of engineers at the request of the U.S. Department of Energy, I hold natural advantages constant and provide evidence that the persistence and amplification of the dam effects in the long-term is due in great part to agglomeration spillovers. In fact, my lower bound of agglomeration spillovers is up to 45 percent five decades after dam construction, representing almost half of the full effect of hydro dams over the same time span. Interestingly, my short-run estimate of agglomeration externalities

is very close to that of Greenstone, Hornbeck and Moretti's (2010). My lower bound nearly a decade after spillovers kick in is around 11.5 percent, while their estimate five years after the opening of a Million Dollar Plant is 12 percent.

I also find that the estimated short- and long-run effects are highly robust to alternative procedures for selecting synthetic controls, to controls for confounding factors such as proximity to transportation networks, and to alternative sample restrictions, such as dropping dams built by the Tennessee Valley Authority or removing control counties with environmental regulations. I also provide evidence of small local agglomeration effects from smaller dam projects, and small spillovers to nearby locations from large dams.

This study opens up the possibility of some other research projects. First, a similar evaluation of short- and long-run impacts can be done with flood control dams. Because the U.S. Army Corps of Engineers has flood maps available for most of the country, a control group can be constructed in the same spirit as this paper. Second, further investigation on the shape of the agglomeration function may be feasible with my database. Given that the eGRID has information of hydro dams of all sizes, and the INL report contains hydropower potential of all possible capacities as well, variation in dam size could be used to identify potential nonlinearities of that function. These are interesting questions for future research.

References

- Abadie, Alberto, and Gardeazabal, Javier. (2003). The Economic Costs of Conflict: A Case Study of the Basque Country, *American Economic Review*, 93(1): 112-132.

- Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program, *Journal of the American Statistical Association*, 105(490): 493-505.

- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. (2004). How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics*, 119(1): 249-275.

- Billington, David P., Donald C. Jackson, and Martin V. Melosi. (2005). *The History of Large Federal Dams: Planning, Design, and Construction in the Era of Big Dams*. Denver: U.S. Department of the Interior/ Bureau of Reclamation.

- Blanchard, Olivier J., and Lawrence F. Katz. (1992). Regional Evolutions, *Brookings Papers on Economic Activity*, 23(1): 1-76.

- Bleakley, Hoyt, and Jeffrey Lin. (2010). Portage and Path Dependence, *Quarterly Journal of Economics*, 127(2): 587-644.

- BPA - Bonneville Power Administration. (n.d.). *History*. Portland, OR: BPA, bpa.gov/news/AboutUs/Page

- Brakman, Steven, Harry Garretsen, and Marc Schramm. (2004). The Strategic Bombing of German Cities during WWII and its Impact on City Growth, *Journal of Economic Geography*, 4(2): 201-218.

- Brown, Matthew H., and Richard P. Sedano. (2004). *Electricity Transmission: A Primer*. Denver/Washington D.C.: National Council on Electricity Policy.

- Bureau of the Census. (1949). *Census of Manufactures 1947: Fuels and Electric Energy Consumed*. Washington, DC: U.S. Department of Commerce.

- Cintrón, Lizette (ed.). (1995). *Historical Statistics of the Electric Utility Industry through 1992*. Washington D.C.: Edison Electric Institute.

- Conner, Alison M., James E. Francfort, and Ben N. Rinehart. (1998). *U.S. Hydropower Resource Assessment - Final Report*. Idaho Falls, ID: Idaho National Engineering and Environmental Laboratory (INL), hydropower.inel.gov/resourceassessment.

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

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

- Davis, Steven J., Cheryl Grim, John Haltiwanger, and Mary Streitwieser. (2013). Electricity Unit Value Prices and Purchase Quantities: U.S. Manufacturing Plants, 1963-2000, *Review of Economics and Statistics*, 95(4): 1150-1165.

- Devine, Warren D. (1983). From Shafts to Wires: Historical Perspective on Electrification, *Journal of Economic History*, 43(2): 347-372.

- Donaldson, Dave, and Richard Hornbeck. (2012). Railroads and American Economic Growth: A "Market Access" Approach, *Working Paper - Harvard University*.

- DuBoff, Richard B. (1979). *Electric Power in American Manufacturing, 1889-1958*, University of Pennsylvania Dissertation, (1964), New York: Arno Press.

- Duflo, Esther, and Rohini Pande. (2007). Dams, *Quarterly Journal of Economics*, 122(2): 601-646.

- Duranton, Gilles, and Henry G. Overman. (2005). Testing for Localization Using Microgeographic Data, *Review of Economic Studies*, 72(4): 1077-1106.

- Duranton, Gilles, and Henry G. Overman. (2008). Exploring the Detailed Location Patterns of U.K. Manufacturing Industries Using Microgeographic Data, *Journal of Regional Science*, 48(1): 213-43.

- Duranton, Gilles, and Matthew A. Turner. (2012). Urban Growth and Transportation, *Review of Economic Studies*, 79(4): 1407-1440.

- EIA - U.S. Energy Information Administration. (n.d.). *History of the U.S. Electric Power Industry, 1882-1991.* Washington, DC: U.S. Department of Energy, eia.gov/cneaf/electricity/page/electric_kid/append_a

- Ellison, Glenn, and Edward L. Glaeser. (1997). Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach, *Journal of Political Economy*, 105(5): 889-927.

- Ellison, Glenn, and Edward L. Glaeser. (1999). The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration? *American Economic Review Papers and Proceedings*, 89(2): 311-316.

- Ellison, Glenn, Edward L. Glaeser, and William R. Kerr. (2010). What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns, *American Economic Review*, 100(3): 1195-1213.

- Emissions and Generation Resource Integrated Database (eGRID). (2010). *eGRID2010: Version 1.1.* Washington, DC: U.S. Environmental Protection Agency, epa.gov/cleanenergy/energy-resources/egrid.

- Greenstone, Michael, Hornbeck, Richard, and Enrico Moretti. (2010). Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Opnenings, *Journal of Political Economy*, 118(3): 536-598.

- Greenstone, Michael, John A. List, and Chad Syverson. (2012). The Effects of Environmental Regulation on the Competitiveness of U.S. Manufacturing, *Working Paper - University of Chicago.*

- Haines, Michael R., and Inter-university Consortium for Political and Social Research (ICPSR). (2010). *Historical, Demographic, Economic, and Social Data: The United States, 1790-2002.* Ann Arbor, MI: Inter-university Consortium for Political and Social Research, icpsr.org.

- Heckman, James J., and V. Joseph Hotz. (1989). Choosing Among Alternatives Nonexperimental Methods for Estimating the Impact of Social Programs, *Journal of the American Statistical Association*, 84(408): 862-874.

- Hornbeck, Richard. (2010). Barbed Wire: Property Rights and Agricultural Development, *Quarterly Journal of Economics*, 125(2): 767-810.

- Hornbeck, Richard. (2012). The Enduring Impact of the American Dust Bowl: Short- and Long-Run Adjustments to Environmental Catastrophe, *American Economic Review*, 102(4): 1477-1507.

- Hornbeck, Richard, and Pinar Keskin. (2012). Does Agriculture Generate Local Economic Spillovers? Short-run and Long-run Evidence from the Ogallala Aquifer, *NBER Working Paper* No. 18416.

- Idaho National Engineering and Environmental Laboratory (INL). (1998). *U.S. Hydropower Resource Assessment.* Idaho Falls, ID: hydropower.inel.gov/resourceassessment.

- Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan. (1993). Earnings Losses of Displaced Workers, *American Economic Review*, 83(4): 685-709.

- Jollata, Pat. (2004). *Images of America: Downtown Vancouver*. Arcadia Publishing.

- Kitchens, Carl T. (2012). The Role of Publicly Provided Electricity in Economic Development: The Experience of the Tennessee Valley Authority 1929-1955, *Working Paper - University of Mississippi*.

- Kline, Patrick. (2010). Place Based Policies, Heterogeneity, and Agglomeration, *American Economic Review Papers and Proceedings*, 100(2): 383-387.

- Kline, Patrick. (2012). The Impact of Juvenile Curfew Laws on Arrests of Youth and Adults, *American Law and Economics Review*, 14(1): 44-67.

- Kline, Patrick, and Enrico Moretti. (2014). Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority, *Quarterly Journal of Economics*, 129(1): 275-331.

- Mansur, Erin T., and Matthew E. Kahn. (2012). Do Local Energy Prices and Regulation Affect the Geographic Concentration of Employment? A Border Pairs Approach, *Working Paper - Dartmouth College*.

- McClary, Daryl C. (2008). Longview - Thumbnail History, *HistoryLink.org Essay 8560*, historylink.org/index.cfm?DisplayPage=output.cfm&file_id=8560.

- McCrary, Justin. (2007). The Effect of Court Ordered Hiring Quotas on the Composition and Quality of Police, *American Economic Review*, 97(1): 318-53.

- McNerneya, James, J. Doyne Farmer, and Jessika E. Trancik. (2011). Historical Costs of Coal-fired Electricity and Implications for the Future, *Energy Policy*, 39(6): 3042-3054.

- Miguel, Edward, and Gérard Roland. (2011). The Long-Run Impact of Bombing Vietnam, *Journal of Development Economics*, 96(1): 1-15.

- Minnesota Population Center. (2011). *National Historical Geographic Information System (NHGIS): Version 2.0*. Minneapolis, MN: University of Minnesota, nhgis.org.

- Moretti, Enrico. (2011). Local Labor Markets, in David Card and Orley Ashenfelter (eds.), *Handbook of Labor Economics*, Volume 4b. New York: Elsevier, Chapter 14, pp. 1237-1313.

- Redding, Stephen J., and Daniel M. Sturm. (2008). The Costs of Remoteness: Evidence from German Division and Reunification, *American Economic Review*, 98(5): 1766-97.

- Redding, Stephen J., Daniel M. Sturm, and Nikolaus Wolf. (2011). History and Industry Location: Evidence from German Airports, *Review of Economics and Statistics*, 93(3): 814-831.

- Rosenthal, Stuart S. and William C. Strange. (2004). Evidence on the Nature and Sources of Agglomeration Economies, in Henderson, J. Vernon. and Jacques François Thisse (eds.), *Handbook of Regional and Urban Economics*, Volume 4. New York: Elsevier, Chapter 49, pp. 2119-2171.

- Schlenker, Wolfram, and Michael J. Roberts. (2009). *Nonlinear Temperature Effects indicate Severe Damages to U.S. Crop Yields under Climate Change*, Proceedings of the National Academy of Sciences, 106(37): 15594-15598.

- Schramm, Gunter. (1969). The Effects of Low-Cost Hydro Power on Industrial Location, *Canadian Journal of Economics*, 2(2): 210-229.

- U.S. Department of the Interior. (2011). *The National Parks: Index 2009-2011.* Washington, DC: nps.gov/history/history/online_books/nps/index2009_11.pdf.

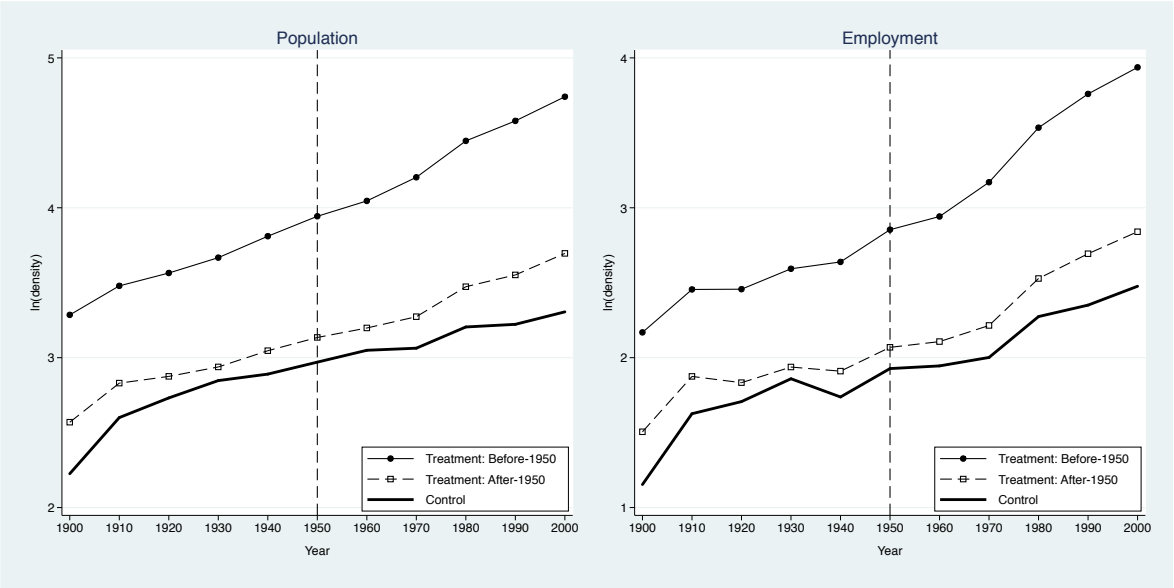
- Voller, Bob. (2010). Early Aluminum Production in the Pacific Northwest, *IEEE Industry Applications Magazine*, 16(3): 10-12.

- World Comission on Dams. (2000). *Dams and Development: A New Framework for Decision-Making.* London and Sterling, VA: Earthscan Publications Ltd.

Appendix

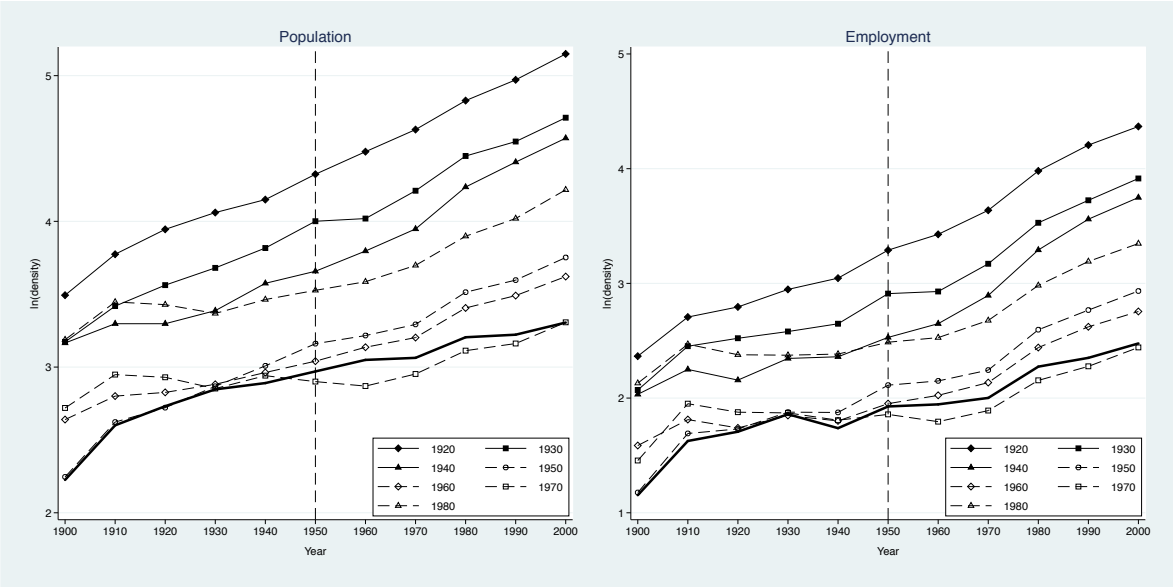
Appendix A: Trends of Population and Employment

Figure A.I: Log Density - 1900-2000



Notes: This Figure displays time series of log density of population (panel A) and employment (panel B) throughout the twentieth century. The dashed vertical line simply separates the twentieth century into pre-1950 and post-1950. The cutoff 1950 is my assumed turning point for the attenuation of the advantage of cheap local hydroelectricity. In each panel, the solid line with solid circles depicts the dynamics for counties that have hydroelectric dams built before 1950. The dashed line with hollow squares depicts the dynamics for counties that have dams built after 1950. Lastly, the thick solid line shows the dynamics of log density for counties in the control group.

Figure A.II: Log Density by Decade Treated - 1900-2000



Notes: This Figure displays time series of log density of population (panel A) and employment (panel B) throughout the twentieth century. The dashed vertical line simply separates the twentieth century into pre-1950 and post-1950. The cutoff 1950 is my assumed turning point for the attenuation of the advantage of cheap local hydroelectricity. In each panel, the solid lines with solid symbols (circles, squares, and triangles) depict the dynamics for counties that have hydroelectric dams built before 1950. The dashed lines with hollow symbols (circles, squares, triangles, and diamonds) depicts the dynamics for counties that have dams built after 1950. Lastly, the thick solid line shows the dynamics of log density for counties in the control group.

Appendix B: Synthetic Propensity Score Reweighting

As an alternative to my reweighting-matching strategy, we can use a propensity score reweighting approach with weights from the synthetic control method (Abadie and Gardeazabal, 2003; and Abadie, Diamond and Hainmueller, 2010). I refer to such an approach as the "synthetic propensity score reweighting". If we treat the synthetic control weights as non-random, then we can just average the weights each untreated county gets across the synthetic control estimation of all treated counties and run the event-study analysis by weighted least squares in the full sample, giving each treated county a weight of one. Intuitively, the synthetic control weights provide us with a set of propensity score weights, and we can use them to compute average treatment effects on the treated.

An advantage to this approach over the reweighting-matching strategy is that we can easily cluster for more aggregated levels, such as state, in the estimation of standard errors. In the reweighting-matching case, the "location" of each synthetic control county is a weighted average of all originally defined control counties. On the other hand, in the reweighting-matching strategy we do not have to assume non-random weights in the estimation of standard errors. Each pair of treated and synthetic control counties is allowed to have arbitrary within-pair dependence, including the one arising from the uncertainty in the estimation of the synthetic control. Furthermore, because each synthetic control is matched to its treated county by construction, the comparison between synthetic control and treated counties becomes much more transparent.

In the Table B.I below, I compare four sets of estimates: (i) reweighting-matching with standard errors clustered at the case level[18], as in my main analysis; (ii) standard event-study with standard errors clustered at the county level (all treated and control counties are assumed to have weight one); (iii) synthetic propensity score reweighting with standard errors clustered at the county level; and (iv) synthetic propensity score reweighting with standard errors clustered at the state level. As we can see, the synthetic propensity score reweighting estimates are similar to the reweighting-matching ones. Regarding the cluster at the state level versus the county level, observe that there is a loss of power when using the state level but the long run estimates are still statistically significant at conventional levels.

[18]Recall that a case is just a pair of a treated and its corresponding synthetic control county.

Table B.I: The Impact of Pre-1950 Hydro Dams on Population Density - Comparison of Methodologies

ln(Pop Density)	TS	TC	TC_SPSR_C	TC_SPSR_S
	(1)	(2)	(3)	(4)
40 years before dam	-0.2428	-0.1441	0.1993	0.1993
	(0.2365)	(0.2740)	(0.2771)	(0.2497)
30 years before dam	-0.2564	-0.1493	-0.0547	-0.0547
	(0.1893)	(0.1941)	(0.2062)	(0.1649)
20 years before dam	-0.1536	-0.1253	-0.0614	-0.0614
	(0.1528)	(0.1740)	(0.1626)	(0.1305)
10 years before dam	-0.0398	-0.0337	-0.0248	-0.0248
	(0.1445)	(0.1544)	(0.1267)	(0.1080)
10 years after dam	0.1103**	0.0652	0.0603	0.0603
	(0.0451)	(0.0461)	(0.0524)	(0.0564)
20 years after dam	0.3063***	0.1574*	0.2146*	0.2146
	(0.0979)	(0.0811)	(0.1120)	(0.1332)
30 years after dam	0.4129***	0.2207**	0.2962**	0.2962
	(0.1131)	(0.0942)	(0.1332)	(0.1749)
40 years after dam	0.6551***	0.4118***	0.5194***	0.5194**
	(0.1408)	(0.1138)	(0.1586)	(0.2000)
50 years after dam	0.7897***	0.5176***	0.6693***	0.6693***
	(0.1681)	(0.1293)	(0.1840)	(0.2177)
60 years after dam	0.8699***	0.6036***	0.7835***	0.7835***
	(0.2047)	(0.1514)	(0.2179)	(0.2494)
70 years after dam	0.9584***	0.6026***	0.8759***	0.8759***
	(0.2454)	(0.1663)	(0.2663)	(0.2964)
80 years after dam	1.0494***	0.6448***	0.9246***	0.9246**
	(0.3052)	(0.2094)	(0.3382)	(0.3611)
Observations	660	935	935	935
R-squared	0.9837	0.9781	0.9836	0.9836

Notes: This table presents the short- and long-run effects of pre-1950 hydroelectric dams on population density. The estimated coefficients are the β's in equation (5) in the text. They are coefficients of event-time dummies. "TS" in the labels of the columns represents treatment vs. synthetic control, "TC" treatment vs. control, and "SPSR" synthetic propensity score reweighting. The synthetic control group consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. The control group contains the originally defined control counties. Standard errors are shown in parentheses. In column 1, they are clustered at a case level. A case is a pair of a treated county and its corresponding synthetic control. In columns 2 and 3, they are clustered at the county level ("C"). In column 4, they are clustered at the state level ("S"). "Observations" reports the number of county-year observations.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Appendix C: Hydroelectric Dams and Farmland Values

Table C.I: Short- and Long-Run Effects of Hydro Dams on Average Value of Farmland

ln(Average Value of Farmland)	TS_b1950	TC_b1950	TS_a1950	TC_a1950
	(1)	(2)	(3)	(4)
80 years before dam			0.1887	0.1267
			(0.1341)	(0.1528)
70 years before dam			0.1425	0.0892
			(0.1140)	(0.1377)
60 years before dam			0.0800	-0.0153
			(0.1030)	(0.1244)
50 years before dam			0.0607	-0.0536
			(0.0858)	(0.1040)
40 years before dam	-0.1932	-0.2882	-0.0113	-0.0972
	(0.3165)	(0.3087)	(0.0771)	(0.0889)
30 years before dam	-0.1767	-0.1984	-0.0031	-0.0610
	(0.2652)	(0.2618)	(0.0717)	(0.0839)
20 years before dam	-0.2911	-0.2825	-0.0015	-0.0390
	(0.2850)	(0.2924)	(0.0705)	(0.0795)
10 years before dam	-0.3019	-0.2482	0.0046	-0.0177
	(0.2262)	(0.2341)	(0.0666)	(0.0733)
10 years after dam	0.0881	0.0689	0.0213	0.0141
	(0.0523)	(0.0551)	(0.0304)	(0.0352)
20 years after dam	0.1678**	0.1554**	0.0824*	0.0432
	(0.0719)	(0.0686)	(0.0417)	(0.0536)
30 years after dam	0.1721*	0.1461*	0.1109**	0.0464
	(0.0976)	(0.0859)	(0.0540)	(0.0709)
40 years after dam	0.1449	0.1786	0.0890	0.0213
	(0.1437)	(0.1224)	(0.0721)	(0.0923)
50 years after dam	0.1916	0.2652**	0.0658	0.0153
	(0.1405)	(0.1328)	(0.1008)	(0.1313)
60 years after dam	0.1789	0.2937*		
	(0.1626)	(0.1508)		
70 years after dam	0.0932	0.3108*		
	(0.1750)	(0.1797)		
80 years after dam	-0.0088	0.1949		
	(0.1398)	(0.1640)		
Observations	660	902	1518	1331
R-squared	0.9685	0.9451	0.9653	0.9452

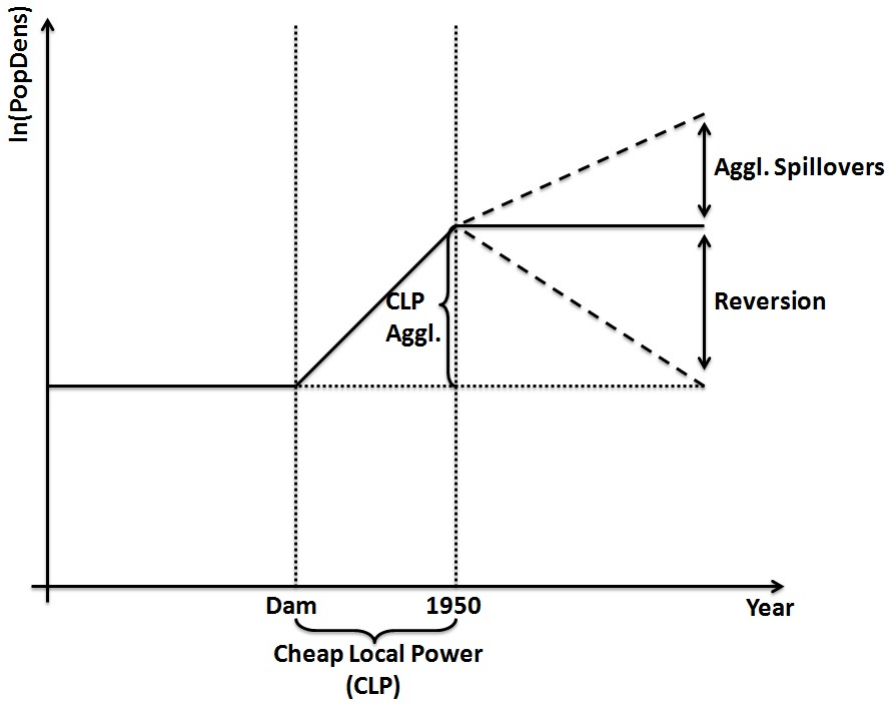
Notes : This table presents the short- and long-run effects of hydroelectric dams on the average value of farmland ($/acre). The estimated coefficients are the β's in equation (5) in the text. They are coefficients of event-time dummies. "TS" in the labels of the columns represents treatment vs. synthetic control, "TC" treatment vs. control, "b1950" before 1950, and "a1950" after 1950. The synthetic control group consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. The control group contains the originally defined control counties. Standard errors are shown in parentheses. In "TS" columns, they are clustered at a case level. A case is a pair of a treated county and its corresponding synthetic control. In "TC" columns, they are clustered at the county level. "Observations" reports the number of county-year observations.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

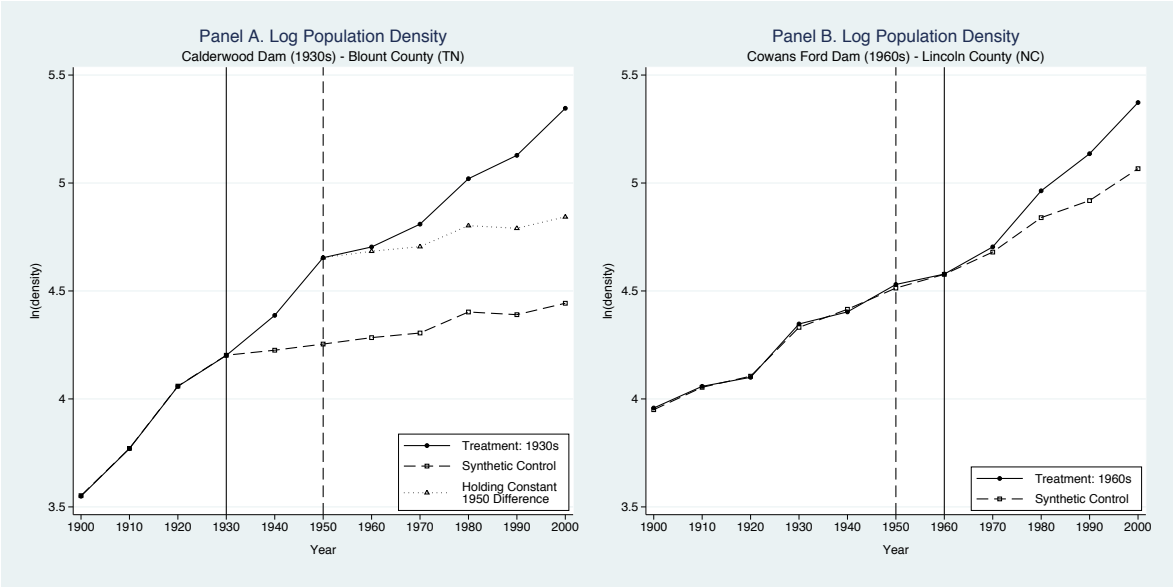
* Significant at the 10 percent level.

Figure I: Identification of Agglomeration Spillovers



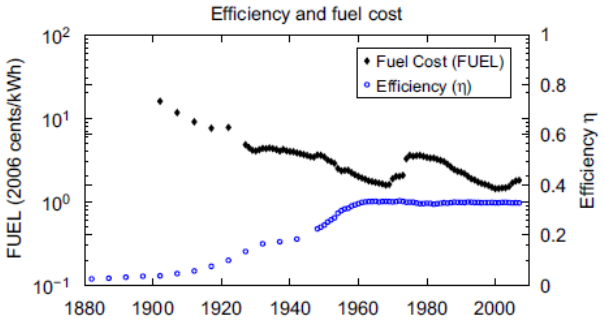
Notes: This Figure depicts my strategy to identify agglomeration spillovers. Time is represented in the horizontal axis. My analysis spans the entire twentieth century. Log of population density, my main outcome variable capturing economic activity, is represented in the vertical axis. Initially, two hypothetical counties - treated and counterfactual - have the same population density. From the completion of the dam (the treatment) until mid century, time period displayed between the vertical dotted lines, the treated county experiences a temporary CLP advantage. This advantage gives rise to a strong concentration of economic activity around dam sites, which is represented by the CLP agglomeration jump between dam completion and 1950. The dynamics of population density in the second half of the twentieth century, represented on the right-hand side of the second vertical dotted line, sheds light on the presence and magnitude of agglomeration spillovers. If the jump in population density that happens between the completion of the dam and the middle of century reverts, then there would be no agglomeration spillovers. If population density stabilizes after the jump, then there would be persistence in economic activity probably due to sunk investments and/or durability of goods such as housing. If population density continues growing after the initial CLP-driven jump, then there would be agglomeration spillovers, and the magnitude would be measured by the difference between the post-1950 level and the 1950 level of population density.

Figure II: Impact of Hydro Dams on Population Density: Pre-1950 Dams vs. Post-1950 Dams



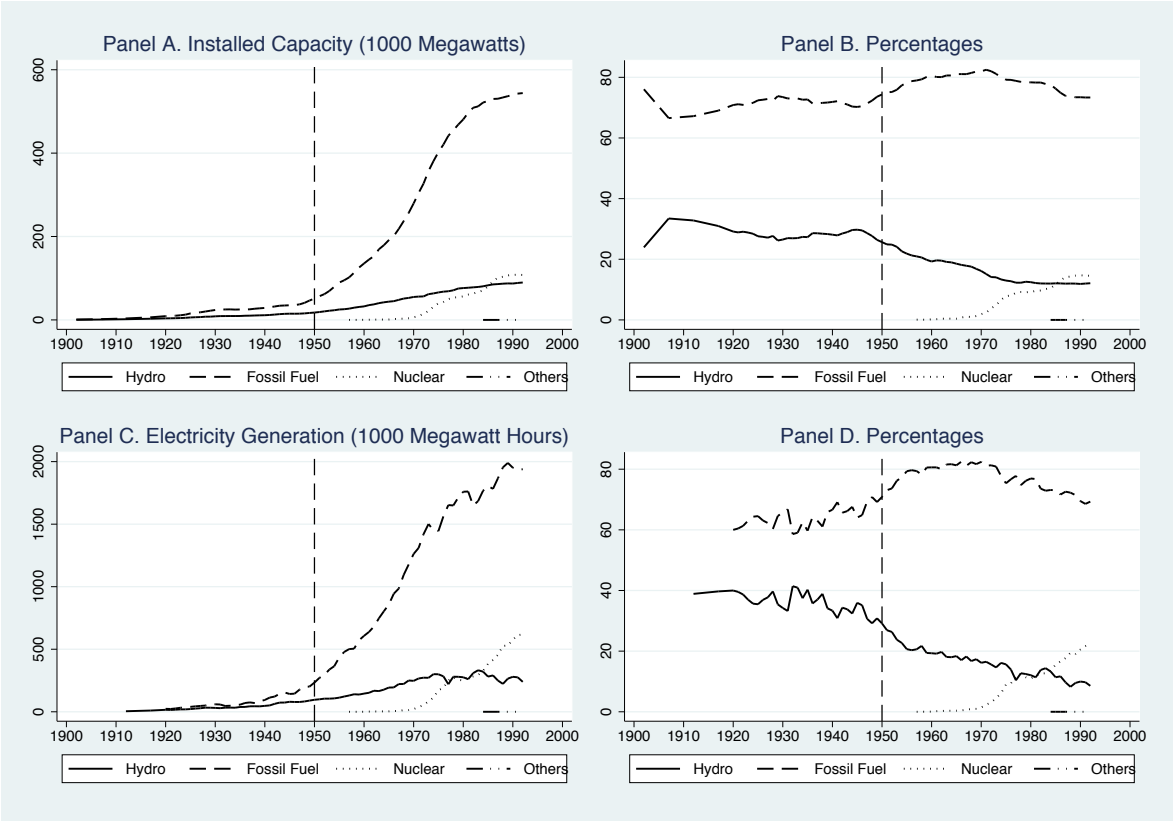
Notes: This Figure plots the log population density from 1900 to 2000 for two counties: one with a hydroelectric dam built before 1950 (panel A), and the other with a dam built after 1950 (panel B). The solid vertical line shows the decade in which the dam was completed. The dashed vertical line simply separates the twentieth century into pre-1950 and post-1950. The cutoff 1950 is my assumed turning point for the attenuation of the advantage of cheap local hydroelectricity. The solid line with solid circles displays the observed time series of log population density for each county, and the dashed line with hollow squares depicts the predicted time series of a corresponding synthetic control county. The synthetic control is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. The dotted line with hollow triangles in panel A displays the simulated times series of the treated county if growth in population density had stopped in 1950.

Figure III: Average efficiency of coal plants and the cost of the fuel component, 1882-2006



Source: McNerneya, Farmer and Trancik (2011, p. 3045).

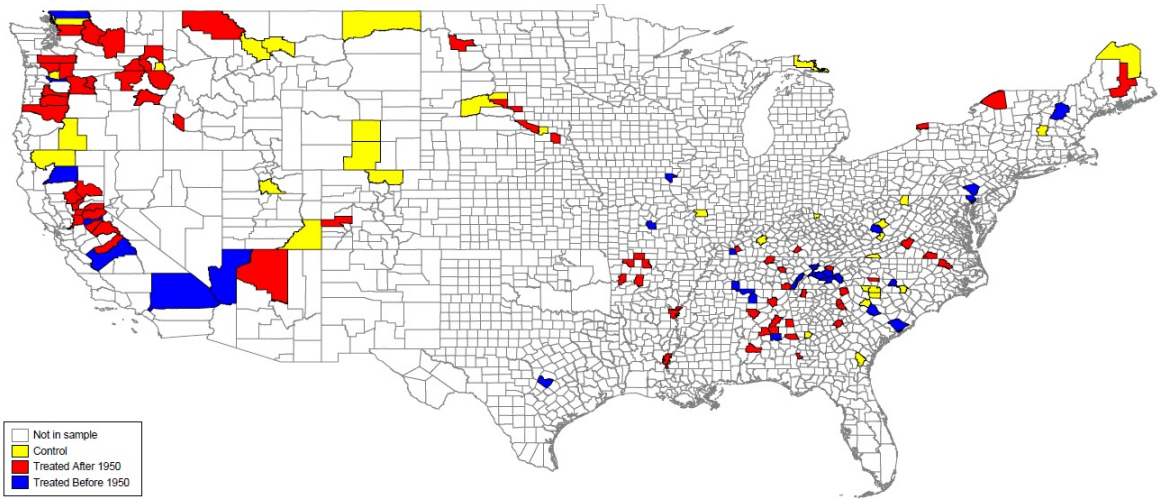
Figure IV: Installed Capacity and Electricity Generation in the U.S.



Notes: This Figure displays the evolution of installed capacity and electricity generation by source in the U.S. throughout the twentieth century. Panel A depicts installed capacity in thousands of megawatts. Panel B shows corresponding percentages. Panel C displays electricity generation in thousands of megawatt hours. Panel D shows corresponding percentages. The dashed vertical line simply separates the twentieth century into pre-1950 and post-1950. The cutoff 1950 is my assumed turning point for the attenuation of the advantage of cheap local hydroelectricity. "Hydro" includes conventional hydroelectricity and pumped storage. "Fossil Fuel" includes coal, natural gas, petroleum, geothermal, other gases, and wood and wood derived fuels. "Others" includes solar and wind.

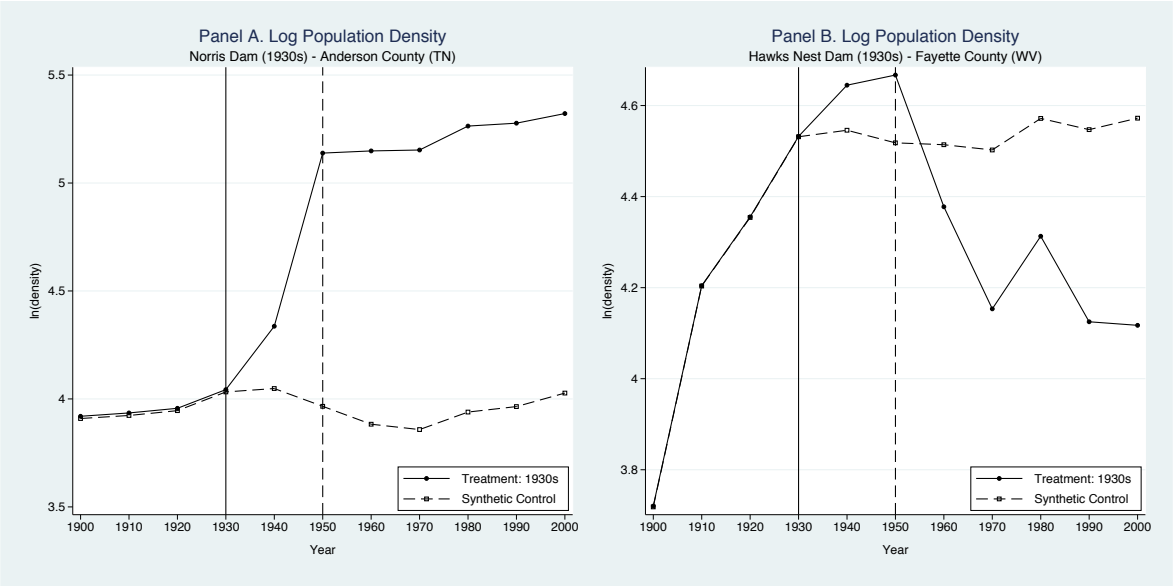
Source: *Historical Statistics of the Electric Utility Industry through 1992,* edited by Lizette Cintrón, 1995.

Figure V: Sample Counties: Treatment vs. Control



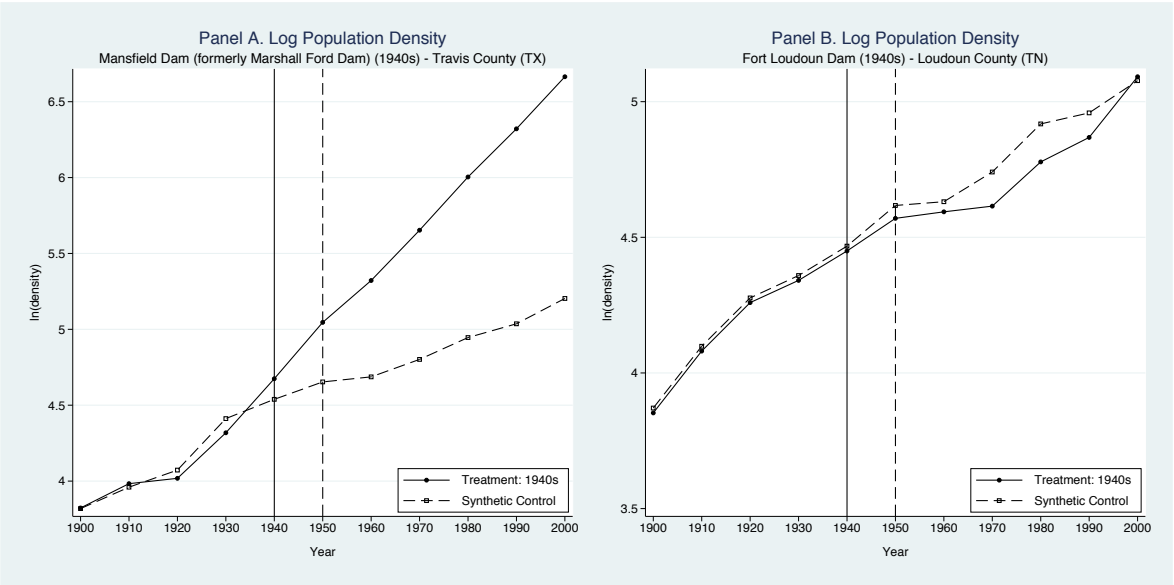
Notes: This Figure displays the counties included in my sample. "Treated Before 1950" are indicated in blue and have hydroelectric dams built in the first half of the twentieth century. "Treated After 1950" are indicated in green and have hydroelectric dams built in the second half of the century. "Control" are indicated in yellow and have hydropower potential comparable to the capacity installed in treated counties, but have no hydroelectric facilities. Counties left blank are not in my sample. Thin gray lines denote 1900 county borders, which are held constant throughout my analysis.

Figure VI: Impact of Hydro Dams on Population Density: Pre-1950 Dams



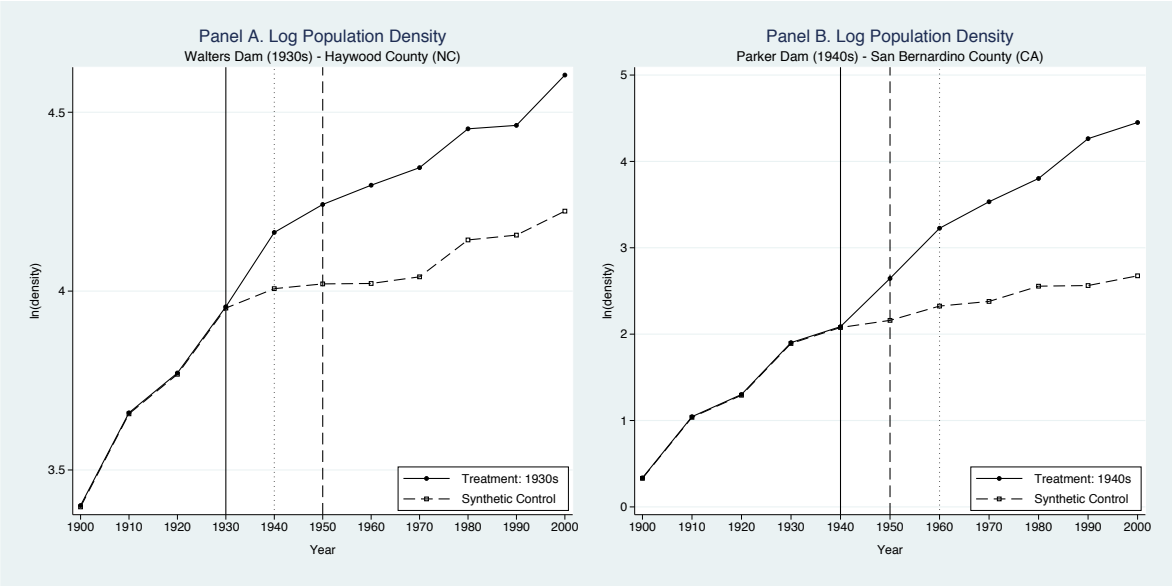
Notes: This Figure plots the log population density from 1900 to 2000 for two counties: one with virtually no agglomeration effects after 1950 (panel A), and the other with negative growth after 1950 (panel B). The solid vertical line shows the decade in which the dam was completed. The dashed vertical line simply separates the twentieth century into pre-1950 and post-1950. The cutoff 1950 is my assumed turning point for the attenuation of the advantage of cheap local hydroelectricity. The solid line with solid circles displays the observed time series of log population density for each county, and the dashed line with hollow squares depicts the predicted time series of a corresponding synthetic control county. The synthetic control is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams.

Figure VII: Impact of Hydro Dams on Population Density: Pre-1950 Dams



Notes: This Figure plots the log population density from 1900 to 2000 for two counties: one with constant growth after dam completion (panel A), and the other with virtually no difference relative to the counterfactual of no dams (panel B). The solid vertical line shows the decade in which the dam was completed. The dashed vertical line simply separates the twentieth century into pre-1950 and post-1950. The cutoff 1950 is my assumed turning point for the attenuation of the advantage of cheap local hydroelectricity. The solid line with solid circles displays the observed time series of log population density for each county, and the dashed line with hollow squares depicts the predicted time series of a corresponding synthetic control county. The synthetic control is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams.

Figure VIII: Impact of Hydro Dams on Population Density: Pre-1950 Dams

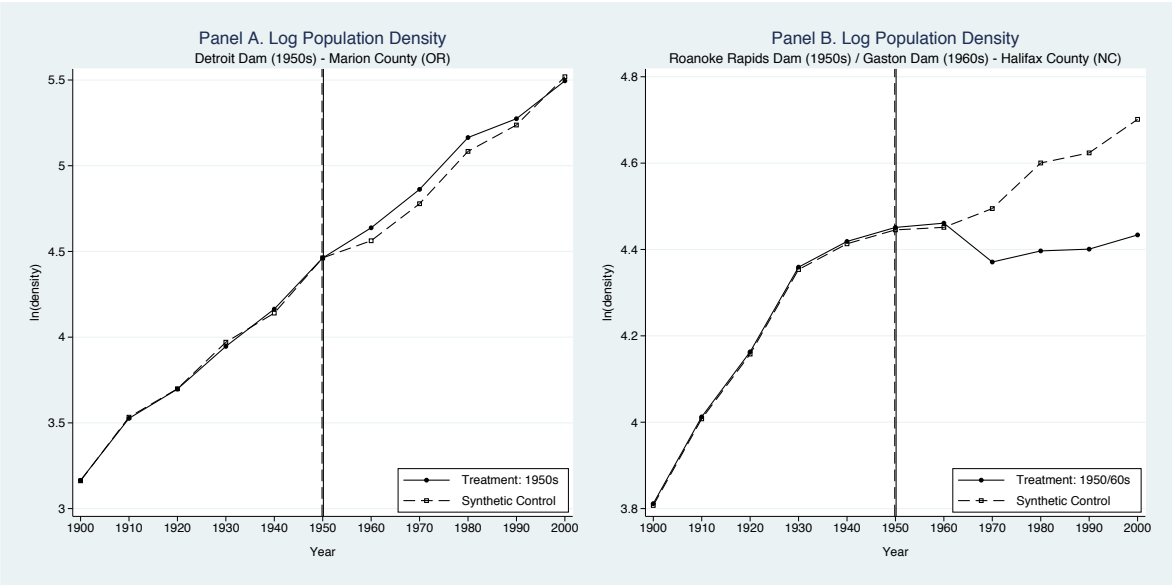


Notes: This Figure plots the log population density from 1900 to 2000 for two counties: one with growth flattening before 1950 (panel A), and the other with growth flattening after 1950 (panel B). The solid vertical line shows the decade in which the dam was completed. The dashed vertical line simply separates the twentieth century into pre-1950 and post-1950. The cutoff 1950 is my assumed turning point for the attenuation of the advantage of cheap local hydroelectricity. The dotted vertical line shows a cutoff that might be a more accurate turning point for the county in question. The solid line with solid circles displays the observed time series of log population density for each county, and the dashed line with hollow squares depicts the predicted time series of a corresponding synthetic control county. The synthetic control is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams.

Figure IX: Impact of Hydro Dams on Population Density: Post-1950 Dams

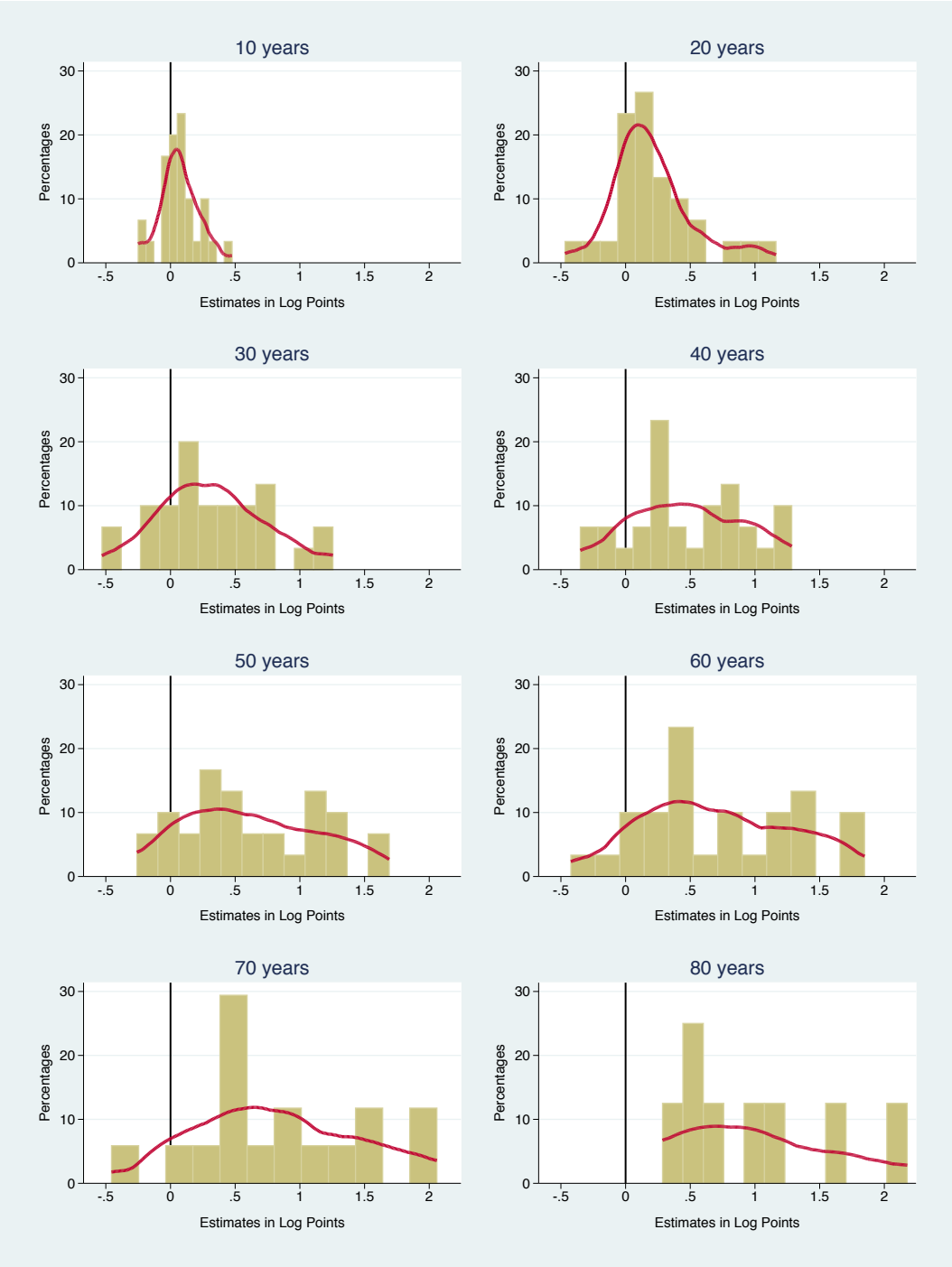
Notes: This Figure plots the log population density from 1900 to 2000 for two counties: one with moderately positive effects of dams (panel A), and the other with small but positive dam effects (panel B). The solid vertical line shows the decade in which the dam was completed. The dashed vertical line simply separates the twentieth century into pre-1950 and post-1950. The cutoff 1950 is my assumed turning point for the attenuation of the advantage of cheap local hydroelectricity. The solid line with solid circles displays the observed time series of log population density for each county, and the dashed line with hollow squares depicts the predicted time series of a corresponding synthetic control county. The synthetic control is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams.

Figure X: Impact of Hydro Dams on Population Density: Post-1950 Dams



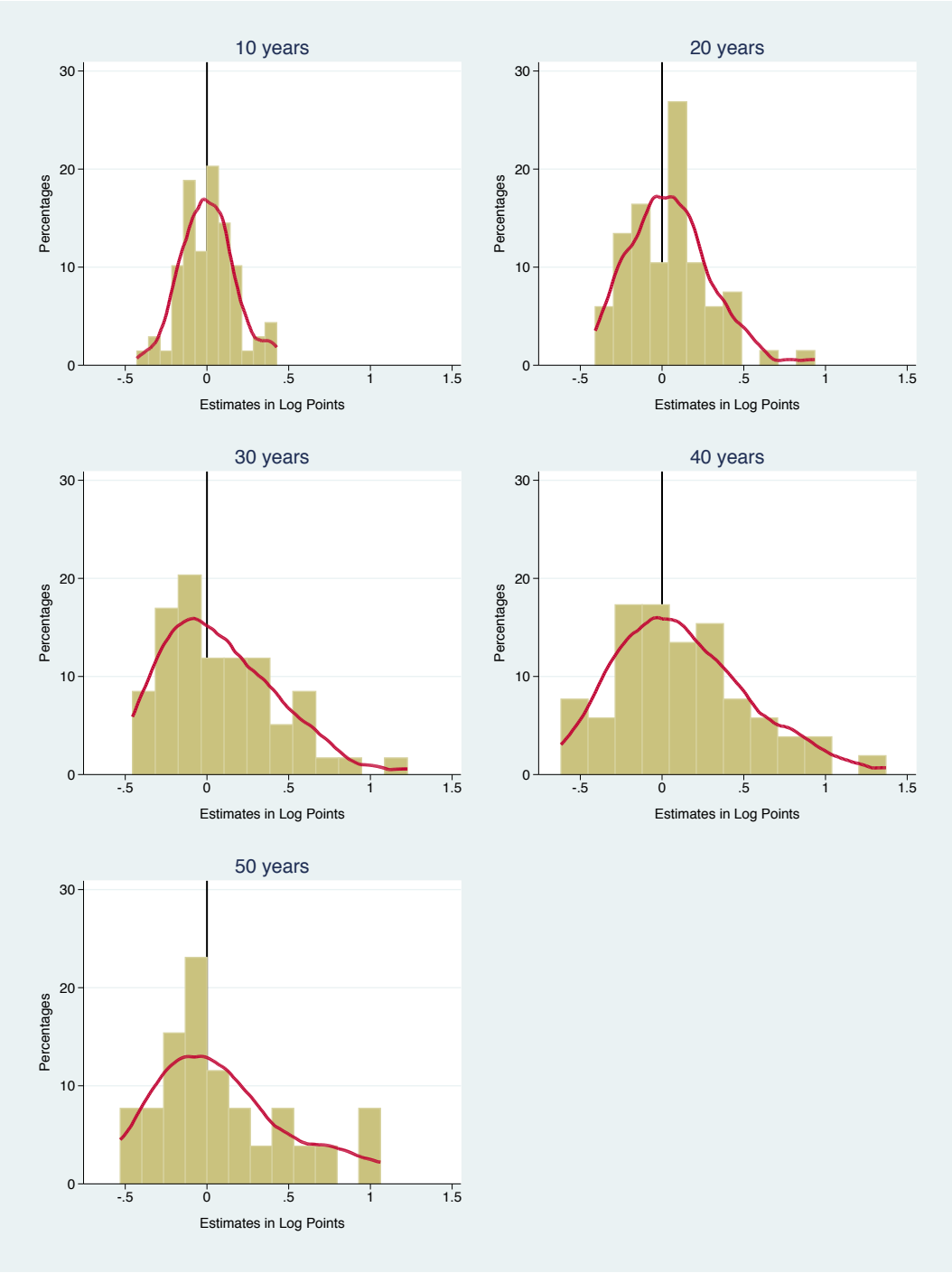
Notes: This Figure plots the log population density from 1900 to 2000 for two counties: one with virtually no dam effects (panel A), and the other with somewhat negative effects of dams (panel B). The solid vertical line shows the decade in which the dam was completed. The dashed vertical line simply separates the twentieth century into pre-1950 and post-1950. The cutoff 1950 is my assumed turning point for the attenuation of the advantage of cheap local hydroelectricity. The solid line with solid circles displays the observed time series of log population density for each county, and the dashed line with hollow squares depicts the predicted time series of a corresponding synthetic control county. The synthetic control is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams.

Figure XI: Impact of Hydro Dams on Population Density: Distribution of County Estimates by Years Since Dam Completion (Pre-1950 Dams)



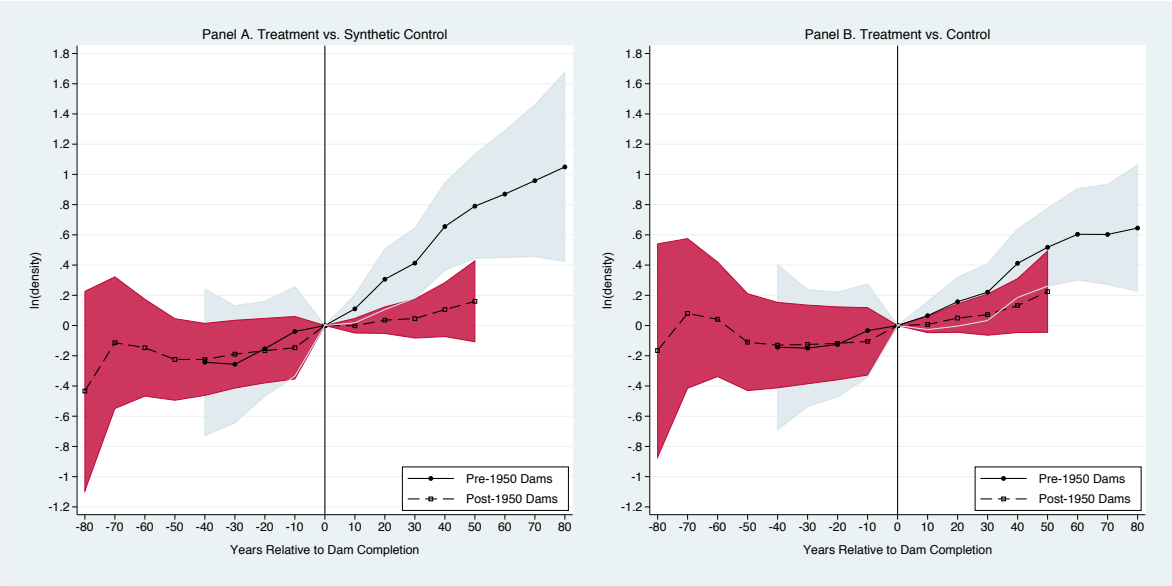
Notes: This Figure plots the distribution of county-specific estimates of the impact of pre-1950 hydro dams on the log of population density. The estimates are found by using the synthetic control estimator for each treated county. The solid vertical line at zero just emphasizes the point of no effects of dams. Each vertical bar in each histogram shows the percentage of counties with effects of dams in a particular bin. The thick solid curved line just smoothes the distribution of effects using a kernel density estimator.

Figure XII: Impact of Hydro Dams on Population Density: Distribution of County Estimates by Years Since Dam Completion (Post-1950 Dams)



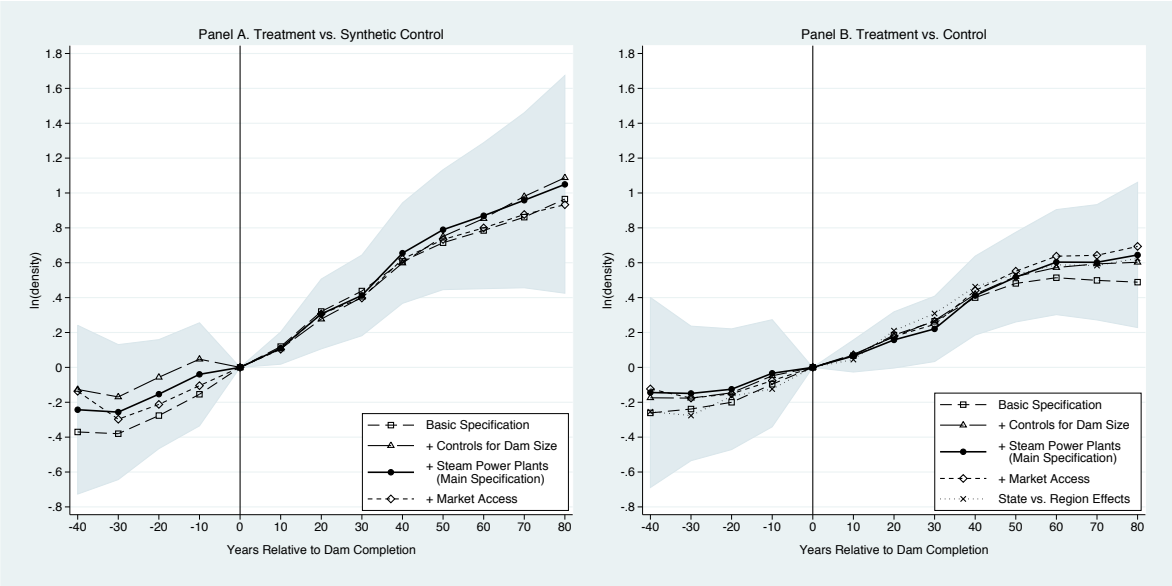
Notes: This Figure plots the distribution of county-specific estimates of the impact of post-1950 hydro dams on the log of population density. The estimates are found by using the synthetic control estimator for each treated county. The solid vertical line at zero just emphasizes the point of no effects of dams. Each vertical bar in each histogram shows the percentage of counties with effects of dams in a particular bin. The thick solid curved line just smoothes the distribution of effects using a kernel density estimator.

Figure XIII: Impact of Hydro Dams on Population Density: Short- and Long-Run Estimates



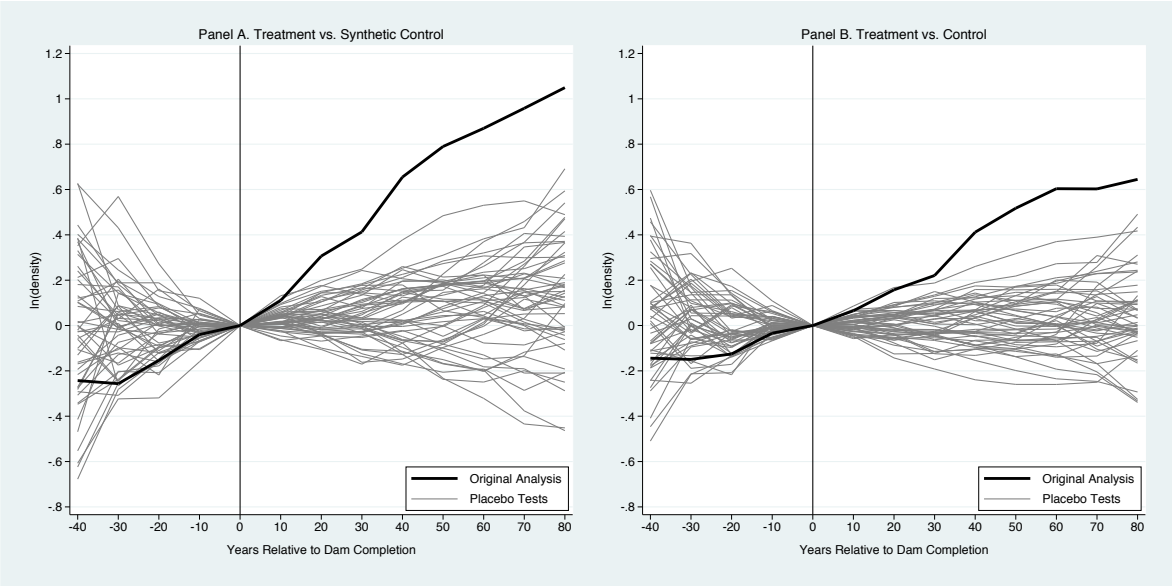
Notes: This Figure presents the short- and long-run effects of large hydroelectric dams on population density. Each panel graphs the estimated coefficients β's of event-time dummies from equation (6) in the text. The vertical solid line at zero facilitates the comparison of the dynamics before and after the treatment. It also points out that the dummy for event time zero is omitted in the estimation. The solid lines with solid circles report differences in the log of population density relative to year of dam completion, for counties that had hydroelectric dams installed before 1950. The light (bluish gray) shade around the solid line depicts the 95% confidence interval for the coefficients represented by solid circles. The dashed lines with hollow squares report similar differences for counties that had dams installed after 1950. The dark (cranberry) shade around the dashed line depicts the 95% confidence interval for the coefficients represented by hollow squares. In panel A, the control group used in the estimation consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. In panel B, the control set contains the originally defined control counties.

Figure XIV: Impact of Hydro Dams on Population Density: Specification Checks - Pre-1950 Dams



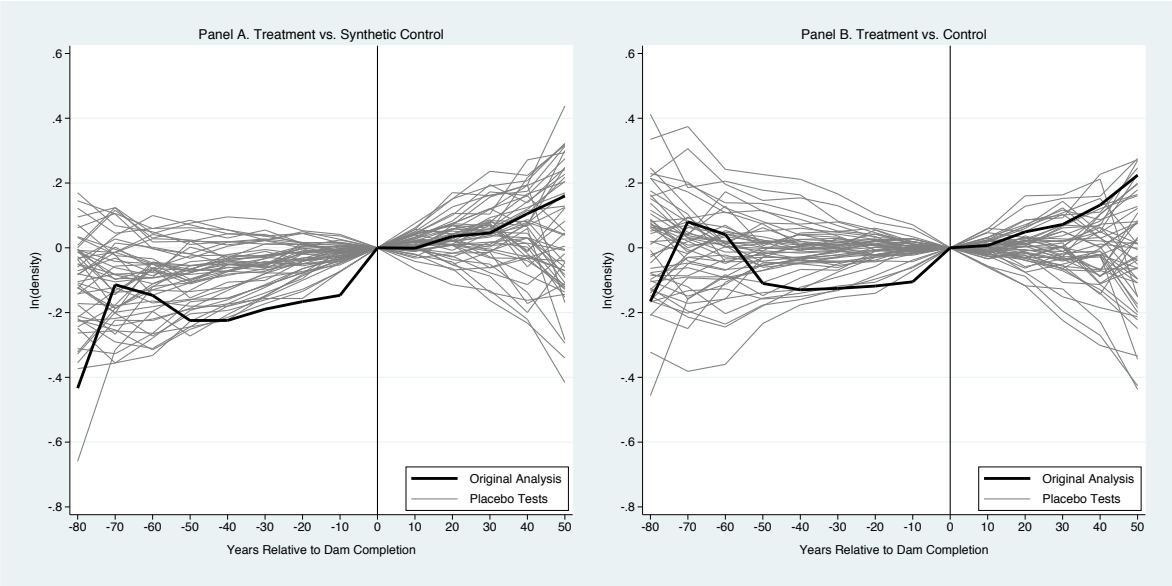
Notes: This Figure presents some robustness checks regarding the specification used in the estimation of equation (6) in the text. Each panel graphs the estimated coefficients β's of event-time dummies from equations like that one. The vertical solid line at zero facilitates the comparison of the dynamics before and after the treatment. It also points out that the dummy for event time zero is omitted in the estimation. Each line reports differences in the log of population density relative to year of dam completion, for counties that had hydroelectric dams installed before 1950. The dashed lines with hollow squares display estimates related to the basic specification. Such specification includes event-time dummies, county effects, region-by-year fixed effects, and time-invariant county characteristics (cubic function in latitute and longitude, and 50-year average rainfall and 50-year average temperature for each season of the year) interacted with year effects. The long-dashed lines with hollow triangles show estimates associated with the basic specification plus a cubic function in dam capacity. The solid lines with solid circles report the estimates related to the main specification, which is basic specification plus controls for dam size plus a cubic function in thermal power plant capacity. The short-dashed lines with hollow diamonds display estimates related to the main specification plus controls for the interaction of year effects with three county-specific measures of market access: mileage of railroad tracks, distance to closest waterway, and log of market access as estimated by Donaldson and Hornbeck (2012). The light (bluish gray) shade around the solid line depicts the 95% confidence interval for the coefficients of the main specification, represented by solid circles. In panel B, the dotted line with little crosses displays estimates associated with the main specification, but replacing region-by-year fixed effects with state-by-year fixed effects. This line does not appear in panel A because the sample size was insufficient to estimate all the parameters. In panel A, the control group used in the estimation consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. In panel B, the control set contains the originally defined control counties.

Figure XV: Impact of Hydro Dams on Population Density: Placebo - Pre-1950 Dams



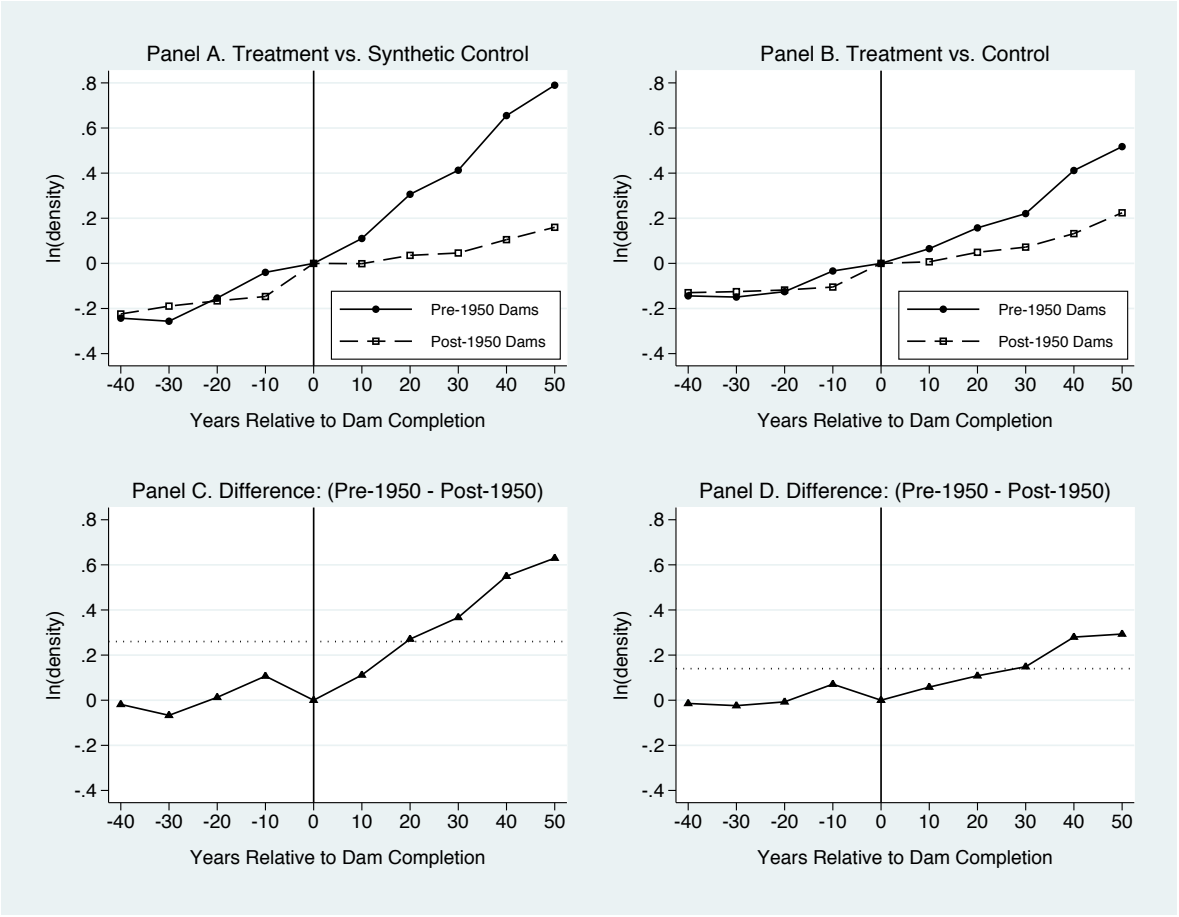
Notes: This Figure plots estimates of placebo tests with pre-1950 dams. Each panel graphs estimated coefficients β's of event-time dummies from equation (6) in the text for actual and artificially assigned treatments. The vertical solid line at zero facilitates the comparison of the dynamics before and after the treatment. It also points out that the dummy for event time zero is omitted in the estimation. The thick solid black line displays the effects of actual dams. The thin solid gray lines show the effects of artificially treated counties. Artificial treatment is assigned through a two-step procedure. First, counties are drawn randomly from the pool of originally treated and control counties. The number of selected counties is identical to the number of originally treated counties. Second, for each artificially treated county, the date of dam completion is set at random as well. In panel A, the control group used in the estimation consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. In panel B, the control set contains the originally defined control counties.

Figure XVI: Impact of Hydro Dams on Population Density: Placebo - Post-1950 Dams



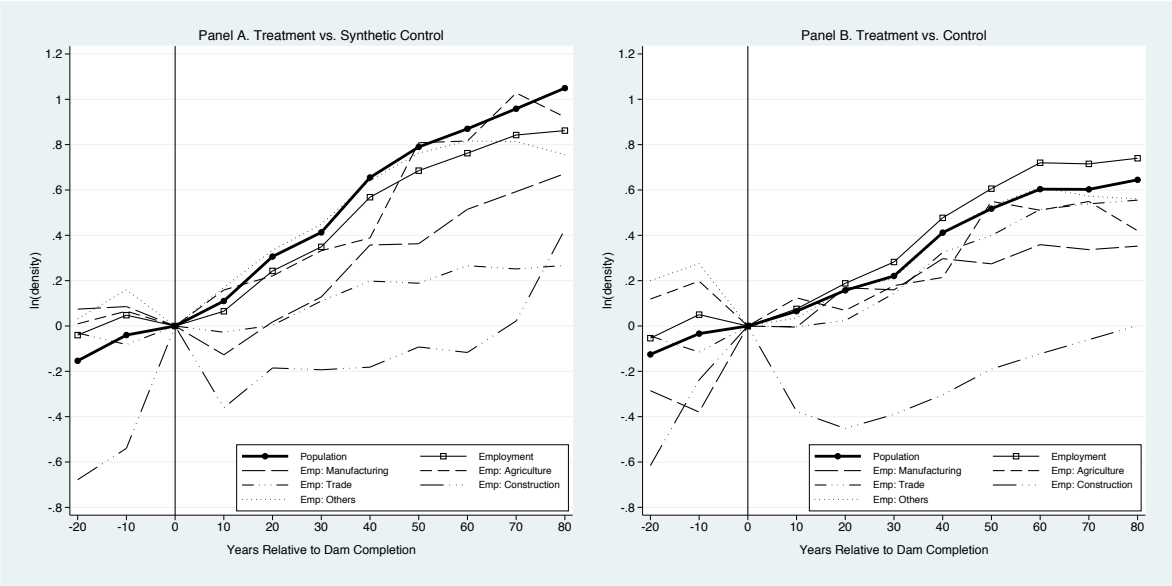
Notes: This Figure plots estimates of placebo tests with post-1950 dams. Each panel graphs estimated coefficients β's of event-time dummies from equation (6) in the text for actual and artificially assigned treatments. The vertical solid line at zero facilitates the comparison of the dynamics before and after the treatment. It also points out that the dummy for event time zero is omitted in the estimation. The thick solid black line displays the effects of actual dams. The thin solid gray lines show the effects of artificially treated counties. Artificial treatment is assigned through a two-step procedure. First, counties are drawn randomly from the pool of originally treated and control counties. The number of selected counties is identical to the number of originally treated counties. Second, for each artificially treated county, the date of dam completion is set at random as well. In panel A, the control group used in the estimation consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. In panel B, the control set contains the originally defined control counties.

Figure XVII: Impact of Hydro Dams on Population Density: Agglomeration Spillovers



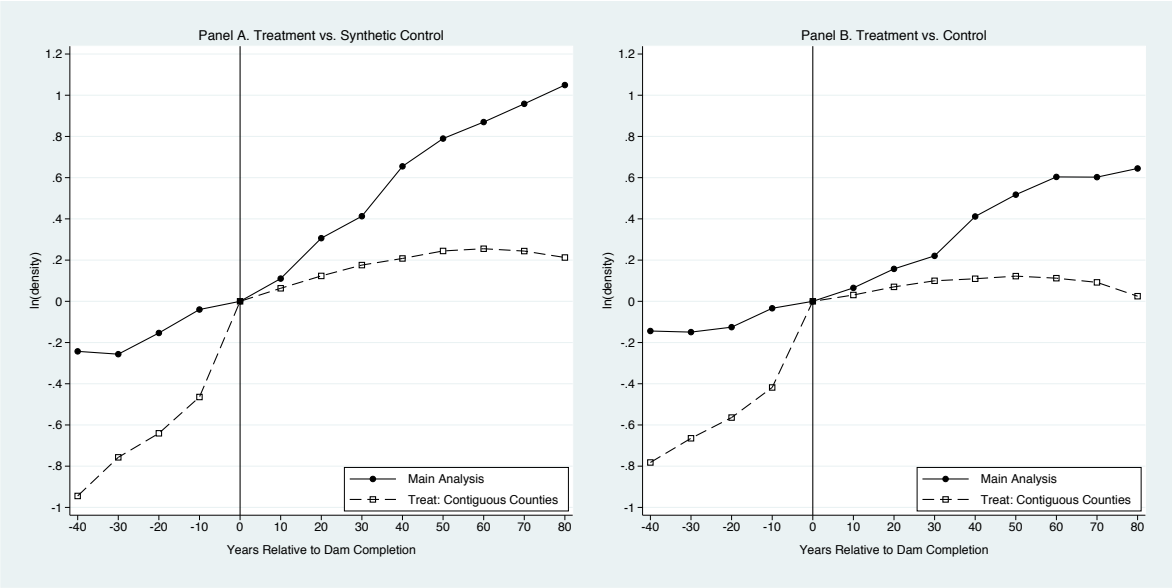
Notes: This Figure presents my estimates of agglomeration spillovers. The top panels plot the estimated coefficients β's of event-time dummies from equation (6) in the text. The bottom panels plot the differences between coefficients associated with pre-1950 dams and post-1950 dams. Each panel shows estimates from 40 years before dam installation until 50 years after the dam: the differences in the bottom charts can be computed only for this range. The vertical solid line at zero facilitates the comparison of the dynamics before and after the treatment. It also points out that the dummy for event time zero is omitted in the estimation of the coefficients. The solid lines with solid circles in the top panels report differences in the log of population density relative to year of dam completion, for counties that had hydroelectric dams installed before 1950. The dashed lines with hollow squares in the top panels report similar differences for counties that had dams installed after 1950. The solid lines with solid triangles in the bottom panels report the difference between the solid line with solid circles and the dashed line with hollow squares of the corresponding top panel. The dotted lines in the bottom panels report the effect of dams attributable to the advantage of cheap local hydroelectricity until 1950. My (lower bound) estimates of agglomeration spillovers consist of the differences between the solid line with solid triangles and the dotted line, in the bottom panels. In panels A and C, the control group used in the estimation consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. In panels B and D, the control set contains the originally defined control counties.

Figure XVIII: Impact of Hydro Dams on Population/Employment Density: Employment by Sector - Pre-1950 Dams



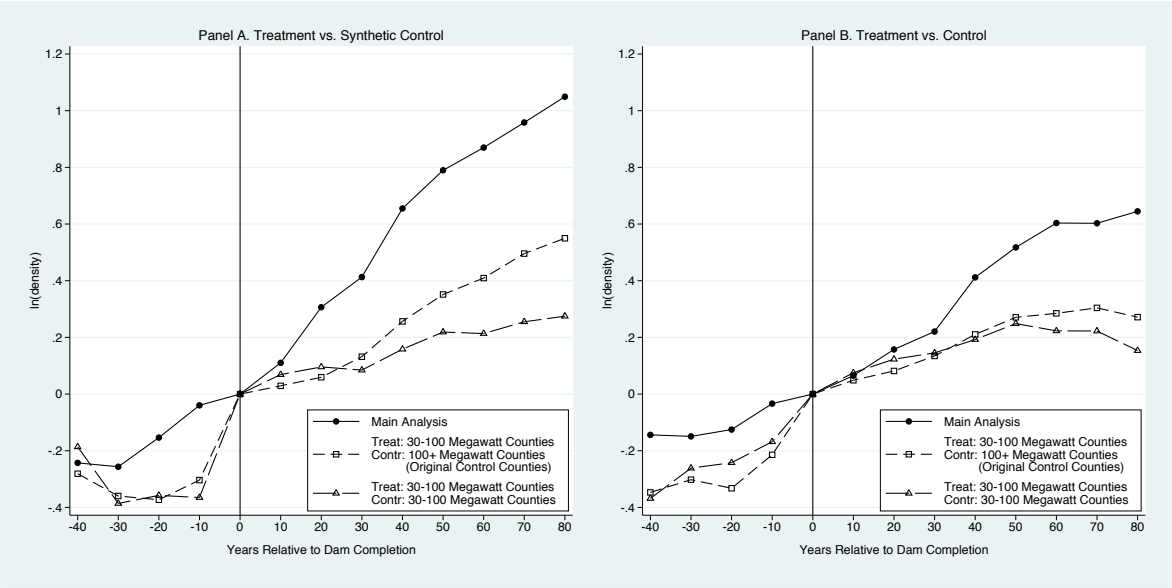
Notes: This Figure presents the impact of pre-1950 dams on population and employment by sector. Each panel graphs the estimated coefficients β's of event-time dummies from equation (6) in the text, for the log density of each outcome mentioned in the legend. The vertical solid line at zero facilitates the comparison of the dynamics before and after the treatment. It also points out that the dummy for event time zero is omitted in the estimation. The thick solid lines with solid circles report differences in the log of population density relative to year of dam completion, for counties that had hydroelectric dams installed before 1950. The thin solid lines with hollow squares report similar estimates, but for log of employment density. More specifically, the long-dashed lines display similar estimates for log of employment in manufacturing, the dashed lines for log of employment in agriculture, the dash-3dotted lines for log of employment in trade (wholesale plus retail), the long-dash-3dotted lines for log of employment in construction, and the dotted lines for log of employment in other sectors. "Other sectors" includes employment in industries that provide local goods and services such as real estate, cleaning services, legal services, medical services, and personal services. In panel A, the control group used in the estimation consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. In panel B, the control set contains the originally defined control counties.

Figure XIX: Impact of Hydro Dams on Population Density: Analysis with Contiguous Counties - Pre-1950 Dams



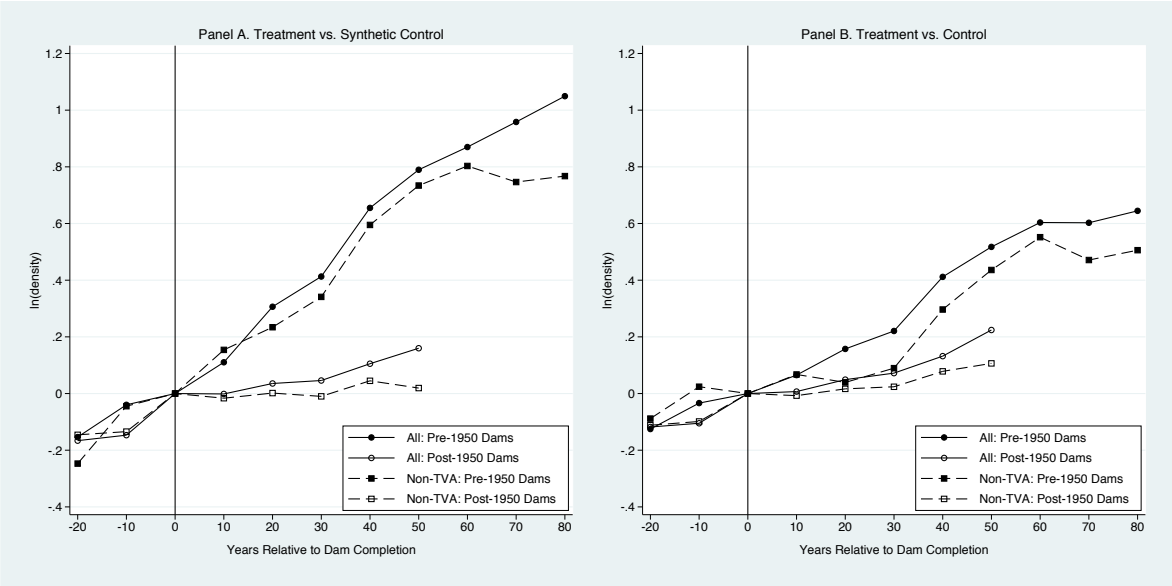
Notes: This Figure presents potential externalities of pre-1950 dams on population density of neighboring counties. Each panel graphs the estimated coefficients β's of event-time dummies from equation (6) in the text. The vertical solid line at zero facilitates the comparison of the dynamics before and after the treatment. It also points out that the dummy for event time zero is omitted in the estimation. The solid lines with solid circles report differences in the log of population density relative to year of dam completion, for counties that had hydroelectric dams installed before 1950. The dashed lines with hollow squares report similar estimates, but with counties contiguous to the treated ones as the treatment. In panel A, the control group used in the estimation consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. In panel B, the control set contains the originally defined control counties.

Figure XX: Impact of Hydro Dams on Population Density: Analysis with Counties with Capacity of 30-100 Megawatts - Pre-1950 Dams



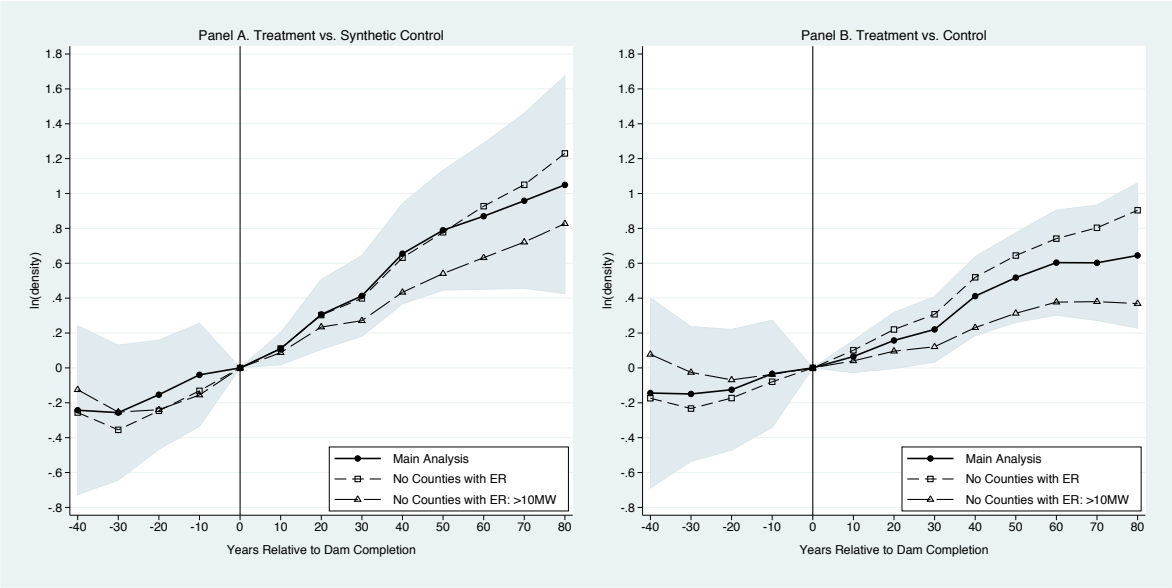
Notes: This Figure presents effects of smaller pre-1950 dams on population density. Each panel graphs the estimated coefficients β's of event-time dummies from equation (6) in the text. The vertical solid line at zero facilitates the comparison of the dynamics before and after the treatment. It also points out that the dummy for event time zero is omitted in the estimation. The solid lines with solid circles report differences in the log of population density relative to year of dam completion, for counties that had hydroelectric dams installed before 1950. The dashed lines with hollow squares report similar estimates, but with treated counties with hydropower capacity in the range of 30-100 megawatts. The long-dashed lines with hollow triangles report comparable estimates, but with treated and control counties with hydropower capacity in the range 30-100 megawatts. In panel A, the control group used in the estimation consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. In panel B, the control set contains the originally defined control counties.

Figure XXI: Impact of Hydro Dams on Population Density: All vs. Non-TVA Counties



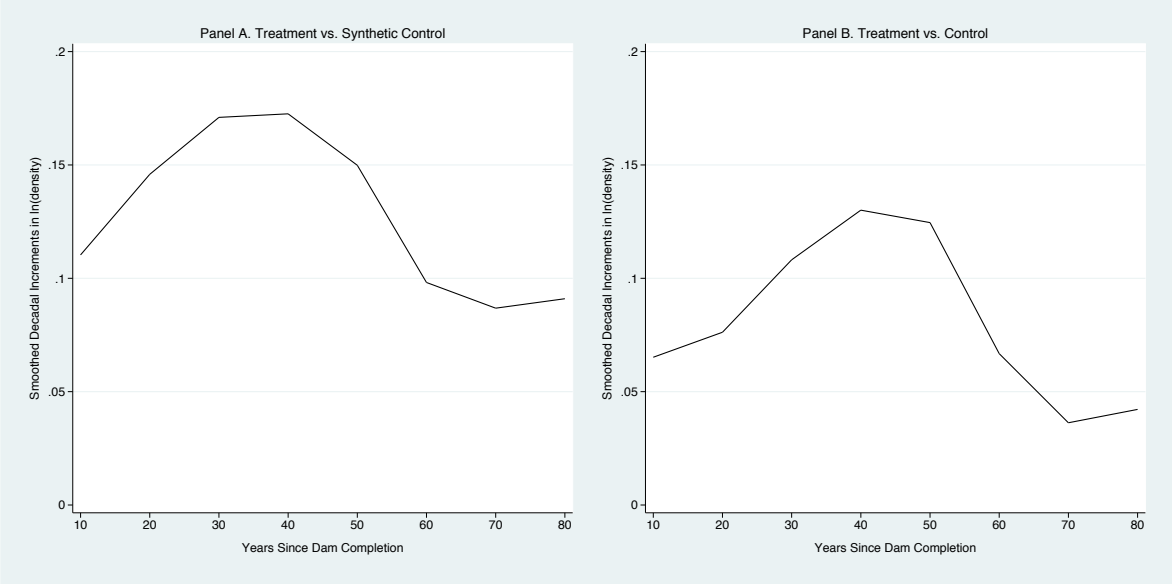
Notes: This Figure presents the effects of dams estimated with all counties in my sample versus with only counties outside the Tennessee Valley Authority (TVA) region. The TVA was one of the most ambitious place based economic development policies in the history of the U.S. Besides hydroelectric dams, it included construction of an extensive network of new roads, canals, and flood control systems. Each panel graphs the estimated coefficients β's of event-time dummies from equation (6) in the text. The vertical solid line at zero facilitates the comparison of the dynamics before and after the treatment. It also points out that the dummy for event time zero is omitted in the estimation. The solid lines with solid circles report differences in the log of population density relative to year of dam completion, for all sample counties that had hydroelectric dams installed before 1950. The solid lines with hollow circles report similar differences for all sample counties that had dams installed after 1950. The dashed lines with solid squares report differences in the log of population density relative to year of dam completion, for non-TVA counties that had hydroelectric dams installed before 1950. The dashed lines with hollow squares report similar differences for non-TVA counties that had dams installed after 1950. In panel A, the control group used in the estimation consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. In panel B, the control set contains the originally defined control counties.

Figure XXII: Impact of Hydro Dams on Population Density: All vs. No Control Counties With Environmental Regulations (ER) - Pre-1950 Dams



Notes: This Figure presents the short- and long-run effects of hydroelectric dams on population density, with and without control counties subject to land regulations aimed at preserving wilderness and wildlife. Such environmental regulations could restrict population growth in those locations. Each panel graphs the estimated coefficients β's of event-time dummies from equation (6) in the text. The vertical solid line at zero facilitates the comparison of the dynamics before and after the treatment. It also points out that the dummy for event time zero is omitted in the estimation. The solid lines with solid circles report differences in the log of population density relative to year of dam completion, for all sample counties that had hydroelectric dams installed before 1950. These are the estimates from the main analysis. The dashed lines with hollow squares report similar differences, but excluding any control counties with environmental regulations from the analysis. The long-dashed lines with hollow triangles also report similar differences as in the main analysis, but redefining the original control group to include all counties without environmental regulations but with hydropower potential starting at 10 megawatts (instead of 100 megawatts, as in the main analysis). In panel A, the control group used in the estimation consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. In panel B, the control set contains the originally defined control counties.

Figure XXIII: Impact of Pre-1950 Hydro Dams on Population Density: Non-Structural Shape of Agglomeration Function



Notes: This Figure presents the shape of the agglomeration function based on the estimates from equation (6). Each panel displays the decadal increments of log of population density after dam installation for counties treated before 1950, smoothed out by locally weighted regressions (lowess). Each decadal increment is the difference between the estimated coefficient of an event-time dummy and the estimated coefficient of the event-time dummy a decade before. In panel A, the control group used in the estimation consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. In panel B, the control set contains the originally defined control counties.

Table I: Chronology of Electrification of Industry and History of the U.S. Electric Power Industry

1870: D.C. Electric Generator
__73: Motors Driven by a Generator
__81: Hydroelectric Dam Built Near Niagara Falls
__82: Modern Electric Utility Industry Launched
__83: Motors Used in Manufacturing
__88: A.C. Motor Developed
__91: A.C. Power Transmission for Industrial Use
__92: A.C. Polyphase Induction Motor Marketed
__95: A.C. Generation at Niagara Falls
1900: A.C. Generator
__01: Right-of-Way Act: Federal Govt. Entitled the Authority to Grant Permits for Hydro Projects
__02: Reclamation Act: Federal Water Development Projects Initiated
__06: General Dam Act: Private Owners Must Build and Operate Navigation Facilities W/O Compensation
__07: State Regulation of Electric Utilities Started (Georgia, New York, and Wisconsin)
__14-18: World War I
__17: Capacity and Generation of Utilities Exceeded That of Industrial Establishments
__20: Federal Water Power Act: Hydroelectric Licenses Revocable After Fifty Years
__27: Supreme Court Ruling: Only Federal Govt. Can Regulate Interstate Wholesale Power Transactions
__29: Wall Street Crash
__33: Tennessee Valley Authority (TVA) Act: Local Development Fostered by Hydroelectricity
__35: Federal Power Act: Interstate Wholesale Power Transactions Regulated
__39-45: World War II
__49: TVA authorized to construct thermal-electric power plants for commercial electricity sale
__50s: Construction of High-Voltage (230kV or more) Transmission Lines Began
__54: Atomic Energy Act: Private Development of Commercial Nuclear Power Allowed
__60: Current Level of Thermal Efficiency of Fossil-Steam Plants Reached
__63: Clean Air Act (Amendments in 1970, 1977, and 1990): Pollution Regulation Initiated
__64: Wilderness Act: Land Preservation Efforts Became Official

Sources: Devine (1983), Brown and Sedano (2004), Billington, Jackson and Melosi (2005), McNerneya, Farmer and Trancik (2011), and EIA (n.d.).

Table II: Domestic Economy: Percentage distribution of total installed electric generating capacity by producer, 1889-1961

Year	Total Capacity	Industrial Establishments	Electric Utilities
1889	100.0	55.6	44.4
1899	100.0	53.4	46.6
1902	100.0	59.6	40.6
1907	100.0	60.2	39.8
1912	100.0	53.0	47.0
1917	100.0	42.0	58.0
1920	100.0	34.6	65.4
1925	100.0	28.6	71.4
1929	100.0	22.9	77.1
1935	100.0	19.6	80.4
1939	100.0	21.4	78.6
1947	100.0	19.7	80.3
1954	100.0	13.7	86.3
1958	100.0	11.0	89.0
1961	100.0	9.0	91.0

Source: DuBoff (1979, p.39).

Table III: Summary Statistics

		All Counties		Treated - Before 1950		Treated - After 1950		Control	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Outcomes and Hydroelectricity		Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Log of Population Density	1900	2.66	*1.52*	3.28	*1.34*	2.57	*1.29*	2.43	*1.80*
	1950	3.26	*1.32*	3.94	*1.31*	3.13	*1.10*	3.04	*1.48*
	2000	3.80	*1.40*	4.74	*1.06*	3.70	*1.27*	3.41	*1.49*
Log of Employment Density	1900	1.60	*1.49*	2.17	*1.37*	1.51	*1.31*	1.41	*1.71*
	1950	2.19	*1.32*	2.85	*1.34*	2.07	*1.10*	1.99	*1.47*
	2000	2.96	*1.45*	3.94	*1.13*	2.84	*1.34*	2.58	*1.52*
Hydroelectric Capacity	1910s			1.25	.				
at Dam Completion	1920s			1.85	*0.75*				
(in 100's of megawatts)	1930s			1.48	*0.69*				
	1940s			2.05	*2.18*				
	1950s					3.21	*3.34*		
	1960s					4.45	*5.25*		
	1970s					3.43	*3.23*		
	1980s					1.70	*0.72*		
	1990s					1.47	*0.63*	Potential Capacity	
	All Years	2.97	*3.53*	1.81	*1.51*	3.48	*4.01*	2.28	*1.97*
Average Hydroelectric	1910	0.78	*0.47*	0.99	*0.23*	0.14	.		
Installed Capacity	1920	1.49	*0.87*	1.77	*0.73*	0.39	*0.09*		
Over time	1930	1.48	*0.86*	1.71	*0.81*	0.58	*0.23*		
(in 100's of megawatts)	1940	1.98	*1.83*	2.29	*1.90*	0.65	*0.25*		
	1950	2.95	*3.05*	2.96	*2.89*	2.94	*3.26*		
	1960	3.99	*4.49*	3.41	*3.66*	4.30	*4.89*		
	1970	4.49	*5.64*	3.47	*3.63*	4.99	*6.37*		
	1980	4.48	*5.54*	3.84	*3.93*	4.76	*6.13*		
	1990	4.43	*5.50*	3.89	*3.92*	4.67	*6.07*		
	2000	4.43	*5.50*	3.89	*3.92*	4.67	*6.07*		
	All Years	3.90	*4.93*	3.18	*3.31*	4.38	*5.71*		
Sample Counties		154		30		69		55	
Panel B. Dam Completion by Decade		Obs.	*Percent*	Obs.	*Percent*	Obs.	*Percent*	Obs.	*Percent*
Dam Completion Date	1910s	1	*1.01*	1	*3.33*				
	1920s	7	*7.07*	7	*23.33*				
	1930s	9	*9.09*	9	*30.00*				
	1940s	13	*13.13*	13	*43.33*				
	1950s	26	*26.26*			26	*37.68*		
	1960s	26	*26.26*			26	*37.68*		
	1970s	7	*7.07*			7	*10.14*		
	1980s	8	*8.08*			8	*11.59*		
	1990s	2	*2.02*			2	*2.90*		
	All Years	99	*100.00*	30	*100.00*	69	*100.00*		

Notes : Panel A reports mean and standard deviation (SD) for my main outcomes - population density and employment density - for selected years (see Figures A1 and A2 for more details), and for hydroelectricity-related variables throughout the twentieth century. Standard deviations are shown in italic. That panel also provides the sample size for each group of counties: all counties in the sample ("All Counties"), counties that have dams built before 1950 ("Treated - Before 1950"), counties that have dams built after 1950 ("Treated - After 1950"), and counties that have hydropower potential but no hydroelectric facilities ("Control"). Panel B presents the distribution of dam completion by decade. For each group of counties, the first column reports the number of counties ("Obs."), and the second column the corresponding percentage ("Percent").

Table IV: Short- and Long-Run Effects of Hydro Dams on Population Density

ln(Pop Density)	TS_b1950	TC_b1950	TS_a1950	TC_a1950	Diff. (TS)	Diff. (TC)
	(1)	(2)	(3)	(4)	(5) = (1) - (3)	(6) = (2) - (4)
80 years before dam			-0.4335	-0.1652		
			(0.3306)	(0.3564)		
70 years before dam			-0.1137	0.0802		
			(0.2182)	(0.2500)		
60 years before dam			-0.1468	0.0411		
			(0.1605)	(0.1914)		
50 years before dam			-0.2245	-0.1100		
			(0.1353)	(0.1620)		
40 years before dam	-0.2428	-0.1441	-0.2243*	-0.1298	-0.0185	-0.0143
	(0.2365)	(0.2740)	(0.1193)	(0.1428)	(0.2061)	(0.2523)
30 years before dam	-0.2564	-0.1493	-0.1895*	-0.1252	-0.0670	-0.0240
	(0.1893)	(0.1941)	(0.1124)	(0.1315)	(0.1743)	(0.1936)
20 years before dam	-0.1536	-0.1253	-0.1663	-0.1180	0.0126	-0.0073
	(0.1528)	(0.1740)	(0.1070)	(0.1218)	(0.1510)	(0.1778)
10 years before dam	-0.0398	-0.0337	-0.1468	-0.1049	0.1070	0.0712
	(0.1445)	(0.1544)	(0.1036)	(0.1126)	(0.1459)	(0.1617)
10 years after dam	0.1103**	0.0652	-0.0014	0.0070	0.1117***	(0.0582)
	(0.0451)	(0.0461)	(0.0241)	(0.0273)	(0.0390)	(0.0443)
20 years after dam	0.3063***	0.1574*	0.0355	0.0491	0.2708***	0.1083
	(0.0979)	(0.0811)	(0.0439)	(0.0475)	(0.0831)	(0.0783)
30 years after dam	0.4129***	0.2207**	0.0460	0.0721	0.3670***	0.1485
	(0.1131)	(0.0942)	(0.0646)	(0.0688)	(0.1021)	(0.0968)
40 years after dam	0.6551***	0.4118***	0.1053	0.1320	0.5497***	0.2798**
	(0.1408)	(0.1138)	(0.0893)	(0.0900)	(0.1382)	(0.1193)
50 years after dam	0.7897***	0.5176***	0.1600	0.2242	0.6298***	0.2934*
	(0.1681)	(0.1293)	(0.1340)	(0.1358)	(0.1845)	(0.1554)
60 years after dam	0.8699***	0.6036***				
	(0.2047)	(0.1514)				
70 years after dam	0.9584***	0.6026***				
	(0.2454)	(0.1663)				
80 years after dam	1.0494***	0.6448***				
	(0.3052)	(0.2094)				
Cheap Local Hydroelectricity Effect					0.2579	0.1407
Observations	660	935	1518	1364		
R-squared	0.9837	0.9781	0.9634	0.9634		

Notes : This table presents the short- and long-run effects of hydroelectric dams on population density. The estimated coefficients are the β's in equation (5) in the text. They are coefficients of event-time dummies. "TS" in the labels of the columns represents treatment vs. synthetic control, "TC" treatment vs. control, "b1950" before 1950, and "a1950" after 1950. The synthetic control group consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. The control group contains the originally defined control counties. The cheap local hydroelectricity effects displayed at the bottom of columns 5 and 6 represent average growth of population density from the time of dam installation until 1950, as expressed in equation (7) in the text. Therefore, to obtain the estimates of agglomeration spillovers, simply subtract the cheap local hydroelectricity effect from the post-dam effects in columns 5 and 6, as expressed in equation (6) in the text. Standard errors are shown in parentheses. In "TS" columns, they are clustered at a case level. A case is a pair of a treated county and its corresponding synthetic control. In "TC" columns, they are clustered at the county level. "Observations" reports the number of county-year observations.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Table V: The Impact of Hydro Dams on Population Density - Specification Checks

ln(Pop Density)	Treatment vs. Synthetic Control				Treatment vs. Control				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
40 years before dam	-0.3703	-0.1265	-0.2428	-0.1369	-0.2610	-0.1744	-0.1441	-0.1235	-0.2542
	(0.2624)	(0.2346)	(0.2365)	(0.1784)	(0.2091)	(0.2911)	(0.2740)	(0.2556)	(0.4039)
30 years before dam	-0.3802**	-0.1692	-0.2564	-0.2974*	-0.2387**	-0.1763	-0.1493	-0.1758	-0.2761
	(0.1521)	(0.2003)	(0.1893)	(0.1459)	(0.1136)	(0.2129)	(0.1941)	(0.1921)	(0.2927)
20 years before dam	-0.2764***	-0.0565	-0.1536	-0.2122	-0.2001***	-0.1457	-0.1253	-0.1516	-0.1652
	(0.0934)	(0.1785)	(0.1528)	(0.1275)	(0.0684)	(0.2005)	(0.1740)	(0.1772)	(0.2679)
10 years before dam	-0.1542***	0.0471	-0.0398	-0.1048	-0.0964**	-0.0482	-0.0337	-0.0758	-0.1240
	(0.0472)	(0.1778)	(0.1445)	(0.1208)	(0.0409)	(0.1799)	(0.1544)	(0.1634)	(0.2220)
10 years after dam	0.1193**	0.1062**	0.1103**	0.1041**	0.0675*	0.0744	0.0652	0.0702	0.0448
	(0.0482)	(0.0471)	(0.0451)	(0.0451)	(0.0363)	(0.0502)	(0.0461)	(0.0456)	(0.0655)
20 years after dam	0.3206***	0.2770***	0.3063***	0.3070***	0.1854***	0.1757*	0.1574*	0.1772**	0.2106*
	(0.0905)	(0.0918)	(0.0979)	(0.1016)	(0.0664)	(0.0887)	(0.0811)	(0.0831)	(0.1194)
30 years after dam	0.4380***	0.4022***	0.4129***	0.3978***	0.2644***	0.2682**	0.2207**	0.2467***	0.3093*
	(0.1264)	(0.1206)	(0.1131)	(0.1044)	(0.0905)	(0.1129)	(0.0942)	(0.0934)	(0.1592)
40 years after dam	0.6079***	0.5985***	0.6551***	0.6209***	0.4000***	0.4197***	0.4118***	0.4409***	0.4628**
	(0.1595)	(0.1557)	(0.1408)	(0.1292)	(0.1104)	(0.1364)	(0.1138)	(0.1131)	(0.1770)
50 years after dam	0.7150***	0.7512***	0.7897***	0.7335***	0.4815***	0.5210***	0.5176***	0.5514***	0.5331***
	(0.1899)	(0.1862)	(0.1681)	(0.1509)	(0.1309)	(0.1518)	(0.1293)	(0.1272)	(0.2016)
60 years after dam	0.7850***	0.8535***	0.8699***	0.7999***	0.5142***	0.5726***	0.6036***	0.6374***	0.5882**
	(0.2245)	(0.2272)	(0.2047)	(0.1858)	(0.1467)	(0.1718)	(0.1514)	(0.1477)	(0.2312)
70 years after dam	0.8614***	0.9804***	0.9584***	0.8767***	0.4990***	0.5936***	0.6026***	0.6427***	0.5839**
	(0.2737)	(0.2823)	(0.2454)	(0.2145)	(0.1783)	(0.1997)	(0.1663)	(0.1681)	(0.2674)
80 years after dam	0.9652***	1.0877***	1.0494***	0.9330***	0.4886**	0.6030**	0.6448***	0.6937***	0.6253**
	(0.3293)	(0.3398)	(0.3052)	(0.3371)	(0.2227)	(0.2562)	(0.2094)	(0.2089)	(0.2983)
Hydro Dam Size		0.2180	0.1263	0.0442		0.0754	0.0807	0.0420	0.0002
		(0.1572)	(0.1217)	(0.1138)		(0.1416)	(0.1193)	(0.1314)	(0.1942)
Hydro Dam Size^2		-0.0443*	-0.0260	-0.0097		-0.0221	-0.0210	-0.0160	-0.0071
		(0.0217)	(0.0167)	(0.0158)		(0.0185)	(0.0153)	(0.0174)	(0.0273)
Hydro Dam Size^3		0.0019**	0.0011*	0.0004		0.0011	0.0010*	0.0008	0.0005
		(0.0008)	(0.0006)	(0.0006)		(0.0007)	(0.0006)	(0.0006)	(0.0010)
Thermal Plant Size			0.1936***	0.1826***			0.1297***	0.1154***	0.1185***
			(0.0401)	(0.0386)			(0.0273)	(0.0284)	(0.0417)
Thermal Plant Size^2			-0.0154***	-0.0146***			-0.0088***	-0.0078***	-0.0078**
			(0.0029)	(0.0031)			(0.0021)	(0.0023)	(0.0032)
Thermal Plant Size^3			0.0003***	0.0003***			0.0001***	0.0001***	0.0001**
			(0.0001)	(0.0001)			(0.0000)	(0.0000)	(0.0001)
Market Access-by-Year	No	No	No	Yes	No	No	No	Yes	No
Region-by-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
State-by-Year FE	No	No	No	No	No	No	No	No	Yes
Observations	660	660	660	660	935	935	935	913	935
R-squared	0.9787	0.9801	0.9837	0.9875	0.9745	0.9751	0.9781	0.9796	0.9872

Notes: This table presents some robustness checks regarding the specification used in the estimation of the effects of pre-1950 hydroelectric dams on population density. The synthetic control group consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. The control group contains the originally defined control counties. Columns 1 and 5 display estimates related to the basic specification. Such specification includes event-time dummies, county effects, region-by-year fixed effects, and time-invariant county characteristics (cubic function in latitute and longitude, and 50-year average rainfall and 50-year average temperature for each season of the year) interacted with year effects. Columns 2 and 6 show estimates associated with the basic specification plus a cubic function in dam capacity. Columns 3 and 7 report the estimates related to the main specification, which is basic specification plus controls for dam size plus a cubic function in thermal power plant capacity. Columns 4 and 8 report the estimates related to the main specification plus controls for the interaction of year effects with three county-specific measures of market access: mileage of railroad tracks, distance to closest waterway, and log of market access as estimated by Donaldson and Hornbeck (2012). Column 9 displays estimates associated with the main specification, but replacing region-by-year fixed effects with state-by-year fixed effects. A similar column does not appear under the label "Treatment vs. Synthetic Control" because the sample size was insufficient to estimate all the parameters. Standard errors are shown in parentheses. In columns 1 through 4, they are clustered at a case level. A case is a pair of a treated county and its corresponding synthetic control. In columns 5 through 9, they are clustered at the county level. "Observations" reports the number of county-year observations.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Table VI: Short- and Long-Run Effects of Hydro Dams on Employment Density

ln(Pop Density)	TS_b1950	TC_b1950	TS_a1950	TC_a1950
	(1)	(2)	(3)	(4)
80 years before dam			-0.3185	-0.1612
			(0.3929)	(0.3784)
70 years before dam			-0.2408	-0.1090
			(0.2599)	(0.2778)
60 years before dam			-0.1482	0.0001
			(0.1934)	(0.2223)
50 years before dam			-0.2665	-0.1751
			(0.1626)	(0.1941)
40 years before dam	-0.0411	-0.0150	-0.2800*	-0.1905
	(0.2235)	(0.3112)	(0.1443)	(0.1706)
30 years before dam	-0.1079	-0.0682	-0.2628*	-0.1758
	(0.1788)	(0.2291)	(0.1338)	(0.1579)
20 years before dam	-0.0402	-0.0538	-0.2515*	-0.1605
	(0.1593)	(0.2221)	(0.1302)	(0.1494)
10 years before dam	0.0478	0.0502	-0.2377*	-0.1343
	(0.1459)	(0.1903)	(0.1199)	(0.1355)
10 years after dam	0.0650	0.0754	0.0145	0.0184
	(0.0466)	(0.0568)	(0.0291)	(0.0306)
20 years after dam	0.2430**	0.1881*	0.0395	0.0552
	(0.0992)	(0.0995)	(0.0526)	(0.0571)
30 years after dam	0.3491***	0.2824**	0.0606	0.0897
	(0.1116)	(0.1212)	(0.0764)	(0.0823)
40 years after dam	0.5681***	0.4770***	0.1188	0.1593
	(0.1230)	(0.1480)	(0.1022)	(0.1075)
50 years after dam	0.6854***	0.6059***	0.1840	0.2710
	(0.1375)	(0.1664)	(0.1729)	(0.1645)
60 years after dam	0.7624***	0.7200***		
	(0.1592)	(0.1931)		
70 years after dam	0.8425***	0.7152***		
	(0.1885)	(0.2127)		
80 years after dam	0.8619***	0.7401***		
	(0.2580)	(0.2613)		
Observations	660	935	1496	1353
R-squared	0.9865	0.9677	0.9554	0.9509

Notes : This table presents the short- and long-run effects of hydroelectric dams on employment density. The estimated coefficients are the β's in equation (5) in the text. They are coefficients of event-time dummies. "TS" in the labels of the columns represents treatment vs. synthetic control, "TC" treatment vs. control, "b1950" before 1950, and "a1950" after 1950. The synthetic control group consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. The control group contains the originally defined control counties. Standard errors are shown in parentheses. In "TS" columns, they are clustered at a case level. A case is a pair of a treated county and its corresponding synthetic control. In "TC" columns, they are clustered at the county level. "Observations" reports the number of county-year observations.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table VII: The Impact of Hydro Dams on Employment Density - Sectoral Decomposition

ln(density)	Total		Manufacturing		Agriculture		Trade		Construction		Other Sectors	
	TS_b1950	TC_b1950	TS_b1950	TC_b1950	TS_b1950	TC_b1950	TS_b1950	TC_b1950	TS_b1950	TC_b1950	TS_b1950	TC_b1950
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
40 years before dam	-0.0411	-0.0150	0.2843	-0.3996	0.2905	0.4334*	0.8197	0.0939	-1.2201**	-1.0967**	-0.3724	-0.1949
	(0.2235)	(0.3112)	(0.5156)	(0.4122)	(0.2412)	(0.2285)	(0.8111)	(0.5670)	(0.4166)	(0.4404)	(0.3732)	(0.4240)
30 years before dam	-0.1079	-0.0682	0.3037	-0.2043	0.0111	0.1761	0.0924	-0.2169	-1.0684**	-0.7967	-0.1229	0.1461
	(0.1788)	(0.2291)	(0.4676)	(0.3264)	(0.1888)	(0.1552)	(0.4330)	(0.4236)	(0.4097)	(0.4836)	(0.3364)	(0.3350)
20 years before dam	-0.0402	-0.0538	0.0744	-0.2853	0.0096	0.1189	-0.0308	-0.0443	-0.6780*	-0.6156*	0.0320	0.2003
	(0.1593)	(0.2221)	(0.4837)	(0.4319)	(0.2291)	(0.1645)	(0.5193)	(0.4480)	(0.3693)	(0.3556)	(0.3015)	(0.3257)
10 years before dam	0.0478	0.0502	0.0858	-0.3795	0.0655	0.1977	-0.0817	-0.1150	-0.5398*	-0.2376	0.1602	0.2753
	(0.1459)	(0.1903)	(0.4006)	(0.3589)	(0.1941)	(0.1564)	(0.3564)	(0.3820)	(0.2545)	(0.3344)	(0.3035)	(0.2792)
10 years after dam	0.0650	0.0754	-0.1271	-0.0049	0.1588***	0.1242**	-0.0275	-0.0031	-0.3605	-0.3758	0.1697**	0.0352
	(0.0466)	(0.0568)	(0.1474)	(0.1562)	(0.0525)	(0.0495)	(0.2096)	(0.1341)	(0.5112)	(0.3130)	(0.0736)	(0.0834)
20 years after dam	0.2430**	0.1881*	0.0180	0.1685	0.2210***	0.0687	0.0034	0.0243	-0.1851	-0.4518	0.3346**	0.1611
	(0.0992)	(0.0995)	(0.1706)	(0.1501)	(0.0799)	(0.0865)	(0.2335)	(0.1695)	(0.5102)	(0.3347)	(0.1247)	(0.1256)
30 years after dam	0.3491***	0.2824**	0.1280	0.1595	0.3322**	0.1780	0.1099	0.1435	-0.1930	-0.3900	0.4488**	0.2096
	(0.1116)	(0.1212)	(0.1902)	(0.1655)	(0.1248)	(0.1149)	(0.2518)	(0.2101)	(0.5476)	(0.3751)	(0.1860)	(0.1654)
40 years after dam	0.5681***	0.4770***	0.3574	0.2970	0.3883**	0.2144	0.1982	0.3247	-0.1815	-0.3033	0.6422**	0.4088*
	(0.1230)	(0.1480)	(0.2083)	(0.1989)	(0.1727)	(0.1344)	(0.2423)	(0.2362)	(0.5997)	(0.3790)	(0.2387)	(0.2056)
50 years after dam	0.6854***	0.6059***	0.3629	0.2743	0.8084***	0.5501***	0.1887	0.3993	-0.0926	-0.1901	0.7627**	0.5287**
	(0.1375)	(0.1664)	(0.2185)	(0.2161)	(0.1822)	(0.1736)	(0.2772)	(0.2618)	(0.5904)	(0.3994)	(0.3152)	(0.2329)
60 years after dam	0.7624***	0.7200***	0.5148**	0.3589	0.8164***	0.5106***	0.2661	0.5141*	-0.1168	-0.1226	0.8168**	0.6122**
	(0.1592)	(0.1931)	(0.2384)	(0.2269)	(0.2003)	(0.1482)	(0.3186)	(0.2911)	(0.6536)	(0.4566)	(0.3698)	(0.2583)
70 years after dam	0.8425***	0.7152***	0.5927**	0.3369	1.0279***	0.5492***	0.2516	0.5389*	0.0223	-0.0603	0.8134*	0.5726*
	(0.1885)	(0.2127)	(0.2660)	(0.2350)	(0.2035)	(0.2002)	(0.3357)	(0.2967)	(0.6954)	(0.4838)	(0.4597)	(0.3047)
80 years after dam	0.8619***	0.7401***	0.6723**	0.3523	0.9216***	0.4202**	0.2674	0.5552	0.4266	0.0024	0.7560	0.5604
	(0.2580)	(0.2613)	(0.2973)	(0.2714)	(0.2086)	(0.2103)	(0.3856)	(0.4014)	(0.6146)	(0.4821)	(0.4635)	(0.3488)
Observations	660	935	506	638	638	924	440	682	352	528	660	902
R-squared	0.9865	0.9677	0.9715	0.9471	0.9664	0.9525	0.9803	0.9536	0.9817	0.9526	0.9693	0.9387

Notes : This table presents the effects of pre-1950 hydroelectric dams on employment density by sector. The estimated coefficients are the β's in equation (5) in the text. They are coefficients of event-time dummies. "Other sectors" includes employment in industries that provide local goods and services like real estate, cleaning services, legal services, medical services, and personal services. "TS" in the labels of the columns represents treatment vs. synthetic control, "TC" treatment vs. control, and "b1950" before 1950. The synthetic control group consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. The control group contains the originally defined control counties. Standard errors are shown in parentheses. In "TS" columns, they are clustered at a case level. A case is a pair of a treated county and its corresponding synthetic control. In "TC" columns, they are clustered at the county level. "Observations" reports the number of county-year observations.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table VIII: The Impact of Hydro Dams on Population Density of Contiguous Counties

ln(Pop Density)	Main Analysis		Treatment: Contiguous Counties	
	TS_b1950	TC_b1950	TS_b1950	TC_b1950
	(1)	(2)	(3)	(4)
40 years before dam	-0.2428	-0.1441	-0.9443***	-0.7819***
	(0.2365)	(0.2740)	(0.1826)	(0.1803)
30 years before dam	-0.2564	-0.1493	-0.7571***	-0.6644***
	(0.1893)	(0.1941)	(0.1597)	(0.1537)
20 years before dam	-0.1536	-0.1253	-0.6404***	-0.5639***
	(0.1528)	(0.1740)	(0.1294)	(0.1282)
10 years before dam	-0.0398	-0.0337	-0.4642***	-0.4180***
	(0.1445)	(0.1544)	(0.1011)	(0.1015)
10 years after dam	0.1103**	0.0652	0.0633*	0.0307
	(0.0451)	(0.0461)	(0.0322)	(0.0288)
20 years after dam	0.3063***	0.1574*	0.1234***	0.0702
	(0.0979)	(0.0811)	(0.0463)	(0.0426)
30 years after dam	0.4129***	0.2207**	0.1759***	0.0996*
	(0.1131)	(0.0942)	(0.0591)	(0.0548)
40 years after dam	0.6551***	0.4118***	0.2080***	0.1096*
	(0.1408)	(0.1138)	(0.0669)	(0.0619)
50 years after dam	0.7897***	0.5176***	0.2444***	0.1222*
	(0.1681)	(0.1293)	(0.0734)	(0.0671)
60 years after dam	0.8699***	0.6036***	0.2552***	0.1123
	(0.2047)	(0.1514)	(0.0765)	(0.0735)
70 years after dam	0.9584***	0.6026***	0.2439***	0.0919
	(0.2454)	(0.1663)	(0.0795)	(0.0798)
80 years after dam	1.0494***	0.6448***	0.2127**	0.0249
	(0.3052)	(0.2094)	(0.0973)	(0.1118)
Observations	660	935	2222	1430
R-squared	0.9837	0.9781	0.9697	0.9656

Notes : This table presents the effects of pre-1950 hydroelectric dams on population density of neighboring counties. The estimated coefficients are the β's in equation (5) in the text. They are coefficients of event-time dummies. "TS" in the labels of the columns represents treatment vs. synthetic control, "TC" treatment vs. control, and "b1950" before 1950. The synthetic control group consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. The control group contains the originally defined control counties. Standard errors are shown in parentheses. In "TS" columns, they are clustered at a case level. A case is a pair of a treated county and its corresponding synthetic control. In "TC" columns, they are clustered at the county level. "Observations" reports the number of county-year observations.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table IX: The Impact of Smaller (30-100MW) Hydro Dams on Population Density

ln(Pop Density)	Main Analysis		Treatment:		Treatment:	
	TS_b1950	TC_b1950	TS_b1950	TC_b1950	TS_b1950	TC_b1950
	(1)	(2)	(3)	(4)	(5)	(6)
40 years before dam	-0.2428	-0.1441	-0.1861	-0.3683*	-0.2811	-0.3470
	(0.2365)	(0.2740)	(0.2068)	(0.2185)	(0.2680)	(0.2700)
30 years before dam	-0.2564	-0.1493	-0.3860*	-0.2614	-0.3597*	-0.3021
	(0.1893)	(0.1941)	(0.1935)	(0.2057)	(0.2056)	(0.2317)
20 years before dam	-0.1536	-0.1253	-0.3576**	-0.2421	-0.3729**	-0.3324
	(0.1528)	(0.1740)	(0.1624)	(0.1849)	(0.1836)	(0.2016)
10 years before dam	-0.0398	-0.0337	-0.3651**	-0.1681	-0.3035*	-0.2142
	(0.1445)	(0.1544)	(0.1489)	(0.1687)	(0.1654)	(0.1794)
10 years after dam	0.1103**	0.0652	0.0689**	0.0753**	0.0290	0.0488
	(0.0451)	(0.0461)	(0.0328)	(0.0289)	(0.0302)	(0.0315)
20 years after dam	0.3063***	0.1574*	0.0959	0.1232**	0.0593	0.0816*
	(0.0979)	(0.0811)	(0.0577)	(0.0475)	(0.0440)	(0.0476)
30 years after dam	0.4129***	0.2207**	0.0847	0.1453**	0.1319**	0.1345**
	(0.1131)	(0.0942)	(0.0691)	(0.0614)	(0.0590)	(0.0635)
40 years after dam	0.6551***	0.4118***	0.1584*	0.1925**	0.2565***	0.2105***
	(0.1408)	(0.1138)	(0.0809)	(0.0744)	(0.0728)	(0.0794)
50 years after dam	0.7897***	0.5176***	0.2189**	0.2487***	0.3514***	0.2709***
	(0.1681)	(0.1293)	(0.0882)	(0.0838)	(0.0819)	(0.0921)
60 years after dam	0.8699***	0.6036***	0.2136**	0.2231**	0.4093***	0.2848**
	(0.2047)	(0.1514)	(0.0971)	(0.0941)	(0.0956)	(0.1082)
70 years after dam	0.9584***	0.6026***	0.2552**	0.2226**	0.4958***	0.3042**
	(0.2454)	(0.1663)	(0.1050)	(0.1106)	(0.1106)	(0.1233)
80 years after dam	1.0494***	0.6448***	0.2748**	0.1535	0.5497***	0.2714*
	(0.3052)	(0.2094)	(0.1337)	(0.1341)	(0.1317)	(0.1445)
Observations	660	935	968	1584	968	946
R-squared	0.9837	0.9781	0.9880	0.9651	0.9869	0.9833

Notes: This table presents the effects of smaller pre-1950 hydroelectric dams on population density. Small dams refers to dams with hydropower capacity between 30 and 100 megawatts (MW). The estimated coefficients are the β's in equation (5) in the text. They are coefficients of event-time dummies. "TS" in the labels of the columns represents treatment vs. synthetic control, "TC" treatment vs. control, and "b1950" before 1950. The synthetic control group consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. The control group contains the originally defined control counties. Standard errors are shown in parentheses. In "TS" columns, they are clustered at a case level. A case is a pair of a treated county and its corresponding synthetic control. In "TC" columns, they are clustered at the county level. "Observations" reports the number of county-year observations.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table X: The Impact of Hydro Dams on Population Density - All vs. Non-TVA Counties

ln(Pop Density)	*With* TVA Counties				*Without* TVA Counties			
	TS_b1950	TC_b1950	TS_a1950	TC_a1950	TS_b1950	TC_b1950	TS_a1950	TC_a1950
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
80 years before dam			-0.4335	-0.1652			-0.3578	-0.1590
			(0.3306)	(0.3564)			(0.3508)	(0.3778)
70 years before dam			-0.1137	0.0802			-0.0214	0.1115
			(0.2182)	(0.2500)			(0.2374)	(0.2708)
60 years before dam			-0.1468	0.0411			-0.0880	0.0520
			(0.1605)	(0.1914)			(0.1753)	(0.2093)
50 years before dam			-0.2245	-0.1100			-0.1780	-0.1062
			(0.1353)	(0.1620)			(0.1490)	(0.1786)
40 years before dam	-0.2428	-0.1441	-0.2243*	-0.1298	-0.6014	-0.1867	-0.1838	-0.1197
	(0.2365)	(0.2740)	(0.1193)	(0.1428)	(0.3714)	(0.4076)	(0.1330)	(0.1588)
30 years before dam	-0.2564	-0.1493	-0.1895*	-0.1252	-0.5027*	-0.1490	-0.1616	-0.1176
	(0.1893)	(0.1941)	(0.1124)	(0.1315)	(0.2690)	(0.2313)	(0.1273)	(0.1475)
20 years before dam	-0.1536	-0.1253	-0.1663	-0.1180	-0.2473	-0.0883	-0.1457	-0.1119
	(0.1528)	(0.1740)	(0.1070)	(0.1218)	(0.1884)	(0.2034)	(0.1250)	(0.1383)
10 years before dam	-0.0398	-0.0337	-0.1468	-0.1049	-0.0446	0.0242	-0.1343	-0.0986
	(0.1445)	(0.1544)	(0.1036)	(0.1126)	(0.1632)	(0.1784)	(0.1214)	(0.1296)
10 years after dam	0.1103**	0.0652	-0.0014	0.0070	0.1540**	0.0673	-0.0164	-0.0073
	(0.0451)	(0.0461)	(0.0241)	(0.0273)	(0.0716)	(0.0594)	(0.0268)	(0.0297)
20 years after dam	0.3063***	0.1574*	0.0355	0.0491	0.2341**	0.0393	0.0017	0.0166
	(0.0979)	(0.0811)	(0.0439)	(0.0475)	(0.1040)	(0.0868)	(0.0502)	(0.0509)
30 years after dam	0.4129***	0.2207**	0.0460	0.0721	0.3411**	0.0896	-0.0099	0.0239
	(0.1131)	(0.0942)	(0.0646)	(0.0688)	(0.1386)	(0.1153)	(0.0730)	(0.0734)
40 years after dam	0.6551***	0.4118***	0.1053	0.1320	0.5950***	0.2966**	0.0450	0.0782
	(0.1408)	(0.1138)	(0.0893)	(0.0900)	(0.1788)	(0.1411)	(0.0965)	(0.0943)
50 years after dam	0.7897***	0.5176***	0.1600	0.2242	0.7340***	0.4361***	0.0194	0.1065
	(0.1681)	(0.1293)	(0.1340)	(0.1358)	(0.2331)	(0.1571)	(0.1350)	(0.1391)
60 years after dam	0.8699***	0.6036***			0.8031**	0.5518***		
	(0.2047)	(0.1514)			(0.2824)	(0.1831)		
70 years after dam	0.9584***	0.6026***			0.7464**	0.4712**		
	(0.2454)	(0.1663)			(0.3142)	(0.1880)		
80 years after dam	1.0494***	0.6448***			0.7673*	0.5060**		
	(0.3052)	(0.2094)			(0.4040)	(0.2337)		
Observations	660	935	1518	1364	440	792	1364	1254
R-squared	0.9837	0.9781	0.9634	0.9634	0.9914	0.9798	0.9637	0.9638

Notes : This table presents the short- and long-run effects of hydroelectric dams on population density, with and without counties of the Tennessee Valley Authority (TVA) region. The TVA was one of the most ambitious place based economic development policies in the history of the U.S. Besides hydroelectric dams, it included construction of an extensive network of new roads, canals, and flood control systems. The estimated coefficients are the β's in equation (5) in the text. They are coefficients of event-time dummies. "TS" in the labels of the columns represents treatment vs. synthetic control, "TC" treatment vs. control, "b1950" before 1950, and "a1950" after 1950. The synthetic control group consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. The control group contains the originally defined control counties. Standard errors are shown in parentheses. In "TS" columns, they are clustered at a case level. A case is a pair of a treated county and its corresponding synthetic control. In "TC" columns, they are clustered at the county level. "Observations" reports the number of county-year observations.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table XI: The Impact of Hydro Dams on Population Density - All vs. No Control Counties With Environmental Regulations

ln(Pop Density)	Treatment vs. Synthetic Control			Treatment vs. Control		
	(1)	(2)	(3)	(4)	(5)	(6)
40 years before dam	-0.2428	-0.2557	-0.1249	-0.1441	-0.1739	0.0777
	(0.2365)	(0.4924)	(0.4001)	(0.2740)	(0.4194)	(0.2881)
30 years before dam	-0.2564	-0.3550	-0.2534	-0.1493	-0.2331	-0.0257
	(0.1893)	(0.3492)	(0.3044)	(0.1941)	(0.3100)	(0.2075)
20 years before dam	-0.1536	-0.2459	-0.2398	-0.1253	-0.1729	-0.0685
	(0.1528)	(0.2585)	(0.2512)	(0.1740)	(0.2502)	(0.1929)
10 years before dam	-0.0398	-0.1315	-0.1552	-0.0337	-0.0794	-0.0370
	(0.1445)	(0.1887)	(0.2004)	(0.1544)	(0.2004)	(0.1669)
10 years after dam	0.1103**	0.1112	0.0877	0.0652	0.1013	0.0406
	(0.0451)	(0.0852)	(0.0605)	(0.0461)	(0.0838)	(0.0526)
20 years after dam	0.3063***	0.3030	0.2345*	0.1574*	0.2202	0.0960
	(0.0979)	(0.1792)	(0.1219)	(0.0811)	(0.1551)	(0.0882)
30 years after dam	0.4129***	0.3987*	0.2708*	0.2207**	0.3076	0.1214
	(0.1131)	(0.2282)	(0.1445)	(0.0942)	(0.2139)	(0.1026)
40 years after dam	0.6551***	0.6321**	0.4334**	0.4118***	0.5196*	0.2309*
	(0.1408)	(0.2705)	(0.1728)	(0.1138)	(0.2690)	(0.1172)
50 years after dam	0.7897***	0.7780**	0.5411**	0.5176***	0.6443**	0.3132**
	(0.1681)	(0.3182)	(0.1964)	(0.1293)	(0.3183)	(0.1258)
60 years after dam	0.8699***	0.9274**	0.6321***	0.6036***	0.7415*	0.3773***
	(0.2047)	(0.3709)	(0.2262)	(0.1514)	(0.3698)	(0.1430)
70 years after dam	0.9584***	1.0502**	0.7215***	0.6026***	0.8035*	0.3799**
	(0.2454)	(0.4392)	(0.2424)	(0.1663)	(0.4110)	(0.1673)
80 years after dam	1.0494***	1.2300**	0.8273***	0.6448***	0.9041*	0.3691
	(0.3052)	(0.5057)	(0.2766)	(0.2094)	(0.4583)	(0.2468)
Observations	660	660	660	935	462	1892
R-squared	0.9837	0.9858	0.9818	0.9781	0.9861	0.9559

Notes: This table presents the short- and long-run effects of hydroelectric dams on population density, with and without control counties subject to land regulations aimed at preserving wilderness and wildlife. Such environmental regulations could restrict population growth in those locations. The estimated coefficients are the β's in equation (5) in the text. The synthetic control group consists of synthetic control counties. Each synthetic control county is a weighted average of the originally defined control counties that reproduces more closely the counterfactual outcome trajectory that the affected county would have experienced in the absence of hydro dams. The control group contains the originally defined control counties. Columns 1 and 4 display the estimates of the main analysis, which includes control counties subject to environmental regulations. Columns 2 and 5 show estimates associated with synthetic control and control counties, respectively, without environmental regulations. Columns 3 and 6 report estimates related to synthetic control and control counties, respectively, without environmental regulations but with hydropower potential starting at 10 megawatts (instead of 100 megawatts, as in the main analysis). Standard errors are shown in parentheses. In columns 1 through 3, they are clustered at a case level. A case is a pair of a treated county and its corresponding synthetic control. In columns 4 through 6, they are clustered at the county level. "Observations" reports the number of county-year observations.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.