Title
Essays on State Capacity and Local Public Goods

Author
Schoenholzer, David

Publication Date
2018

Peer reviewed|Thesis/dissertation

Essays on State Capacity and Local Public Goods

by

David S. Schoenholzer

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Fred Finan, Chair
Professor Ernesto Dal Bó
Professor Victor Couture
Professor Patrick Kline
Professor Edward Miguel

Spring 2018

Essays on State Capacity and Local Public Goods

Abstract

Essays on State Capacity and Local Public Goods

by

David S. Schoenholzer

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Fred Finan, Chair

The availability of public goods – public safety, infrastructure, rule enforcement, education, and others – is one of the hallmarks of a well-functioning society. These public goods are tightly linked to the state, defined as the institutionalized hierarchy of public decision-making (Brumfiel, 1994): only when the state attains sufficient capacity, these public goods are effectively provided. However, while the state plays an indispensable role in the provision of public goods, there is no reason to believe it does so efficiently. Thus, developing new ways to evaluate the role of the state in efficient public goods provision and long-run development is a core theme of this dissertation. Moreover, given the enormous variability across the globe in the capacity of the state to fulfill its core mandate of public goods provision, understanding the development of the state and its linkages to the economic environment continues to pose an important and understudied challenge to scholars in political economy, public economics, economic history and development economics.

In the first chapter of this dissertation, in joint work with Calvin Zhang, we study the role of the state in providing public goods through local governments. Across counties, cities, special districts and school districts, the U.S. has almost 100,000 local governments employing more than ten million people and making up more than 10% of the U.S. economy (U.S. Census Bureau Government Division, 2013). Strikingly, local state capacity is distributed highly unevenly, especially between those within municipal boundaries and those in unincorporated parts of cities. 36 million people live in unincorporated communities without separate municipal government, instead being served by counties and special districts. These places typically have limited local electoral representation, lower levels of public infrastructure, weaker code enforcement and poorer public safety provision (Anderson, 2008). We study whether local state services are underprovided in this context. To do so, we combine

administrative boundary changes from Californian cities with the universe of individual real estate sales prices for the state over the years 1988-2013. In this way, we can estimate the change in residential home prices and home construction activity in the aftermath of municipal annexation. We then interpret these changes in a spatial equilibrium model with heterogeneous households to estimate households' willingness to pay for public goods provided by local governments. We find that households value a dollar of public goods expenditures by more than a dollar, despite most of the benefits of annexation accruing to landowners and developers.

Given that the state is an important element of public goods provision, it is natural to wonder why the state has developed sooner and more effectively in some places than in others. While there is a large literature across the social sciences on the causes of early state formation, two key puzzles remain: first, why did states first form in peculiar locations like Mesopotamia, the Nile Valley, or the Valley of Mexico and not elsewhere? Second, why did incipient subjects accept the extraction by the state instead of evading its power? In the second chapter of my dissertation, I answer these questions using an old idea from cultural anthropology: states arose in regions that offered no refuge to dissidents, such as lush river valleys circumscribed by deserts, mountains, or the ocean (Carneiro, 1970). To evaluate this idea quantitatively, I collect data on archaeological excavation sites relating to early states and combine these sites with a large array of agricultural, climatic and other environmental datasets. I then show that the location of early state sites is closely associated with high land quality but low surrounding land quality.

After the initial formation of the state, it began its slow but inexorable conquest of human societies across the globe. This process was rapidly accelerated with the development of the state in Europe in the course of the middle ages. In the third and final chapter of my dissertation, in joint work with Eric Weese, we ask: what was the role of the European state in the economic growth unleashed in the the run-up to the Industrial Revolution? To this end, we employ a newly available dataset showing every single boundary change of all European states between AD 1000-2000, amounting to about 9,000 boundary changes. We combine these boundary changes with data on urban growth across Europe (Bairoch et al., 1988). Doing so, we find that cities that were subject to more changes in sovereigns saw significantly lower population growth than cities that were subject to a more stable state. In counterfactual simulations, we establish that the urban population in Europe would have been around 9% larger if European states had been more stable, offering an environment with more effective public goods and more conducive to long-run growth.

This dissertation is dedicated to: my grandfather Hans Schönholzer, whose love of history and science was a great inspiration to me at a young age; my parents Elisabeth and Wolfgang Schönholzer, who have supported my education unconditionally; and my partner Fanling Chen, who has always been there for me when obstacles seemed insurmountable and helped me overcome them.

Contents

List of Figures

List of Tables

Acknowledgments

I thank the chair of my committee, Fred Finan, for his countless hours of support and encouragement. Ernesto Dal Bó, Pat Kline, Ted Miguel and Danny Yagan have also gone far beyond the call of duty in their guidance of my work. I have also greatly benefited from conversations with Alan Auerbach, David Card, Victor Couture, Brad DeLong, Jan DeVries, Bruce Fuller, Martha Olney, Christopher Palmer, Gerard Roland, Jean-Laurent Rosenthal, Jesse Rothstein, Kurt Schmidheiny, Raúl Sánchez de la Sierra, Jeff Vincent, Joachim Voth, Nancy Wallace, Eric Weese, Noam Yuchtman, and Gabriel Zucman. I have also learned a great deal from my classmates, co-authors and friends at Berkeley, especially Erin Kelley, Julien Lafortune, Gregory Lane, Kirsten Vacca, and Calvin Zhang.

A number of people and institutions have also made significant contributions to my research progression: I thank Paulo Issler and Nancy Wallace at the Fisher Center for Real Estate and Urban Economics at Berkeley for generously allowing me access to Center resources. I gratefully acknowledge financial support from the Swiss National Science Foundation, Fellowship numbers P1SKP1-158639 and P1SKP1-165020 for my work on early states, and from the Lincoln Institute of Land Policy for my work on municipal annexations. I also thank Ike Allibhai, Stephen English, Günther Fischer, Jeremy Gordon, James Langat, Boris Maguire, James Mwangi, Frank Reed, Tom Rubin, and Nico Voigtländer for their important contributions in moving my research forward by giving advice, dedicating their time and effort, or making data available for research.

Chapter 1

Valuing Local Public Goods using Municipal Annexations

1.1 Introduction

The provision of local public goods such as schooling, public safety, roads, utilities, and public transit form the bedrock of a well-functioning state. In the United States, more than 90,000 independent local governments spend around a trillion dollars each year on the provision of these goods, making up about 10% of U.S. GDP and employing almost 14 million people (U.S. Census Bureau Government Division, 2013). Despite the importance of local public goods, there is a long-standing and unresolved debate about household valuation of local public goods going back to Samuelson (1954) and Tiebout (1956). This is especially true for public goods other than schooling: while the provision of school resources has been studied extensively (see e.g. Card and Krueger (1996)), much less is known about the value of other local public goods.[1]

There are at least two reasons why so little is known about the value of these local public goods. First, it is difficult to find settings in which households experience significant changes in local public goods. Compare this to schooling, where individual programs can lead to changes worth hundreds or even thousands of dollars per recipient over a short time period (e.g. Cellini et al. (2010); Jackson et al. (2016); Lafortune et al. (2017)). Second, local public goods are provided by a complex structure of overlapping local governments (i.e. counties and municipalities) – again in contrast to schooling, which is almost exclusively provided by a single layer of local

[1] Henceforth, unless stated explicitly, we refer to local public goods other than schooling simply as "local public goods".

government (i.e. school districts).

This paper overcomes these difficulties by exploiting a unique and widespread phenomenon in U.S. local governance: changes in local government boundaries. Concretely, because many locations in the U.S. lie outside of the boundaries of any incorporated town, more than a third of Americans (121 million people) receive local public goods primarily from counties. When municipalities annex unincorporated county areas, this leads to an abrupt change in service responsibility for local public goods. As municipalities typically provide substantially higher levels of local public goods than counties, annexed areas experience a substantial and permanent increase in public goods, with no concomitant change in schooling. Thus, changes in population and house prices caused by annexation provide information on how households value local public goods.

Using the variation induced by municipal annexations, we establish three key findings about the value of local public goods: first, we find that one dollar of public expenditures on local public goods is valued at between \$1.04 and \$1.32, suggesting substantial valuation of local public goods, with the weight of evidence leaning towards underprovision of public goods. This finding stands in stark contrast to models of excess local government spending such as Brennan and Buchanan (1980) and recent empirical empirical evidence in support of it (see e.g. R. Diamond (2017)). Second, we find that while expanding public goods through municipal annexations leads to substantial benefits, the incidence of increased public goods falls almost entirely on landowners and property owners, with few benefits accruing to renting households or local government budgets. And third, we benchmark our estimates for public goods to the value of schooling by comparing the size of house price discontinuities across municipal boundaries to those across school districts. We find that differences in the quality of local public goods generate substantially larger price discontinuities than differences in the quality of schooling.

We begin by developing a model that yields a simple expression for how changes in housing supply, house prices and public goods after boundary changes inform the value of local public goods. The expression shows that both increases in house prices *and* increases in housing quantities imply higher valuation, generalizing the standard valuation approach developed by Brueckner (1979) and Barrow and Rouse (2004) that assume fixed quantities. The intuition for this result is that, because agents in our model are mobile and have idiosyncratic preferences for residential choices, utility does not equalize across space, but rather households derive idiosyncratic economic rents from their optimal choices. Thus, if households choose to (re)locate into an area after annexation to take advantage of an increase in public goods, our generalized expression takes into account that households forego these economic rents, driving up the implied value of public goods above and beyond an increase in the willingness to pay for housing in the area. Moreover, due to the presence of preference heterogeneity, we can use our valuation estimates to infer the incidence of increased public goods through annexations on households. We also derive expressions for the incidence on landowners and on government budgets, allowing us to quantify the distribution of benefits and costs of annexation as well as overall welfare.

To estimate the model, we require estimates for three key parameters: the change in housing supply, the change in house prices, and the change in public goods. We estimate these parameters using an event study design, exploiting the uncertainty in the timing of annexation. Specifically, we compare real estate transactions and housing construction in small areas both before and after annexation, compared to trends in other small areas both inside and outside of the annexing municipality as well as areas that are annexed at a different time.

To this end, we combine fine-grained administrative boundary change data from 189 Californian cities showing the complete history of boundaries since incorporation with the universe of residential real estate transactions for 1988-2013. In California, more than 1,300 square miles were annexed since 1991, making it the state with the fifth largest area to be annexed in the last 25 years (U.S. Census Bureau, 2015). The frequency of annexations, the existence of fine-grained administrative boundary data, the availability of high-quality real estate sales data, and the fact that the California constitution essentially fixes property tax rates irrespective of local government make California an ideal setting to study the value of local public goods through annexation.

Our difference-in-difference estimates of the key parameters suggest that annexations lead to (a) large housing supply growth, (b) moderate price increases, and (c) substantial increases in public goods, relative to trends in similar areas. Specifically, the housing stock increases by about 38% five years after annexation; house prices increase annually by around $1,471; and public goods expenditures per household increase by about $1,347. As a robustness check, we show that changes in housing supply and house prices are not driven by changes in the socio-economic composition of neighborhoods: using the surnames of buyers and sellers of each transaction, we see no change in the share of buyers or sellers with white surnames. We also see no change in the loan-to-value ratio of buyers, suggesting that annexation does not attract systematically wealthier households.

Using these reduced form estimates and calibrating the extent of preference heterogeneity, we derive estimates for the marginal valuation of local public goods concomitant with municipal annexations. In our baseline estimate, we assume households have homogeneous preferences for residential location, arriving at an estimate of $1.04 for one dollar of public expenditures. For moderate values of preference heterogeneity, valuation increases to $1.32. The latter estimate suggests underprovision of local public goods, and hence the presence of frictions in adjusting boundaries efficiently. Thus, our evidence implies that there are too many unincorporated areas, offering empirical support for recent programs seeking to simplify annexation procedures (Caballero, 2009).

Our incidence estimates provide one explanation for why there may be a shortage of local public goods provided through municipal annexations: under all plausible parameter configurations, landowners capture 90% or more of the incidence of annexations, with gains between \$385-\$500 million for 502 annexations in California over 1988-2013. Landowners benefit much more than other stakeholders because they derive profits from the value of the total housing stock, benefiting both from price and quantity increases. In contrast, in our most pessimistic scenario, the incidence on households is zero because all benefits from more public goods get eroded by corresponding price increases; even in the optimistic scenario, aggregate households benefits reach only \$20 million. In the pessimistic scenario, local government budgets suffer a loss of \$41 million due to the costs of providing additional public goods while taxes rise only moderately. Only in our most optimistic scenario do local governments benefit from municipal annexation, gaining net revenues of \$12 million. Given the negligible benefits accruing to households and the potential costs to government budgets, some annexations may be difficult to implement in the absence of efficient transfer schemes. Adding up the incidence across stakeholders, the total welfare impact of the observed annexations is between \$344-\$551 million.

Finally, we contextualize our estimates of the value of local public goods relative to schooling. Since boundary changes of school districts are rare, we use static differences across school district boundaries and municipal boundaries, compared to differences across municipal-county boundaries. Building on the approach pioneered by Black (1999), we use boundary fixed effects to account for differences in neighborhood quality and isolate the difference in jurisdictional quality. We use average student achievement in school districts and crime rates in municipalities as quality proxies for the public goods provided by adjacent jurisdictions. To deal with the fact that local government boundaries may separate multiple layers of local government, we identify jurisdictional boundaries that uniquely identify one type of local government across eleven U.S. states. We then form pairs of adjacent jurisdictions and estimate spatial price discontinuities when crossing from the lower-quality jurisdiction to the higher-quality jurisdiction. We find that price discontinuities across municipal boundaries are substantially larger than those between school districts, even after controlling for a large array of housing characteristics and Census block group characteristics. We also show that municipal price discontinuities are about the same size as municipal-county discontinuities, which in turn are similar to the long-run annexation effect on house prices.

This paper contributes to a number of related literatures in local public finance and political economy. We contribute to the literature assessing the valuation and efficiency of local public goods originating in Samuelson (1954). Using a new approach, our study goes against models assigning little value to local public goods expenditures (Brennan and Buchanan, 1980). Accordingly, our estimates are also at odds with much of the recent empirical evidence pointing towards little value of local public goods other than schooling (Bradbury et al., 2001; Boustan, 2013; R. Diamond, 2017) or excessive regulation by municipal governments (Turner et al., 2014). In contrast, our estimates suggest local public goods are highly valued by households and likely underprovided by local governments. These findings are in line with other evidence using a dynamic estimation framework (Bayer, McMillan, et al., 2016). Moreover, unlike the controversial contingent valuation approach (Carson, 2012; Hausman, 2012) relying on stated preferences, our approach allows for an estimation of public goods valuation using revealed preferences.

Our work also adds to the literature on the distribution of local public goods initiated by Tiebout (1956). The empirical assessment of local public goods in equilibrium models has typically used only cross-sectional data from a small number of jurisdictions (Epple and Sieg, 1999; Epple et al., 2001; Calabrese et al., 2006). Our reduced form results using rich micro data provide a useful complement to these structural approaches, and our theoretical contribution allows for a straightforward and intuitive interpretation of how changes in public goods affect the distribution of equilibrium quantities and prices.

We are the first to study unincorporated areas and municipal annexations using fine-grained real estate data, despite its importance for service quality and U.S. local public finance more generally. Earlier work studying annexations typically worked with state-level data (Facer, 2006) or data from Census tracts (Austin, 1999), which rarely coincide with the changes in municipal boundaries induced by annexation. We also overturn the finding that boundary changes are insufficiently flexible to affect house prices (Epple and Romer, 1989).

The rest of this paper is organized as follows: in section 1.2, we describe the structure of U.S. local governments and municipal annexations; section 1.3 describes the data; section 1.4 develops the model we use to derive our key expression for the value of local public goods; in section 1.5, we present our research design to estimate the core parameters of the model; section 1.6 presents estimation results and robustness; section 1.7 combines estimates to inform valuation, incidence and welfare; we contextualize results using boundary discontinuities in section 1.8; and finally, section 1.9 concludes.

1.2 Context: Local Public Goods Provision in the U.S.

Local governance in the U.S. is typically divided into two classes: general purpose governments and special purpose governments (special districts and school districts). Local public goods are mainly provided by counties and municipalities, while school districts provide mainly education services.[2]

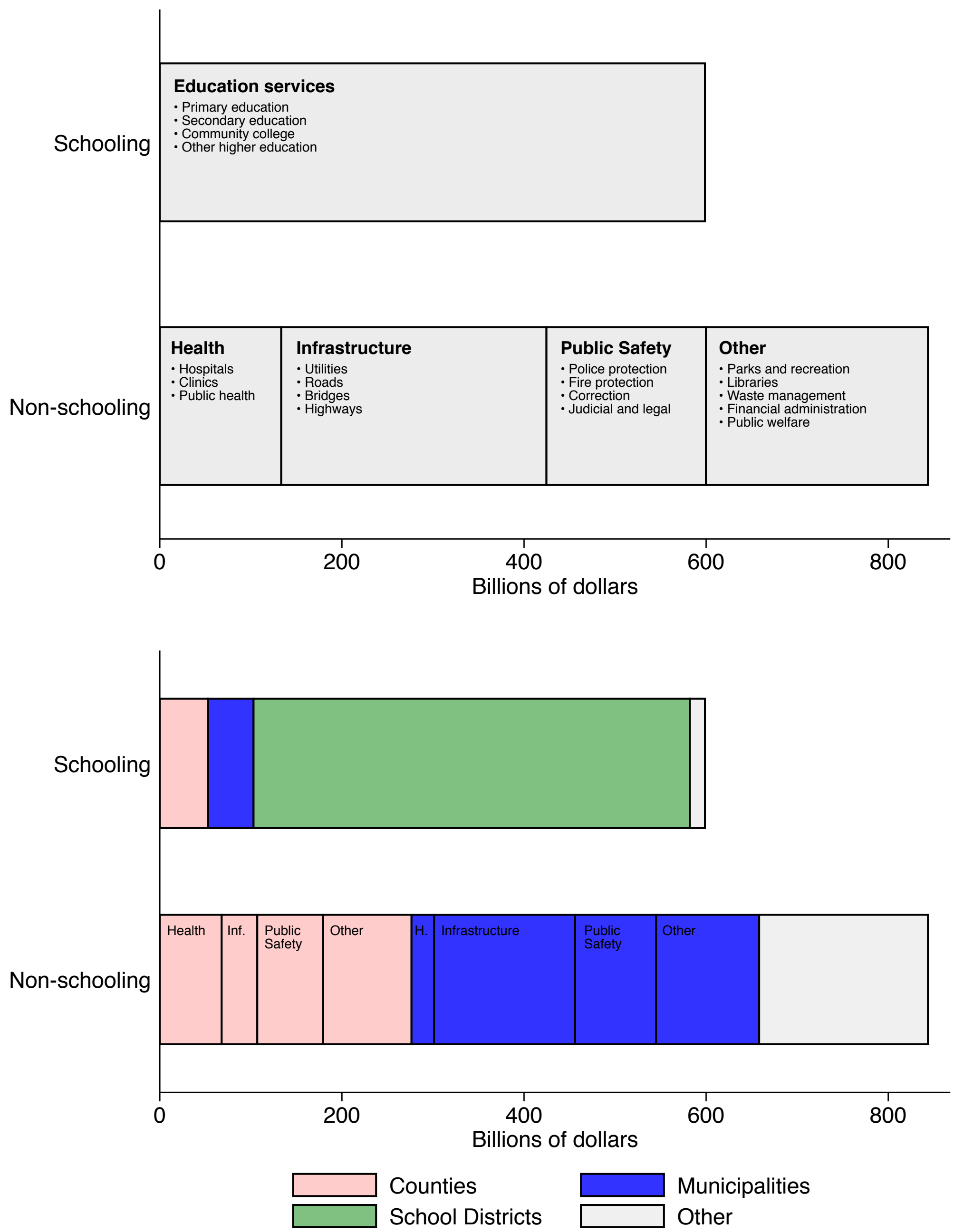
General purpose governments provide a wide variety of services (see figure 1.1). Counties provide three types of services: state-mandated services, such as public welfare and health services; county-wide services such as courts, property assessment, election administration, and correction (jails); and finally, they provide local services in places without municipal government or where municipalities outsource services to them. In contrast, municipalities provide only local services. Their largest expenditure item and their most important responsibility is public safety, including police and fire protection. On average, municipalities spend about $300 per capita on police protection and $100 per capita on fire protection. Local services include a number of other items too, such as street maintenance, utilities, parks and recreation, sewerage maintenance, and solid waste management.

County governments play an important role in local service provision. To see this, it is useful to examine local governance across "places". The U.S. Census defines a place as a concentration of people, irrespective of its local government structure. According to the American Community Survey (U.S. Census Bureau, 2016a), 9,691 places (32%) in the U.S. are unincorporated, with a total population of more than 36 million (16% of total place population). County governments and special districts are in charge of local service provision in unincorporated places. They also provide local services to the 85 million people living in rural areas outside of population concentrations. Service levels and code enforcement are typically lower in counties than in municipalities, especially in unincorporated neighborhoods interspersed between collections of municipal governments in metropolitan areas (Anderson, 2008).

In contrast to the limited powers of county governments, municipalities enjoy "home rule". Municipal home rule grants local governments substantial autonomy from state governments, allowing them to regulate matters of local interest without interference from higher levels of government. These municipal powers have a long tradition of support in court on the basis of federal and state constitutions alluding to "an inherent right of local self-government" (McBain, 1916).

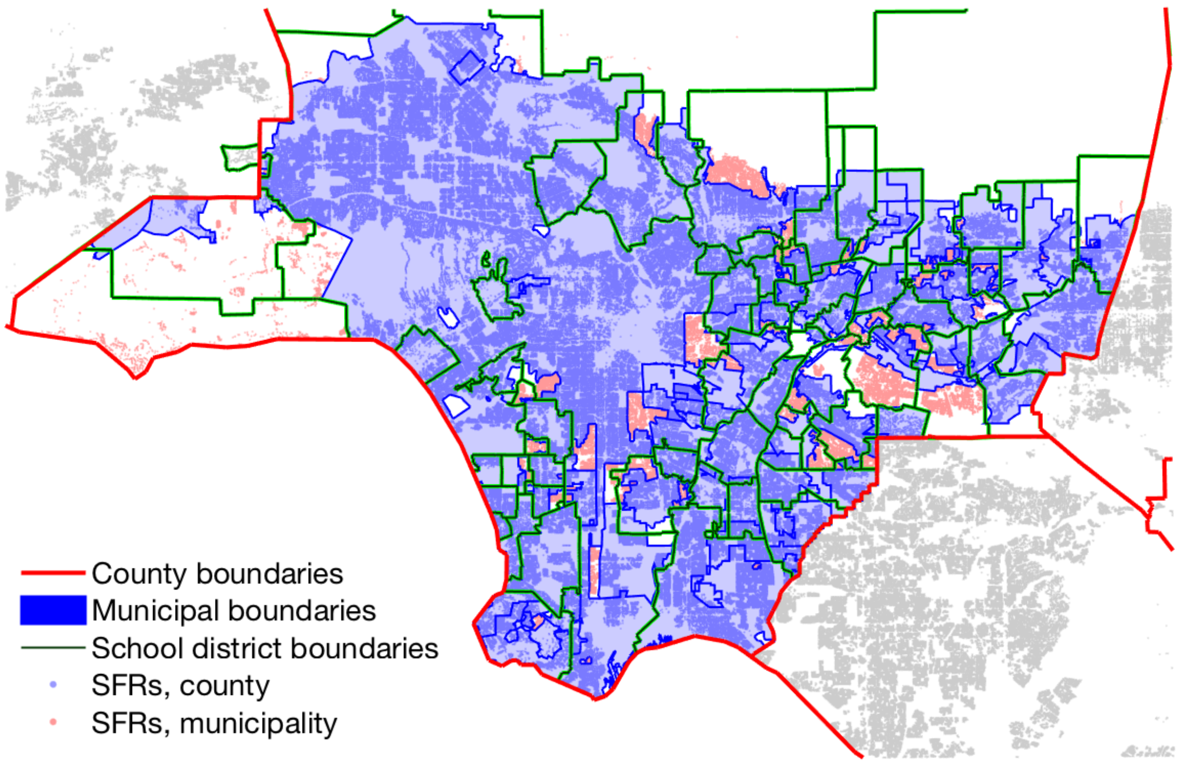
[2]The Census of Governments also defines a second class of sub-county governments – towns or townships – which "provide services to an area without regard necessarily to population" (U.S. Census Bureau Government Division, 2013:p. viii). Since this type of local government does not exist in our setting, there is no need to adjudicate whether they are more similar to counties or municipalities, our main comparison.

Figure 1.1: Schooling and local public goods expenditures.



Note: Types and quantities of local public goods expenditures, divided into schooling and local public goods. In the top panel, we show expenditures by expenditure category. In the bottom panel, the same categories are organized by the type of local government providing them. "Other" includes townships and special districts.

Figure 1.2: Territorial division of local services into county and municipality.



Note: Los Angeles metro area and local public goods provision by service provider. Each dot represents a single-family residence (SFR) in our database. Those colored in red receive local public goods directly from Los Angeles County; those in blue receive services from one of the 88 municipalities in Los Angeles county. School district boundaries are delineated in green.

Even in urban areas, counties play an important role in local service provision. For example, figure 1.2 shows all areas in the Los Angeles metropolitan area receiving services by municipalities (in blue) or from the county directly (in red). In total, more than one million people in Los Angeles county live in unincorporated territories, depending on the county for service provision (see table 4.1).

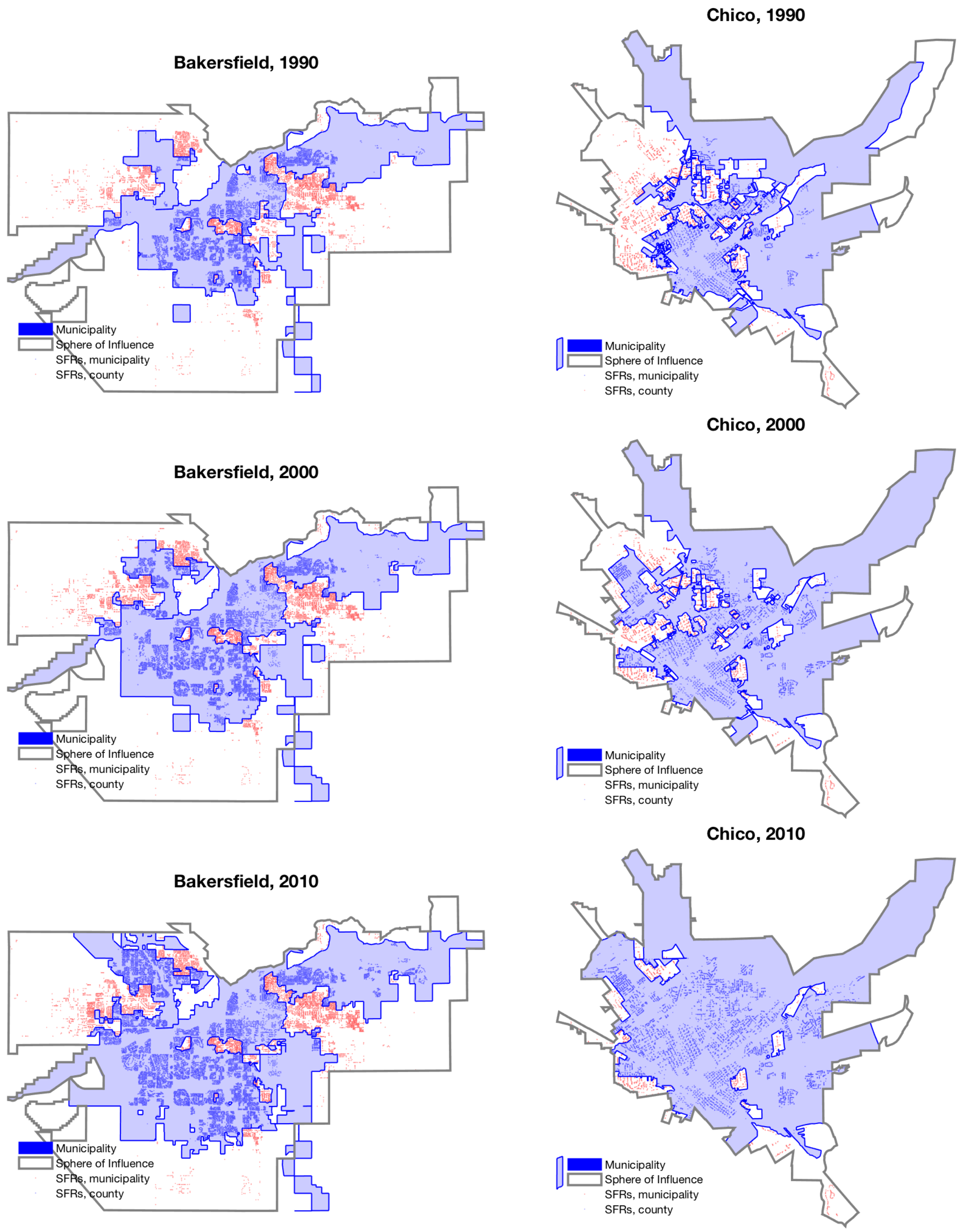
Municipal boundaries change frequently through annexation of unincorporated county territory. The Boundary and Annexation Survey (U.S. Census Bureau, 2015) of the U.S. Census recorded almost 200,000 annexation events covering almost 25,000 square miles since 1990. This process is regulated by state law and typically involves agreement from the municipality, the land or property owner, or both; elections may be required in both the annexing and the annexed territory; and commissions may also adjudicate the process according to particular goals set by the state government (Facer, 2006; Edwards, 2008).

Figure 1.3 shows two examples of municipal annexations. Looking at areas annexed by 1990, it can be seen that local governance is organized in a haphazard way, with the municipality and the county dividing service provision responsibility in a complex assignment of neighborhoods to jurisdictions. This pattern of incomplete municipal governance can be found across U.S. metropolitan areas. In California, it roots in the rapid expansion of municipal boundaries during the boom years after the Second World War, when municipalities often leapfrogged across areas to reach neighborhoods generating high tax revenue. The state has since then put laws into place to combat this pattern of local governance. However, these laws had only partial success at integrating urban areas under a single jurisdiction: even by 2010, there are still a number of residential areas that are under county governance.

It is important to note that the municipal boundary change process is largely independent from school district boundary changes. School districts typically cover the entire sphere of a municipality, encompassing both incorporated and unincorporated areas (see figure 1.2). Even if they cut through municipal spheres, setting school district boundaries is a completely separate process, changing much more rarely than municipal boundaries.

California exhibits a number of peculiarities that make it a particularly suitable setting to study questions of municipal governance versus county governance. First, and most importantly, due to proposition 13, property taxes are essentially fixed across all locations at 1%. Excess property taxes are mostly due to school bonds, which apply across municipal boundaries. Second, the state mandates that every municipality has a "sphere of influence" in which it has the exclusive right to annex territory. This mandate was put in place to avoid what was perceived as harmful competition over territory (Caballero, 2009).

Figure 1.3: Examples of municipal annexations.



Note: Examples of municipal annexation in California. The left column of panels shows Bakersfield in 1990, 2000, and 2010; the right column shows Chico in the same years. The sphere of influence of each municipality is shown in gray. Blue areas are governed by the municipality. Non-blue areas are governed by the county. Single-family properties colored according to service provider.

1.3 Data

We combine three data sources to estimate the relationship between municipal governance, house prices and public goods. First, we collect a novel dataset on the universe of boundary changes for 189 municipalities in California. Second, we combine these boundary changes with the universe of property sales in the state for 1988-2013. Finally, we include detailed public finance data from counties and cities in California. We describe each of these in turn.

To measure jurisdictional boundaries, we use administrative municipal boundary change data from individual counties and cities in California. These data precisely document the evolution of municipal boundaries, often all the way back to the original municipal incorporation. In addition to the boundaries, these data capture the year a particular area (neighborhood) was annexed to a given municipality. Unlike alternative data sources such as the TIGER/Line place database by the U.S. Census, these data are collected for administrative purposes, often in the context of property assessment. In contrast, TIGER/Line is based on an annual voluntary survey, and so the timing of municipal boundary changes is of poor quality, especially before 2007 (U.S. Census Bureau, 2016b).

We were able to obtain the administrative boundary change data from 189 municipalities (out of 482) across 18 counties (out of 58) in California. Not all counties and cities have their entire boundary history in electronic format which is readily available for our purposes. There is no obvious pattern of selection into having this type of data: both small and large places, relatively wealthy and poor places, and places with few or many boundary changes according to the Boundary and Annexation event counts appear in the data. The sample is somewhat biased towards Southern California and towards more urban places (most rural counties with very few municipalities are missing). We include a complete list of all counties in the appendix.

Based on the administrative boundary data, we can then partition each city into its constituent areas according to the year they joined (or originally formed) the municipality. An area can be as small as a few properties (as small as $40m^2$) or as large to encompass several neighborhoods. Unincorporated areas that have not (yet) joined the municipality but are within its sphere of influence are also included. An example of the area data structure for a single sphere can be seen in figure 4.2.

We supplement these boundary data with data on the universe of home transactions obtained by DataQuick from each county's assessor office between 1988 and 2013. The DataQuick data contains information on the characteristics of each home that sold and on each transaction for that home. Home characteristics include the home address, lot size, number of bedrooms, number of bathrooms, square footage, number of stories, and the year the home was built, which do not change from transaction to transaction.[3] However, we also observe time-varying transaction values and dates.

While the DataQuick data has the complete address of each home, it does not provide information on if a home is located in a municipality or an unincorporated area. However, having the address allows the home to be geocoded and merged with the boundary data so we can observe if a home transacted in a municipality or not. We keep only transactions that fall within the spheres of influence of the 189 municipalities in our sample. Our merged DataQuick sample contains 4,119,959 transactions on 2,190,313 homes.

To understand the expenditure patterns associated with counties and municipalities, we use the county and city finance data from the California State Controller's Office. It shows complete finances of counties, cities and special districts for 1991-2014. These data allow us to construct local service expenditure measures of counties and municipalities that are comparable to one another. Specifically, we compare the municipality and the county (with a focus on activities targeted at unincorporated areas) along the following dimensions: total per capita expenditures and police protection per capita expenditures. We also compare county and municipal police protection performance by computing adjusted clearance rates using the California Department of Justice's Criminal Justice Statistics Center (CJSC) database (more details below).

Table 3.1 shows summary statistics for spheres, areas and properties in our dataset. In Panel A we show sphere characteristics, focusing on the extent to which the municipality in the sphere has taken over local governance. We see that most spheres had at least some unincorporated areas (83%) in 1988 and in 2013 (71%). In Panel B, we can see characteristics of areas. Unincorporated areas are typically larger, which is why they have on average more homes and larger built territory. We can also see that prices per lot size are typically lower in unincorporated areas. Finally, in Panel C, we look at properties within 500 meters of a boundary dividing the municipality and the county. This is the relevant sample of properties in our boundary discontinuity design.

1.4 Model

In this section, we present a spatial equilibrium model that provides predictions about the impact of annexations on house prices and housing supply in the presence

[3] In reality, these characteristics may change over time through renovations, but in our data, we observe the characteristics from the most recent transaction only.

Table 1.1: Summary statistics for spheres, areas and properties.

Panel A: Sphere characteristics	Mean	P(50)	SD	P(5)	P(95)
Any unincorporated areas, 1988	0.83	1	0.38	0	1
Any unincorporated areas, 2013	0.71	1	0.45	0	1
Share incorporated areas, 1988	0.73	0.84	0.29	0	1
Share incorporated areas, 2013	0.89	0.95	0.15	0.50	1
Number of areas	55.5	14	140.6	1	229
Square miles of total area	28.8	10.8	53.9	1.62	117.5
Municipal expenditure p.c., 1000s	1.87	0.91	9.42	0.35	3.41
County expenditure p.c., 1000s	0.83	1.06	0.34	0.28	1.13
Municipal adjusted clearance rates	1.15	1.12	0.33	0.72	1.72
County adjusted clearance rates	1.03	0.99	0.29	0.72	1.78
N	189				
Panel B: Area characteristics	Mean	P(50)	SD	P(5)	P(95)
Number of homes, annexed areas	231.2	39.5	2467.8	1	737.1
Number of homes, uninc. areas	267.3	11.8	1184.8	1	1305.8
Share built acres, annexed areas	0.48	0.40	0.55	0.0013	1.11
Share built acres, uninc. areas	0.60	0.11	3.34	0	1.23
Avg. sales price, annexed areas	348.2	290.5	228.4	105.7	806.8
Avg. sales price, uninc. areas	314.6	255.3	212.7	100.1	736.5
Puplic expenditures p.c., annexed areas	1209.4	1060.3	2635.5	590.3	1984.2
Public expenditures p.c., uninc. areas	512.5	404.4	334.0	151.9	1098.8
N	10,359				
Panel C: Property characteristics	Mean	P(50)	SD	P(5)	P(95)
Sales price (thousands)	337.3	277.0	223.8	100.6	774.6
Building square feet	1659.7	1522	598.2	936	2879
Lot size square feet	7934.5	6942	4552.3	4356	15246
Bathrooms	2.10	2	0.73	1	3
Bedrooms	3.23	3	0.79	2	5
Stories	1.21	1	0.41	1	2
Age	34.3	35	22.1	2	73
N	2,821,291				

Note: P(x) denotes the xth percentile. All spheres always include areas with municipal governance; most spheres also have areas with county governance (unincorporated areas).

of heterogeneous preferences for areas. The model is based on the classic equilibrium models developed by Rosen (1979) and Roback (1982), extended with the discrete choice framework as in Bayer et al. (2007). It also borrows elements from Busso et al. (2013) and Kline and Moretti (2014) to allow for households to capture economic rents due to their idiosyncratic preferences, which allows the model to quantify household incidence.

The model is populated by two types of agents operating in two markets: households and absentee landlords who interact in the residential housing market. Landlords also pay fixed property taxes to the municipality or the county, which in turn provide public goods to households. The key parameter of interest is the value of public goods, which we assume to be constant across households. Having described the equilibrium impact on prices and supply, we can derive the incidence of annexations on households, property owners and local governments, which provides a basis to assess the welfare effect of annexations.

Importantly, the goal in this model is *not* to offer a full characterization of the distribution of households, prices, tax rates, and public goods in a Tiebout equilibrium. While this type of fully specified model is intellectually appealing and captures dynamics potentially relevant to the subject matter, we aim for the more modest goal of analyzing the impact of changing jurisdictional boundaries on households, prices and public goods under a given set of taxes. Tax rates are severely constrained in the context of tax limitation laws such as Proposition 13 in California. We also focus our attention on household heterogeneity across areas, not with respect to public goods. We do this because our primary interest is to estimate *average* valuation, allowing for households to capture economic rents from idiosyncratic preferences of some kind. Allowing for sorting according to heterogeneous preferences for public goods seems to be substantially harder.

Environment

Consider a city divided into $j = 1, ..., J$ *areas*, each of which is assumed to be small relative to the city. A municipality and a county share service provision across areas in the following way. Let $S : \{1, ..., J\} \to \{0, 1\}$ define the assignment of areas to service providers, where $S(j) = 0$ means the county provides services to area j, and $S(j) = 1$ means the municpality provides services to j. We define a *service arrangement* as a vector $\mathbf{s} = (S(1), ..., S(J))$, which describes the full mapping of areas to service providers. For example, if $\mathbf{s} = [1, 0, 1]$, the first and the third area in the city receive services from the municipality, and the second area from the county.

A unit interval of households choose one of the J residential locations in the city. Households inelastically demand one unit of housing in their area of choice at the equilibrium market price. In making their residential choice, households trade-off public goods by the service provider, the cost of housing, fixed amenities, and idiosyncratic preferences for areas. Thus, their indirect utility function is given by

$$u_{ij}(\mathbf{s}) = \lambda g_j(\mathbf{s}) - p_j(\mathbf{s}) + A_j + \varepsilon_{ij} \equiv v_j(\mathbf{s}) + \varepsilon_{ij} \tag{1.1}$$

where i indexes households, λ is the value of one unit of the public good $g_j(\mathbf{s})$ provided to area j under arrangement $\mathbf{s}$, and $p_j(\mathbf{s})$ is the market price of a unit of housing in j. A_j are amenities other than those provided by the local government, such as distance to the city center or the quality of restaurants in the neighborhood. Finally, the choice-specific error term ε_{ij} describes heterogeneity of households across areas, capturing aspects such as distance to the household's workplace or the area's idiosyncratic appeal to the household.

Taking all of these elements together, we can summarize household utility from choosing residential location j under arrangement $\mathbf{s}$ in two terms: a systematic component $v_j(\mathbf{s})$ describing mean utility from living in j under arrangement $\mathbf{s}$, and the idiosyncratic component ε_{ij}. We assume that ε_{ij} is independently and identically distributed across households according to a type I extreme value distribution with scale parameter σ and mean zero. The scale parameter σ captures to what extent idiosyncratic considerations matter for a household's choice: if σ is zero, residential choice is completely determined by the systematic component of utility; on the other hand, if σ is large, the idiosyncratic component is important to residential choice, changes in public goods, prices and amenities need to be large for households to move in or out of an area. In other words, σ is the inverse elasticity of substitution between different areas.

We now turn to housing supply. In each area j, landlords provide housing competitively at marginal cost. Since land in each area is fixed, we assume marginal cost rises with the share (i.e. number) of housing units provided:

$$(1-\tau)\, p_j(\mathbf{s}) = z_j H_j\,(\mathbf{s})^{\kappa}$$

where τ is the fixed property tax rate, z_j is housing productivity (e.g. depending on the available land in j), and κ is the inverse elasticity of housing supply. Landlord profit is then simply a fixed fraction of total housing cost in an area:

$$\pi_j(\mathbf{s}) = \frac{\kappa}{1+\kappa}\,(1-\tau)\, p_j\,(\mathbf{s})\, H_j\,(\mathbf{s})\,. \tag{1.2}$$

and the overall landlord profits made across the city are given by $\Pi(\mathbf{s}) = \sum_{j=1}^{J} \pi_j(\mathbf{s})$.

Finally, the county and the municipality produce public goods with a constant returns to scale technology subject to the budget given by their property tax revenue:

$$\sum_{j:S(j)=s} \tau_s p_j(\mathbf{s}) N_j(\mathbf{s}) = \sum_{j:S(j)=s} g_j(\mathbf{s}) N_j(\mathbf{s})$$

for each local government $s \in \{0, 1\}$. While landowners are being charged the fixed property tax rate τ, county and municipal government only receive a rate $\tau_s < \tau$ since around half of property tax revenue goes to school districts. Thus, the overall net local government budget can be expressed as

$$R(\mathbf{s}) = \sum_{j=1}^{J} N_j(\mathbf{s}) \left[\tau_s p_j(\mathbf{s}) - g_j(\mathbf{s})\right]. \tag{1.3}$$

Equilibrium

On the basis of this household utility specification, we can now derive housing demand in each area. When choosing their residential location, households take the service arrangement and the corresponding public goods and prices as given to maximize their utility given by (1.1). According to standard results from discrete choice analysis (McFadden, 1973), the share (i.e. number) of households whose optimal residential choice is j is given by

$$N_j(\mathbf{s}) = \frac{\exp\left(v_j(\mathbf{s})/\sigma\right)}{\sum_{k=1}^{J} \exp\left(v_k(\mathbf{s})/\sigma\right)}.$$

Housing prices $p_j(\mathbf{s})$ are pinned down by the housing market clearing condition in each area: $N_j(\mathbf{s}) = H_j(\mathbf{s})$. Since demand for housing in j monotonically falls with price and supply monotonically rises, the system of equations defined by the optimal residential choice and the optimal provision of housing yields a unique solution for the distribution of households and prices across areas. Public goods are then determined by the combination of equilibrium prices and fixed property tax rates.

We can check immediately that the model predicts a straightforward distribution of households and prices for a given set of exogenous parameters: ceteris paribus, areas with nicer amenities attract more households; so do areas with more public goods, which is reinforced the higher the valuation of public goods; higher prices repel households from j; and stronger idiosyncratic preferences for particular areas (i.e. higher σ) mute the effects of the aforementioned factors.

Valuation under Changing Service Arrangement

So far, we have studied the model for a fixed service assignment $\mathbf{s}$. We now show how equilibrium values change in response to a (small) change in the service arrangement. Specifically, we compare the equilibrium under a baseline arrangement $\mathbf{s}_0 = [S(1), ..., 0, ..., S(J)]$ in which the jth area receives services from the county against the arrangement $\mathbf{s}_1 = [S(1), ..., 1, ..., S(J)]$, in which area j now receives services from the municipality – while all other areas continue to receive services from the same provider as before. In other words, the municipality annexes j, leaving other areas unchanged.

To study the impact of a change in service arrangement, we define the following objects: the effect on house prices $\Delta p = p_j(\mathbf{s}_1) - p_j(\mathbf{s}_0)$; the effect on (the dollar value of) public goods $\Delta g = g_j(\mathbf{s}_1) - g_j(\mathbf{s}_0)$; and the effect on the share of households $\Delta \ln N = \ln N_j(\mathbf{s}_1) - \ln N_j(\mathbf{s}_0)$. Notice that, for simplicity, we are making the strong assumption that these effects are constant across areas. We explore relaxing this further below. We can then express the relationship between these effect in the following way:

$$\begin{aligned} \Delta \ln N &= \frac{v_j(\mathbf{s}_1) - v_j(\mathbf{s}_0)}{\sigma} \\ &= \frac{\lambda \Delta g - \Delta p}{\sigma} \end{aligned} \tag{1.4}$$

where we assumed that $\ln \left[\sum_k \exp\left(v_k(\mathbf{s}_1)/\sigma\right) / \sum_k \exp\left(v_k(\mathbf{s}_0)/\sigma\right)\right] \approx 0$, that is, that the total impact of a small change in service arrangement is negligible. This formula says that the log change in the population due to annexing an area equals, on average, the cost-of-living adjusted value of additional public goods, scaled by the elasticity of substitution between areas. We can solve this for the value of public goods to get

$$\lambda = \frac{\sigma \Delta \ln N + \Delta p}{\Delta g}. \tag{1.5}$$

This equation is the key relationship we use to estimate the value of public goods. It is reminiscent of the type of relationship derived in Brueckner (1979). To see this, note that if $\sigma \to 0$, the formula approaches $\lambda = \Delta p / \Delta g$, which says that the marginal rate of substitution between public goods and private consumption (i.e. λ) equals the ratio of the change in house prices and the change in public goods. The intuition is that the extent to which households are willing to pay more to live in an area with more public goods must reflect the value of these goods.

Equation (1.5) is a generalization of this standard relationship in the following sense: allowing households to have idiosyncratic preferences for areas implies that most households are inframarginal with respect to a change in the attractiveness of living in an area. Thus, if households choose to relocate to an area in response to more public goods being available even when they are inframarginal, it must be the case that the valuation of the public goods is even higher than just reflected by the willingness to pay of marginal households.

Incidence and Welfare

Having described how changes in log population, prices and public goods inform the value of public goods, we now turn to describing incidence and welfare of annexations. Average (i.e. total) household welfare in this model can be expressed as

$$\begin{aligned} V(\mathbf{s}) &= E_{\varepsilon}\left[\max_j u_{ij}(\mathbf{s})\right] \\ &= \sigma \ln \left[\sum_j \exp\left(v_j(\mathbf{s})/\sigma\right)\right] \end{aligned}$$

where the second equation follows from standard results in discrete choice analysis. Thus, the impact of annexation on household welfare is given by

$$\begin{aligned} V(\mathbf{s}_1) - V(\mathbf{s}_0) &= \sigma \ln \left[\frac{\sum_j \exp\left(v_j(\mathbf{s}_1)/\sigma\right)}{\sum_j \exp\left(v_j(\mathbf{s}_0)/\sigma\right)}\right] \\ &= \sigma \ln \left[\sum_j \left(\frac{\exp\left(\left[v_j(\mathbf{s}_1) - v_j(\mathbf{s}_0)\right]/\sigma\right)}{1 + \sum_{k \neq j} \exp\left(v_k(\mathbf{s}_0)/\sigma\right)}\right)\right] \\ &= \sigma \ln \left[\sum_j N_j(\mathbf{s}_0) \exp\left(\frac{\lambda \Delta g - \Delta p}{\sigma}\right)\right] \\ &= \lambda \Delta g - \Delta p = \sigma \Delta \ln N \end{aligned} \tag{1.6}$$

where the last equality follows from (1.4) and the fact that $\sum_j N_j(\mathbf{s}_0) = 1$. That is, the welfare change due to a change in service arrangement is just equal to the population movement induced by the change, scaled by the extent of preference heterogeneity. Moreover, this willingness to relocate exactly equals the net benefits of the boundary change (that is, more public goods net of the cost of living).

Next, we define the incidence on landowners. Using (1.2), we can express the change in profits as

$$\Pi(\mathbf{s}_1) - \Pi(\mathbf{s}_0) = \frac{\kappa}{1+\kappa}(1-\tau)\sum_j \left[p_j(\mathbf{s}_1) N_j(\mathbf{s}_1) - p_j(\mathbf{s}_0) N_j(\mathbf{s}_0)\right]. \tag{1.7}$$

Finally, we turn to the impact on government revenue of a change in service provision:

$$R(\mathbf{s}_1) - R(\mathbf{s}_0) = \sum_{j=1}^{J} \left\{N_j(\mathbf{s}_1)\left[\tau_s p_j(\mathbf{s}_1) - g_j(\mathbf{s}_1)\right] - N_j(\mathbf{s}_0)\left[\tau_s p_j(\mathbf{s}_0) - g_j(\mathbf{s}_0)\right]\right\}. \tag{1.8}$$

It should be noted that although we assumed in the setup of the model that each government (county and municipality) faces its own budget constraint, the expression in (1.8) evaluates the impact of the change in service arrangement on the joint budget of the two governments. This is the relevant object for the study of potential fiscal externalities from annexation.

Taken together, the model describes the relationships between observables and parameters that allows us to quantify the impact of annexation on (a) the value of additional public goods provided, (b) the incidence on households, (c) on landowners, and (d) on government budgets. The core parameter estimates required for this purpose are the effect of annexation on housing supply in a given area $\Delta \ln N$ (or equivalently, on the number of households choosing to reside in an area); the effect on average house prices in an area Δp; and the effect on the dollar value of additional public goods available in an area Δg. We also need to make assumptions about the extent of heterogeneous preferences for areas, the housing supply elasticity, and the share of property tax revenue accruing to municipalities.

1.5 Econometric Design

In this section, we describe the econometric design to estimate the aforementioned core parameters $\Delta \ln N$, Δp and Δg. These parameters can be interpreted as the causal effect of annexation on housing, prices and public goods due to municipal annexation. The econometric challenge in estimating these causal effects is that areas that are annexed may be systematically different than those that are not. To address this challenge, we use an event study design, in which we compare outcomes in annexed areas after annexation to those in areas that continue to be unincorporated, those that are part of a municipality throughout our study period, and to annexed areas that are annexed at another point in time.

Concretely, we estimate regressions of the following form for all 10,358 areas over the period 1988-2013:

$$y_{jt} = \alpha_j + \mu_{c(j),t} + \sum_{k=-15}^{15} \mathbf{1}[a_j + k = t]\theta_k + \varepsilon_{jt} \tag{1.9}$$

where y_{jt} is an area j outcome of interest in year t; α_j are area fixed effects; $\mu_{c(j),t}$ are sphere-by-year fixed effects; a_j is the year of annexation of area n, so that θ_k measures the difference in outcome relative to a reference event year. We use $k = -1$ as the omitted reference year, so we can interpret all effects relative to the year before annexation. ε_{jt} is an error term. We set the lower and upper bound relative to the annexation event to 15 years; the endpoints (-15 and +15 years from annexation) are binned to capture the mean effects beyond the endpoints.

The extent to which pre-trends are absent (i.e. $\theta_k = 0$ for $k < 0$) tells us whether municipal annexation is a precursor or a consequence of changes in the outcome, relative to trends. The absence of pre-trends lends itself to a causal interpretation insofar as systematic changes in outcomes occur once additional public goods are available in an area. We argue that this is because of the intensification of services taking place after municipal annexation and provide suggestive evidence for this line of reasoning further below.

Since we have rich information on individual real estate transactions, we can also run a "pure" event study design to estimate the price effect using each property's time series of sales, restricting to properties that are annexed at some point in our study period. We can do this either in a property-by-year panel as in (1.9), or in a property-by-sales-instance panel, in which a property's individual sales instances are used as the relevant event time. In the case of the property-by-sales-instance regression, we are running:

$$y_{is} = \alpha_i + \mu_{c(i),t(s)} + \sum_{k=-2}^{3} \mathbf{1}[e_i + k = s]\beta_k + \xi_{is} \tag{1.10}$$

where y_{is} is the sth sale of property i in our data and e_i is the first sales instance after annexation. The fixed effect α_i captures observed and unobserved property-specific attributes; $\mu_{c(i),t(s)}$ captures city-by-year specific deviations; and ξ_{is} is an error term. Thus, the β_k coefficients pick up the effect of annexation k sales after annexation, unlike in (1.9), in which the coefficients pick up the effect of annexation k years after annexation.

In addition to the nonparametric evolution of changes in the aftermath of the annexation event as estimated by the event study methods, we also use a generalized difference-in-difference design to express the magnitude in a simple two-parameter specification. To this end, we run the following regression:

$$y_{jt} = \alpha_j + \mu_{c(j),t} + \mathbf{1}[a_j \leq t]\theta_1 + (\mathbf{1}[a_j \leq t] \times [t - a_j])\,\theta_2 + \varepsilon_{jt} \tag{1.11}$$

where variable definitions are as before. The coefficient θ_1 estimates the average difference after annexation, compared to before annexation, controlling for trends and unobserved area characteristics with the fixed effects, and θ_2 captures a linear trend break after annexation. We may use specifications with only θ_1, only θ_2, or both, depending on the extent to which effects accumulate over time.

To estimate the event study and the generalized difference-in-difference, we usually include all three types of areas: those that are always in municipalities, those that are unincorporated throughout our data, and those that are annexed and thus switch from unincorporated county governance to municipal governance. Due to the set of sphere-by-year fixed effects we include in all specifications, we identify our parameters of interest from the switching areas; other areas provide precision to estimate the fixed effects.

1.6 Results

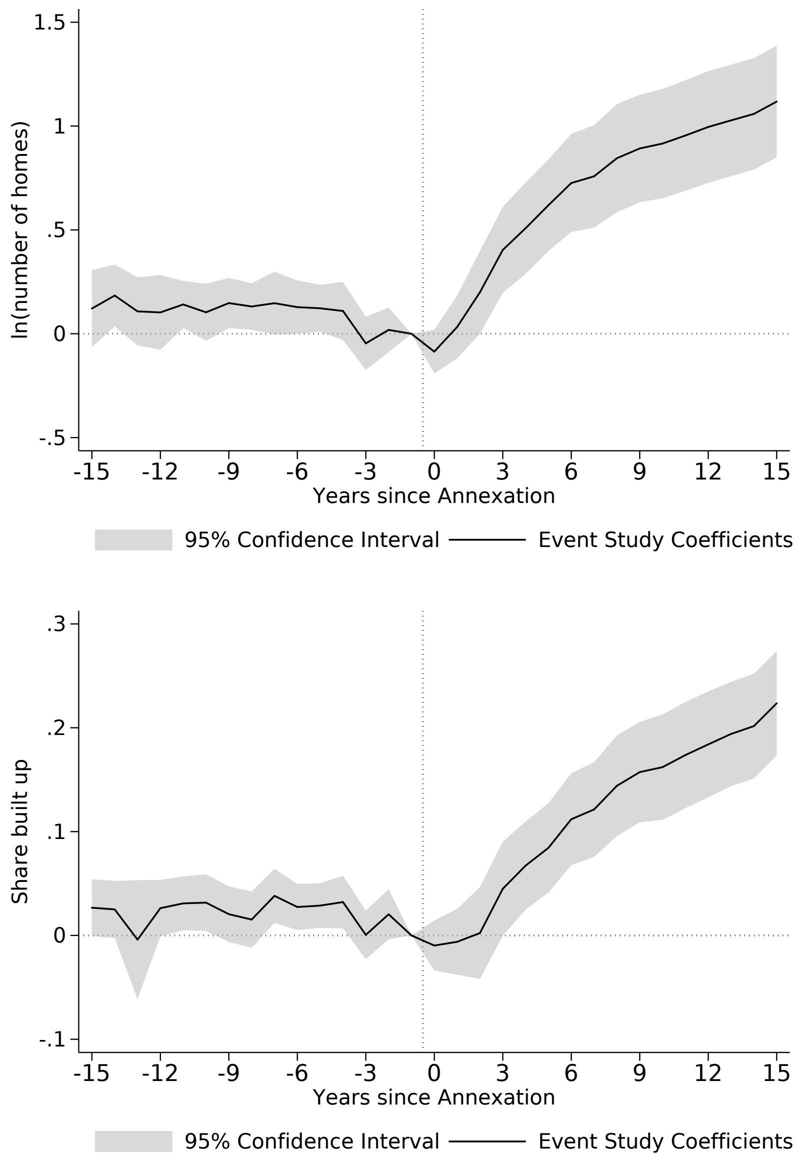
Our theoretical model requires estimates for the change in housing supply $\Delta \ln N$, the change in house prices Δp, and the change in public goods Δg. We structure the presentation of results accordingly. After having pinned down a set of estimates for each of these parameters, we turn to alternative hypotheses for the effects, arguing that the changes in public goods are responsible for the supply and price changes that we document.

Housing Supply Effect

We begin our results on the causal impact of annexation by documenting a strong increase of housing supply starting right after – and not before – annexation. In figure 1.4 we plot the event study coefficients β_k for two outcomes measuring changes in housing supply in a given area: the log of the number of homes in an area, and the total built-up land area. We see that for both measures, the growth in housing supply is largely parallel in municipalities and counties before annexations, but once an area is annexed, its growth in housing supply rapidly accelerates, surpassing supply growth in other areas by more than 30 log points (35%) after three years, relative to before the annexation. Ten years after the annexation, housing supply in an annexed area has grown by more than 150% relative to the year before.

Table 1.2 shows the generalized diff-in-diff parameterization as well as robustness of these estimates to varying comparison groups. In columns (1) and (4), we include all three types of areas – always in municipalities, always unincorporated, and annexed at some point during 1988-2013. This is the same sample we use to estimate the event study coefficients in figure 1.4. We estimate an increase in housing supply of around 40% and in the share built up of about 45% relative to before annexation. Since most of the post-event observations are within five years after the annexation event, at a point when the event study estimates are on a steep trajectory but still on a relatively low level compared to before the annexation, the post-annexation estimate is lower than most of the event study coefficients.

Figure 1.4: Housing supply effect of annexations.



Note: Event study coefficients of log(number of homes) and share built up on annexation. The observations are areas (i.e. neighborhoods defined by municipal boundary changes) for the 189 municipalities for which we observe the complete boundary change history in California. Standard errors clustered on the sphere level.

Table 1.2: Housing supply effect of annexations.

	ln(Houses)			Share built up		
	(1)	(2)	(3)	(4)	(5)	(6)
Annexed	0.383***	0.406***	0.385***	0.045**	0.060***	0.039**
	(0.099)	(0.092)	(0.103)	(0.018)	(0.014)	(0.019)
Fixed Effects:						
Area FE	X	X	X	X	X	X
City-Year FE	X	X	X	X	X	X
Areas included:						
Annexed	X	X	X	X	X	X
Always incorporated	X		X	X		X
Never incorporated	X	X		X	X	
Model Statistics:						
Area-year N	156,336	17,735	146,676	156,879	17,816	147,200
Unique area N	9,552	1,211	8,933	9,640	1,224	9,015
R-squared	0.97	0.95	0.97	0.99	1.00	0.98

Note: Generalized difference-in-difference estimation of the log(number of homes) and share of area that is built up on a post-annexation indicator. Standard errors clustered on sphere level for the 189 municipalities for which we observe the complete boundary change history in California.

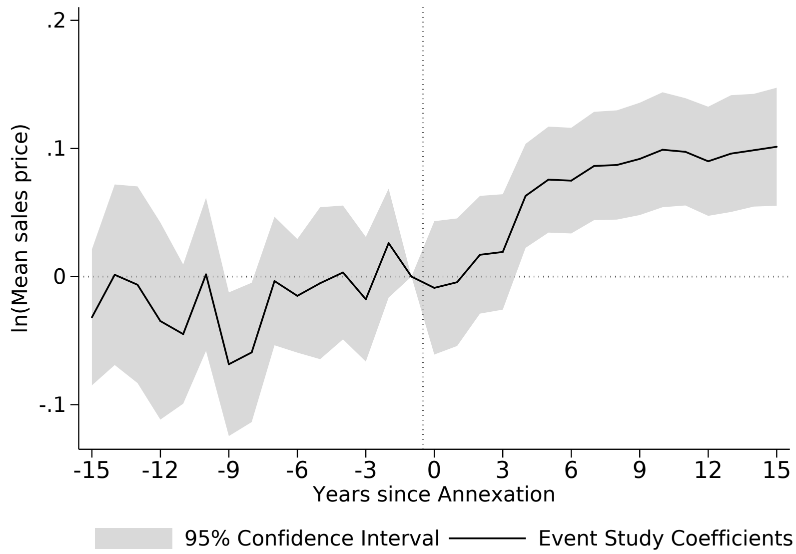
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

In columns (2) and (5), we now only include areas that are unincorporated throughout 1988-2013 and that are annexed during this period. This is the classic diff-in-diff setup with two groups in the same condition (under county governance), and one of the groups getting treated at some point during the study period. We see that coefficients are slightly larger, although we cannot reject that they are the same as in columns (1) and (4). In columns (3) and (6), we restrict to areas that are always incorporated and those that are annexed during 1988-2013. Coefficients are slightly smaller than before but again statistically indistinguishable from those in column (1).

Price Effect

We now turn to the effect of annexation on housing prices. We begin with the same design as with the supply effects, nonparametrically estimating the average house price trend after annexation, as shown in figure 1.5. We see no pre-trend in average house prices for the seven years before annexation, although there are two coefficients (nine and eleven years before annexation) that are significantly negative, suggesting

Figure 1.5: Mean house price effect of annexations.



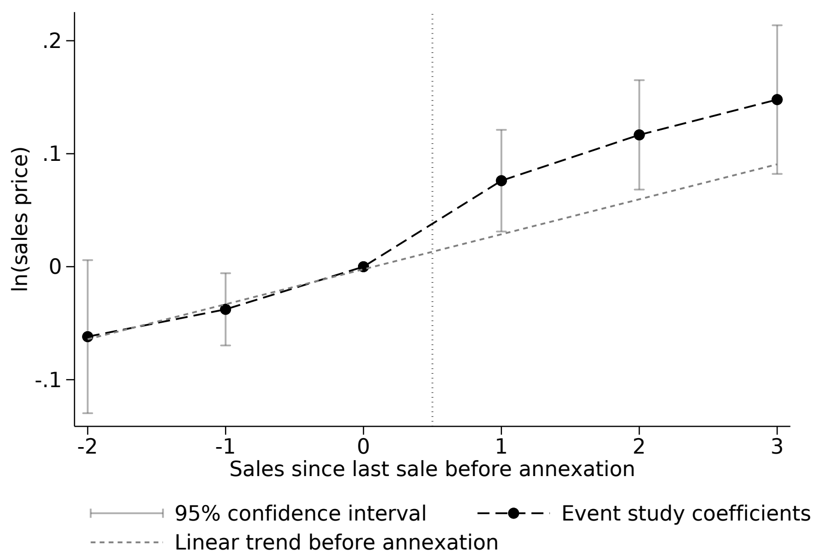
Note: Event study coefficients of average log(sales price) on annexation. The observations are areas (i.e. neighborhoods defined by municipal boundary changes) for the 189 municipalities for which we observe the complete boundary change history in California. Standard errors clustered on city level.

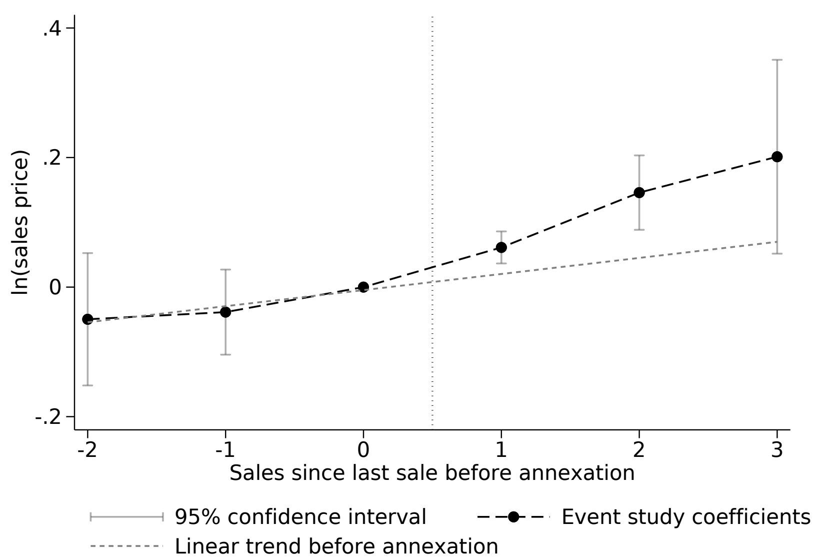
the potential for a slight pre-trend. After annexation, we see average house prices rising slowly until turning significant and reaching around 5% after five years, after which the increase begins to level out at around 7%.

We can again summarize these effects in a simple one-parameter specification, as reported in table 1.3. Columns (1)-(3) show effects on log prices, and columns (4)-(6) on actual sales prices, with specifications again differing in terms of the control groups included (always incorporated and/or never incorporated). With an eye on the desired estimate for the model, we transform each price estimate into a per-annum increase in housing values due to annexation using the average number of years after annexation an area is typically observed. Ranging from $541 to $1,684, the median estimate is $1,111.

While the event study in the area-by-year panel provides evidence for increasing prices, it is useful to corroborate the evidence with the event study in property-by-sales-instance panel, not least because of the issue of potential pre-trends before annexation. Figure 1.6 shows results from this alternative design as well as the counterfactual linear trend based on the sales price before annexation. The top panel

Figure 1.6: Sales-instance event study of house price.





Note: Event study coefficients of log(sales price) on annexation. The observations are individual real estate transactions. Event time is determined through the number of sales since the sale before annexation. In the top panel, all properties that sell at least once before and once after annexation are included. In the bottom panel, all properties that sell at least twice both before and after annexation are included. Standard errors clustered on the sphere level.

Table 1.3: Price effect estimates using average sales prices.

	ln(sales price)			sales price		
	(1)	(2)	(3)	(4)	(5)	(6)
Annexed	0.044***	0.045***	0.029**	9494.2**	10565.1**	4770.7
	(0.013)	(0.013)	(0.011)	(4237.1)	(4371.7)	(4138.3)
Implied annual growth	$1,670	$1,684	$1,111	$1,077	$1,198	$541
Fixed Effects:						
Area FE	X	X	X	X	X	X
Sphere-Year-Acregroup FE	X	X	X	X	X	X
Areas included:						
Annexed	X	X	X	X	X	X
Always incorporated	X	X		X	X	
Never incorporated	X		X	X		X
Model Statistics:						
Area-year N	156,713	147,052	17,789	156,713	147,052	17,789
Unique area N	9,633	9,010	1,222	9,633	9,010	1,222
R-squared	0.94	0.94	0.90	0.93	0.93	0.89

Note: Generalized difference-in-difference estimation of the log(sales price) and sales price on a post-annexation indicator. Standard errors clustered on sphere level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

shows the figure for all annexed properties; the bottom panel restricts to properties for which we observe at least two sales before and after annexation.

It can be seen that, controlling for property fixed effects and nonparametric city-by-year trends, house prices of annexed properties typically rise throughout the study period (they do this despite having converted all sales prices into 2010 real dollars). However, importantly, we see a substantial jump in sales prices in the first sale after annexation, relative to the pre-annexation sales trend. When restricting to properties that sell at least twice both before and after annexation, the price growth after annexation seems to be sustained beyond the first sale after annexation. The latter figure suggests that, in addition to testing for an upward shift in prices after annexation, the trend break specification provides an important test for the price effect of annexation.

We present estimates for trend breaks and shifts using the panel of annexed properties in table 1.4. The first two columns show results for the log sales price, and the last two for the actual sales price; we summarize the key parameter estimate again in the form of the implied annual price growth underneath the coefficient estimates. Results are remarkably similar as in the area-year panel and leaning

Table 1.4: Price effect estimates using individual sales.

	ln(sales price)		sales price	
	(1)	(2)	(3)	(4)
Annexed × Event-time	0.0043***	0.0048***	1642.7**	1654.1*
	(0.0016)	(0.0018)	(717.0)	(836.9)
Annexed		0.011		293.7
		(0.012)		(4595.8)
Implied annual growth	\$1,471	\$1,616	\$1,643	\$1,654
Fixed Effects:				
Property FE	X	X	X	X
Sphere-Year FE	X	X	X	X
Model Statistics:				
Property-year N	122,113	122,113	122,113	122,113
Unique property N	47,934	47,934	47,934	47,934
R-squared	0.96	0.96	0.95	0.95

Note: Generalized difference-in-difference estimation of the ln(sales price) and sales price on a linear trend break and the trend break interacted with a post-annexation indicator. Sample includes properties for which we observe sales before and after annexation. Standard errors clustered on sphere level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

towards the higher end relative to the specification presented earlier: the lowest estimate suggests that annual price growth due to annexation is \$1,471, and the highest estimate is \$1,654.

Changes in Public Goods

The final core parameter we need to estimate is the dollar value change in public goods due to annexation. Ideally, we would be able to measure the public expenditures targeted at each individual area in each moment in time. However, public goods are rarely provided in such a spatially explicit way within a jurisdiction: a jurisdiction typically has, for instance, one police department which responds to all service requests within the jurisdiction. Thus, we assign to each area the per capita expenditure of the corresponding jurisdiction before and after annexation. Concretely, in years before annexation, we assign the expenditures per capita for unincorporated areas of the county corresponding to area j; and in years after annexation, we assign the expenditures per capita of the corresponding municipality. More details on the construction of county expenditures per capita for unincorporated areas are discussed in the appendix. In addition to the impact on expenditures per capita,

Table 1.5: Changes in public goods.

	Expenditures per capita			Adjusted clearance rate		
	(1)	(2)	(3)	(4)	(5)	(6)
Annexed	481.276***	514.490***	487.908***	0.142**	0.106	0.136**
	(68.467)	(81.039)	(66.463)	(0.064)	(0.072)	(0.065)
Fixed Effects:						
Area FE	X	X	X	X	X	X
Sphere-Year FE	X	X	X	X	X	X
Areas included:						
Annexed	X	X	X	X	X	X
Always incorporated	X		X	X		X
Never incorporated	X	X		X	X	
Model Statistics:						
Area-year N	141,320	16,183	132,670	155,977	17,687	146,429
Unique area N	9,465	1,190	8,863	9,603	1,218	8,986
R-squared	1.00	0.94	1.00	0.94	0.83	0.98

Note: Generalized difference-in-difference estimation of public expenditures per capita and adjusted crime clearance rates. Public expenditures per capita before annexation comes from county expenditures items mainly targeted at unincorporated areas divided by the unincorporated population; and after annexation from all municipal expenditures divided by the population in the municipality. The adjusted crime clearance rate captures the number of crimes cleared (i.e. charges being laid against a subject) over the number of reported crimes, corrected for potential differences in the types of crimes reported. See appendix 4.1 for details. Standard errors clustered on sphere level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

we also provide evidence on the impact of the quality of public goods provided. We do this using the crime clearance rate of the associated jurisdiction. In essence, we are comparing the quality of county sheriff departments, which are in charge before annexation, to the quality of municipal police departments, which typically take over service provision after annexation.

Columns (1)-(3) in table 1.5 show the change in public goods expenditures associated with municipal annexation for the three combinations of control groups in the area panel. Estimates range from \$481 to \$514. This means that areas experience an increase in public goods expenditure of about 58%, starting off from a county expenditure per capita average of around \$830.

The results presented in columns (4)-(6) of table 1.5 provide evidence that a change in jurisdiction is associated with a change in public service quality: the adjusted clearance rate increases by about 14% in the aftermath of annexations. In

Table 1.6: Changes in sociodemographic composition.

	white buyer		white seller		share loaned	
	(1)	(2)	(3)	(4)	(5)	(6)
Annexed × Event-time	-0.00079	-0.00049	-0.0018	-0.0015	-0.00033	-0.0010
	(0.0011)	(0.0013)	(0.0066)	(0.0065)	(0.0033)	(0.0034)
Annexed		0.0076		0.0071		-0.018
		(0.014)		(0.014)		(0.011)
Fixed Effects:						
Property FE	X	X	X	X	X	X
Sphere-Year FE	X	X	X	X	X	X
Model Statistics:						
Property-year N	98,055	98,055	61,672	61,672	120,414	120,414
Unique property N	39,702	39,702	25,827	25,827	47,262	47,262
R-squared	0.57	0.57	0.54	0.54	0.51	0.51

Note: Generalized difference-in-difference estimation of socio-economic characteristics of house buyers and sellers. "White seller" is an indicator for whether the seller name is classified as white (average: 74%) with probability greater than 50%; "White buyer" is an indicator classified as white (average: 55%) with probability greater than 50%; loan share is the loan-to-value ratio of the buyer (proxying buyer wealth). Sample includes properties for which we observe sales before and after annexation. Standard errors clustered on sphere level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

other words, controlling for the likelihood of a given type of crime to be cleared, municipal police departments are significantly more likely to do so than county sheriff's offices. This is also consistent with evidence shown in Fujioka (2014), which analyzes staffing levels and emergency response times of the Los Angeles Sheriff's Office and finds that response times are about 20% longer in unincorporated areas than in municipalities.

Robustness

In addition to changes in public goods, it is possible that annexations lead to changes in the socioeconomic composition of an area. We test this possibility in this section, evaluating changes along two dimensions: the ethnic composition of buyers and sellers in an area, and wealth of buyers in an area. To construct the ethnic composition of buyers and sellers, we use the surname of the buyer and seller listed on the transaction record and combine it with the Census database on surnames. In this database, surnames are probabilistically classified according to their reported race and ethnicity. Buyer names that are at least 50% white make up 55% of the buyers of properties that are annexed at some point; reflecting demographic change in California, seller names that are at least 50% white make up 74% percent of annexed properties.

Table 1.6 shows results for whether the probability of buyer and seller ethnicity being white changes systematically after annexation using the panel of property transactions of annexed properties, using trend-break and post-dummy specifications. Concretely, the dependent variable in column (1) and (2) is an indicator for whether the buyer is white; and in (3) and (4) for whether the seller is white. The estimated changes are all below 1%, implying that there are no meaningful changes in the ethnic composition of areas induced by annexation.

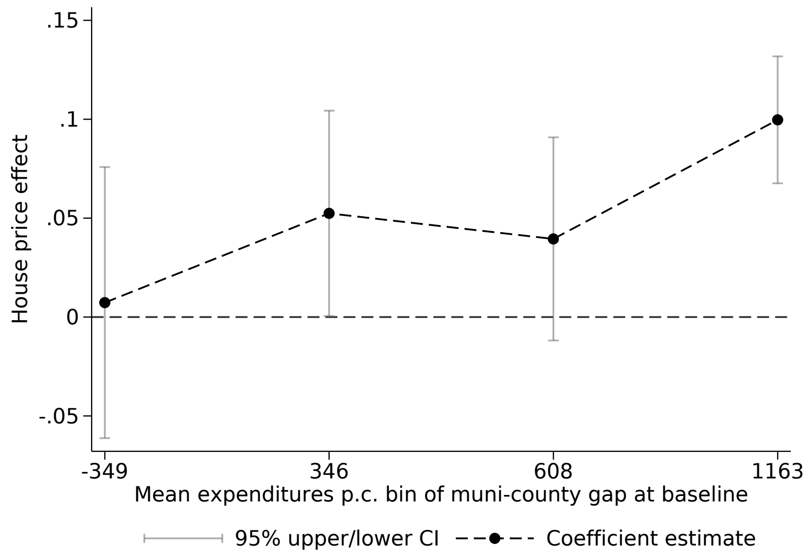
In addition to changes in ethnic composition, we study whether more wealthy households are attracted to an area after annexation. As a proxy for wealth, we use the loans reported on the transaction record to compute the share of the loan-to-value ratio, assuming that the remainder is covered out of their own funds. Using the loan-to-value ratio as the dependent variable, we again find no significant changes as reported in columns (5) and (6) in table 1.6. Together, these estimates render it unlikely that the observed changes are mainly induced by socioeconomic change as opposed to changes in public goods.

Since the bundle of changes induced by annexations includes changes in regulation and zoning beyond changes in public goods, it is useful to estimate the house price effect separately by the magnitude of changes in public expenditures between the county and the municipality. If the price effect is mainly driven by increased public expenditures, a larger difference between county and municipal expenditures should be associated with a larger house price effect. In addition, when the change in expenditures is near zero, there should be little house price effect, unless households value the changes in regulation and zoning directly. We present results testing these hypotheses in figure 1.7. First, we can see that the house price effect is generally increasing as we move up the quartiles, with the exception of moving from the second to the third quartile. This is consistent with a valuation of the public expenditures, as opposed to other changes. Additionally, the coefficient in the lowest quartile (for which the change in public expenditures is slightly negative, i.e. the municipality spends a bit less per capita than the county) is very close to zero, suggesting that the valuation of changes in zoning and regulation are near zero as well.

1.7 Valuation, Incidence and Welfare

Based on the parameter estimates described in section 1.6, we can now evaluate the model expressions derived in section 1.4. We proceed as follows: first, we combine model estimates to estimate the value of the additional public goods provided through

Figure 1.7: House price effect by quartile of change in public expenditures.



Note: Diff-in-diff coefficients of log(sales price) on four annexation indicators, each capturing a quartile of the baseline (i.e. before 1993) difference in municipal and county per capita expenditures. The observations are areas (i.e. neighborhoods defined by municipal boundary changes) for the 189 municipalities for which we observe the complete boundary change history in California. Standard errors clustered on city level.

annexation. We do this under a number of different assumptions about the extent of preference heterogeneity. Second, we show the implications of these assumptions for valuation per affected household. Finally, we quantify the overall incidence of the 502 annexations we observe in our data over the course of 1988-2013 for households, landowners and governments.

In terms of the model parameters of interest, the empirical results suggest that while there is a fairly tight range of plausible estimates, there is still some flexibility which estimate to employ for the model exercise. We opt for estimates near the median so as to ensure our main conclusions on valuation, incidence and welfare are not driven by implausibly large estimates. Concretely, we use for the housing supply effect the baseline estimate in table 1.2, which is $\widehat{\Delta \ln N} = 38\%$; for price effect, we use the lowest estimate from the individual sales estimates in table 1.4, which is $\widehat{\Delta p} = \$1,471$; and for the public goods effect we use the baseline estimate in table 1.5, which is \$481, which we multiply by the average number of people per household in our data (2.8), to arrive at our parameter estimate of $\widehat{\Delta g} = \$1,347$.

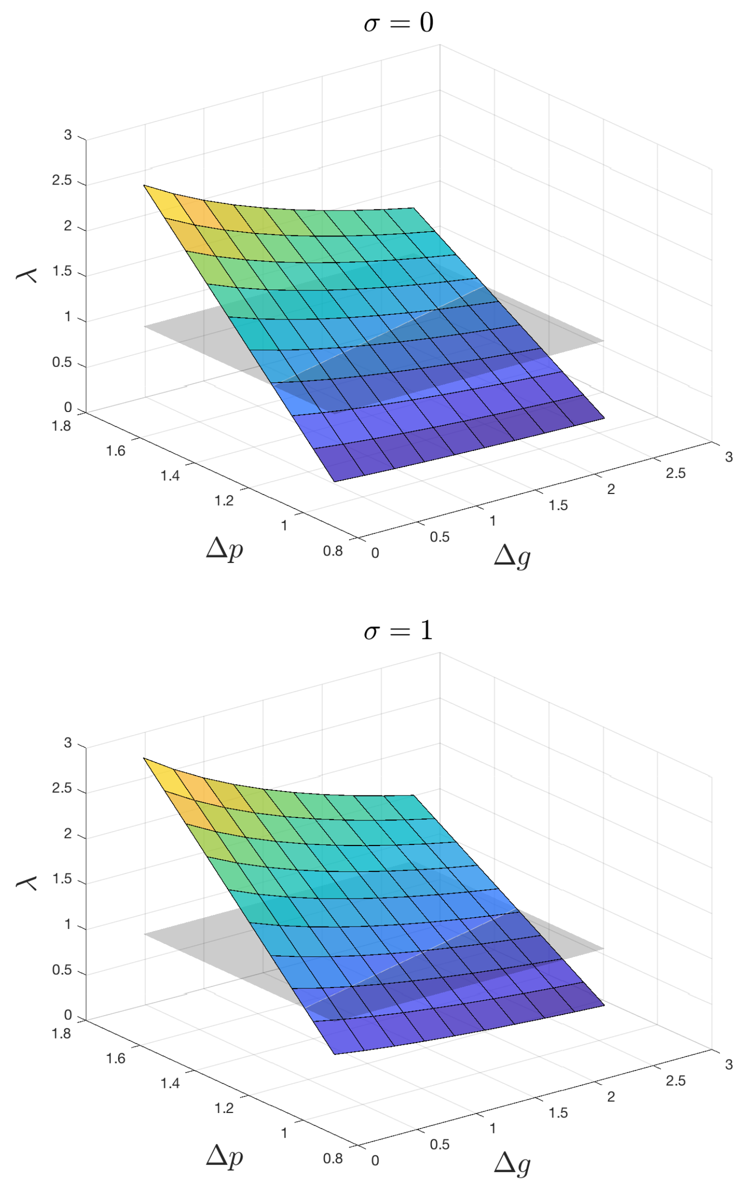
To complete the model estimates, we require a magnitude for σ. Ideally, we'd be able to estimate this directly from the data, or to calibrate it using estimates from previous work. Unfortunately, we lack variation to separately identify σ directly, and since the elasticity of substitution between areas is not scale free, we cannot simply use an estimate from previous work, since no other work uses the fine-grained areas for the cities included in our data. Instead, in the absence of a direct estimate, we allow σ to take on a range of plausible values beginning at zero, which corresponds to homogeneous preferences. To define a plausible upper bound for σ, we use the magnitude of the annual price effect of annexation as a guide. The standard deviation of ε_{ij} equals $\pi/\sqrt{6}\sigma$, which means that a value of 2,000 would correspond to a standard deviation in the annual willingness to pay to live in a particular area of around \$2,566, which is about 75% larger than the price effect of annexation. As two further points of comparison, $\sigma = 2$ (in thousands) corresponds to about 10% of annual mortgage payments over this period, and 76% of annual property tax payments in our data. We thus use $\sigma = \{0, 0.5, 1, 2\}$ in thousands as our calibrated values.

We first turn our attention to the value of public goods as implied by our estimates of the supply effect, the price effect and the public goods effect. Figure 1.8 shows the range of plausible estimates for two values of σ. Each of the surface plots maps out the 90% confidence intervals for Δp and Δg, with the corresponding estimate of λ plotted on the z-axis, and the gray transparent surface delineates the points on which the value of a dollar of public goods expenditures is exactly valued at one dollar.

In the top panel, we show implied values for λ with $\sigma = 0$, that is, under the assumption that preferences for areas are homogeneous. The resulting surface suggests that one dollar of public goods expenditure per household is valued at about half a dollar in the least favorable quadrant of the plane (that is, low price effect reflecting low willingness to pay for the additional public goods, and high public goods effect, reflecting a substantial increase in public goods) up to more than 2.5 dollars in the most favorable quadrant (high price effect and low public goods effect). The point estimates imply a value of λ very close to one.

In the bottom panel, we show the same surface but with $\sigma = 1$, reflecting moderate heterogeneity in the preferences for areas. As a result of the heterogeneity, the plane is shifted upwards with a magnitude determined by the product of σ and the supply effect $\Delta \ln N$. We can see that the gray trapezoid hovering over values below one has shrunk substantially, yielding estimates of λ that are greater than one throughout most of the range of the parameter estimates. The intuition for why heterogeneous preferences for areas drive up the implied value of public goods is the following: Households accrue economic rents from their initial residential choice,

Figure 1.8: Surface plot of public goods valuation.



Note: Surface plots of the estimate for the value of local public goods λ. In each panel, the annual price estimate of annexation Δp and the annual public goods estimate Δg make up the plane, with corresponding values of λ plot on the z-axis. The transparent gray plane delineates at what point λ passes a value of one. In the top panel, we show estimates assuming that there preferences for areas are homogeneous ($\sigma = 0$), and in the bottom panel we assume heterogeneous preferences with a standard deviation about the size of the annual price effect. Values for Δp, Δg and σ in thousands of dollars.

Table 1.7: Household valuation and incidence.

	$\sigma = 0$	$\sigma = 0.5$	$\sigma = 1$	$\sigma = 2$
Household valuation: λ	1.04	1.18	1.32	1.61
Household incidence: ΔV	\$0	\$192	\$383	\$766

Note: Household valuation and incidence with calibrated elasticity of substitution between areas. Household incidence expressed in per household terms. The values for σ are in thousands of dollars.

meaning that most households are inframarginal. The intensity with which they relocate into an area reflects their willingness to give up these rents in exchange for the benefits (better public goods net of living cost) of living in the annexed area.

We summarize these results in table 1.7. For the calibrated values of σ, we get a household valuation of public goods ranging from 1.04 to 1.61. While these estimates concentrate around one or slightly above, we cannot reject the possibility that λ is smaller than one; there are certainly plausible parameter combinations that would yield such values. However, since the weight of evidence leans towards values greater than one, we think of this as suggestive evidence that the additional public goods provided through municipal annexation are valued more highly than the cost of providing them.

The second row of table 1.7 also shows the implied per household incidence from annexation. As shown in equation (1.6), this is simply the supply effect (i.e. the increased willingness for households to move into an area after annexation) scaled by the elasticity of substitution σ. When σ is zero, households derive no economic rents from their residential choice, and there is no way for the annexation to have a welfare effect on households. In contrast to this, once we allow σ to be nonzero, we see household incidence increase into the range of several hundred dollars due to annexation. This amount reflects the willingness of the average household to forego the economic rents they accrue in exchange for the benefits from the increase in public goods net of living costs.

Finally, with some additional assumptions, we can evaluate total household incidence, landowner incidence and government budget incidence as implied by equations (1.6), (1.7) and (1.8). We use the total number of affected households (51,540), baseline house prices (\$277,268), baseline public goods expenditure (\$400) and the fixed property rate (1%) and combine these with three increasingly optimistic scenarios about the remaining parameters σ, κ and τ_s, starting from pessimistic assumptions. Baseline scenarios are based on data for the ten years preceding each annexation. We use estimates of Δp and $\Delta \ln N$ to infer the average impact on price and housing quantities.

Table 1.8: Incidence and welfare.

	Low	**Med**	**High**
Household incidence: ΔV	0	20	39
Landowner incidence: $\Delta \Pi$	385	442	500
Government incidence: ΔR	-41	-13	12
Welfare effect: ΔW	344	449	551

Note: Incidence of annexations on households, landlords and governments for all 502 annexations over 1988-2013 in our sample of Californian municipalities. Columns reflect pessimistic (Low), average (Med), and optimistic (High) scenarios. σ takes on values 0, 1 and 2 (in thousands) in the three scenarios. Housing supply elasticity κ is inferred from profit margin data in Taylor (2015) describing net profit rates of housing construction over eight years, with values of 3.6%, 6% and 8.3% in the three scenarios. We take the share of property taxes accruing to counties and municipalities to be 31%, 44% and 56% in the three scenarios. All values in millions of dollars.

We assume the following values for these parameters in each of the respective scenarios: we let σ take on values 0, 1 and 2 (in thousands), as discussed above. To infer a housing supply elasticity κ, we use data from Taylor (2015) describing net profit rates of housing construction over eight years, with values ranging from 3.6-8.3%. We use these two values as well as the average of the two and solve for the corresponding housing supply elasticity using the property tax rate to arrive at values for κ ranging from 0.069-0.092. Finally, we take the share of property taxes accruing to counties and municipalities to be between 31% and 56%, based on data from Coleman (2015).

Results of the resulting welfare analysis are shown in table 1.8. The first row shows the total incidence on affected households under the three scenarios. As discussed earlier, the incidence is zero when preferences are homogeneous. In the average and optimistic scenario, household incidence is \$20 million and \$30 million respectively. The next row shows landowner incidence. It can be see that this makes up the bulk of welfare increases from annexation. Even in the pessimistic scenario, landowners made \$385 million in profits from annexation. This rises up to \$500 million in the optimistic scenario. Finally, we examine the impact on joint government budgets, finding that local governments make losses in two out of the three scenarios, at \$41 million and \$13 million, respectively. Only in the optimistic scenario, the impact of annexations on government budget is positive at \$12 million. These magnitudes are small for compared to the total budget of the 71 municipalities and corresponding counties that engaged in annexations over this period, which is on the order of \$110 billion.

The total welfare effect of the 502 annexations observed in California over 1988-

2013 is simply the sum of these incidence components, which ranges from $344 million to $551 million. Importantly, landowners reap nearly all benefits. In the pessimistic scenario, the welfare impact would be negative were it not for the profits accruing to landowners. Even in the optimistic scenario, they capture almost 90% of the surplus generated through annexations.

1.8 Extension: Benchmarking Annexations

So far, we have shown that (a) local public goods are valued by households about dollar-for-dollar or slightly higher, (b) extending these public goods generates substantial benefits, and (c) most of these benefits accrue to landowners. We have arrived at these conclusions using variation generated by municipal annexations, exploiting the differences in service provision between municipalities and counties. In this final section, we address the external validity of our estimates by benchmarking differences between municipalities and counties to those between school districts and between municipalities.

Ideally, we would study similar boundary changes in which municipalities take over territory from other municipalities, and school districts take over territory from other school districts. However, these types of boundary changes are extremely rare. Instead, we benchmark our estimates using static differences across jurisdictional boundaries. Specifically, we estimate house price discontinuities across school district boundaries, across municipal boundaries, and across municipality-county boundaries, using the approach pioneered by Black (1999). We first show how price discontinuities between adjacent school districts compare to those between municipalities that differ in the quality of public goods provided: the average achievement of students in the school district, and the crime burden faced by the police department; we then discuss how our dynamic long-run estimates of the price effect of annexation is broadly comparable to the price discontinuity across municipality-county boundaries.

This exercise faces two key challenges. First, boundaries between adjacent jurisdictions may separate more than one type of local government. For instance, in New England, the boundaries of municipalities and school districts typically coincide perfectly; thus, it would be unclear whether any price discontinuity would be due to differences in the quality of the school district or in the quality of the municipality. However, it turns out that outside of the Northeast, municipal boundaries and school district boundaries frequently do *not* coincide, creating spatial discontinuities that uniquely differ in one type of jurisdiction, holding other jurisdictions constant. As an example, figure 1.2 shows all jurisdictional boundaries for Los Angeles county. It can be seen that municipal boundaries (in blue) frequently do not overlap with school district boundaries (in green).

The second challenge to this exercise is household sorting across boundaries. This issue has been studied carefully by Bayer et al. (2007), showing that above and beyond controlling for neighborhood amenities as captured by boundary fixed effects as in Black (1999), controlling for neighborhood socio-economic characteristics further lowers discontinuities across boundaries. We address this challenge by controlling for an array of house characteristics and Census block group characteristics capturing fine-grained variation in neighborhood demographics. While this is not as powerful as using restricted Census micro data as in Bayer et al. (2007), to the extent to which sorting across one type of boundaries is broadly comparable to sorting across other types of boundaries, the relative size of the price discontinuities is informative about the relative value of the public goods provided and thus the external validity of our annexation setting.

Data and Approach

To estimate house price discontinuities across these three types of boundaries – between school districts, between municipalities, and between municipalities and counties – we assemble a new dataset consisting of four components: jurisdictional boundaries, jurisdictional quality measures, Census block group data, and assessment records of houses located near jurisdictional boundaries.

We use jurisdictional boundaries from two sources: the 2010 TIGER/Line county and place boundaries for counties and municipalities, and the 2010 file of the NCES School District Boundaries database. To measure the quality of the associated jurisdictions, we use crime rates calculated from the Uniform Crime Reporting (UCR) statistics averaged over the years 2005-2014 for municipalities. For school districts, we use assessment proficiency in Mathematics and Reading/Language Arts averaged over 2010-2015 (2010 is the first year of data available), specifically the average share of students who scored above proficient on state assessments across math and English. We use five-year averages of Census block group data from the American Community Survey (ACS), including information on each block group on total population, population below 17, population above 65, block group size, share hispanic, share black, average household income, and shares of five educational achievement categories (less than high school, high school, some college, college, postgrad). Finally, we again use DataQuick to extract information on 36 million single-family home locations, sales price, and house characteristics (age, number of rooms, number of bathrooms, lot size, square feet of building space). All house price sales are transformed into 2010 real dollars.

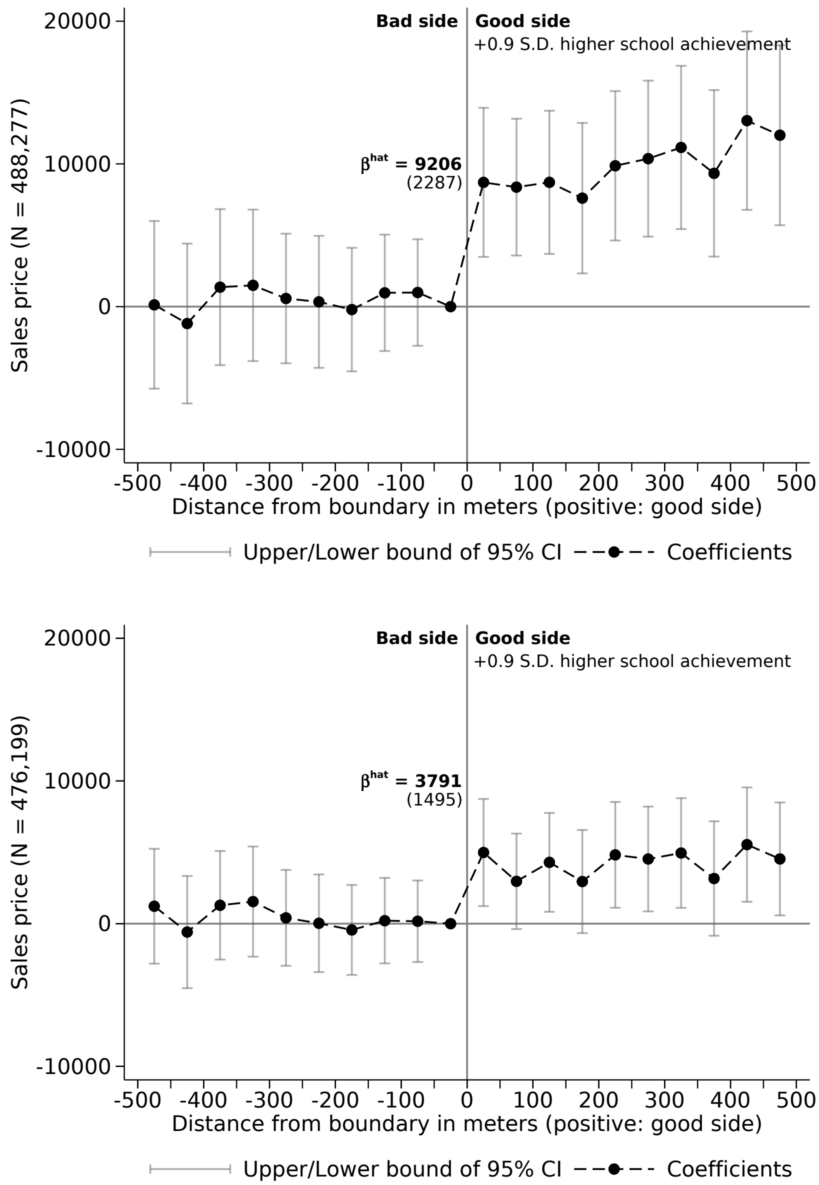
For each of the 36 million properties, we spatially match it to the corresponding county, school district, and municipality (unless the property is in unincorporated territory), and we compute the distance of each property to the nearest jurisdictional boundary. We then identify which types of jurisdictions are separated by this boundary, yielding ten categories of boundaries, four of which separate only one layer of local government: municipality-municipality boundaries, municipality-county boundaries, county-county boundaries, and boundaries between school districts. The other six categories combine two or more of these categories.[4]

To estimate the house price discontinuities of interest, we restrict our sample of properties to those within 500 meters of boundaries separating only municipalities, only school districts, or only municipalities from (county serviced) unincorporated territories. We then flip the distance-to-boundary sign for all properties that are on the worse side of an adjacent jurisdiction pair, such that the discontinuity when crossing a distance of zero reflects an increase in the quality of the jurisdiction. This leaves us with 1.32 million properties from eleven states[5]: 511,924 near school district boundaries; 302,550 near municipal boundaries; and 508,986 near municipality-county boundaries. We are then ready to estimate regressions of the following form for properties indexed by i:

$$y_i = \phi_{b(i)} + \sum_k \mathbf{1}[i \text{ in } k\text{th distance bin}]\beta_k + \mathbf{X}_i\gamma + \varepsilon_i \tag{1.12}$$

where y_i is either the sales price or the net-of-taxes sales price, $\phi_{b(i)}$ is a one-kilometer boundary segment by sales-year fixed effect, β_k is the regression-adjusted mean of 50-meter bin near the boundary, $\mathbf{X}_i$ is a set of hedonic controls and Census block group controls, and ε_i is an error term. The inclusion of boundary-segment by sales-year fixed effects removes variation in the quality of neighborhoods near a particular boundary segment as well as any kind of neighborhood-specific trend in house prices. Standard errors are clustered on both jurisdictions as well as adjacent jurisdiction pairs.

Figure 1.9: Boundary discontinuity around school district boundaries.



Note: Regression-adjusted means of the single-family residence sales price in 50-meter bins near a boundary separating two school districts with the higher achieving district to the right of the boundary, as described in equation (1.12). Both the top and the bottom panel include 1km-boundary-segment by sales-year fixed effects; the bottom panel uses the net-of-taxes sales price and includes hedonic and sociodemographic control variables: age, number of rooms, number of bathrooms, lot size, square feet of building space; total population, population below 17, population above 65, block group size, share hispanic, share black, average household income, and shares of five educational achievement categories.

Results

We begin by conducting a graphical analysis of price discontinuities across the three categories of boundaries, summarized in figures 1.9, 1.10, and 1.11, with price discontinuities summarized in table 1.9. Regression-adjusted mean sales prices near adjacent school districts are displayed in figure 1.9. Having ordered all school district pairs according to student achievement, crossing the boundary from left to right corresponds to a jump in achievement of about 0.9 standard deviations (going from an average of about 62% to 74% proficient across grades, subjects and states). The corresponding price discontinuity is about $9,206 dollars. As shown in the bottom panel of the figure, once we control for house characteristics, Census block group characteristics, and use the net-of-taxes sales price, the discontinuity shrinks considerably, to a value of about $3,791. In both cases, the discontinuity is highly significant, and regression-adjusted means are remarkably stable as one moves further away from the boundary.

These discontinuities around school district boundaries suggest that an approximate one-standard deviation increase in the quality of the school district as measured by student achievement is valued at around $3,791 dollars. While Census block group socio-economic data may deal with some of the potential sorting, the work by Bayer et al. (2007) suggests that about half of the average willingness to pay for better schools may be due to preferences for characteristics of neighbors, which would lower our estimate for the willingness to pay to around $2,000.

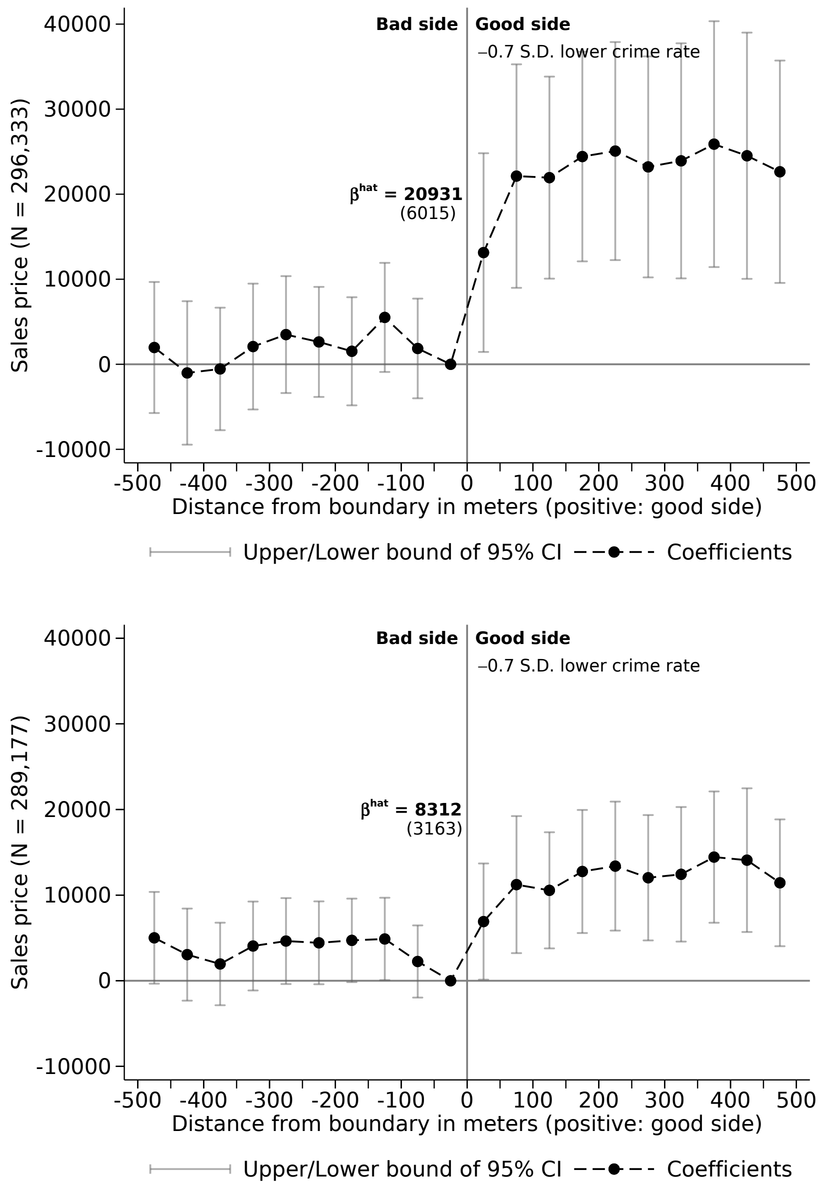
How do these estimates between school districts compare to those between adjacent municipalities? As discussed in section 1.8, we use the crime rate to order municipalities by quality. Even if the frequency of crime is continuous across municipal boundaries, the capacity of the corresponding municipality to respond to crime events may be discontinuous: response times and the quality of the response in case of an emergency may vary considerably. Generally, we use crime rates to order municipalities as a proxy for municipal quality more broadly, in the sense that the quality of other municipal services like street maintenance or fire services is likely to be correlated with the crime rate.

Results for price discontinuities across municipal boundaries can be seen in figure 1.10. Now, moving from the worse to the better side corresponds to a drop in the

[4]Specifically, let M denote municipality, U unincorporated territory, C county, and S school district boundaries. The "separately identified" types of boundaries separating only one layer of local government are M-M, M-U, C-C, and S-S. The six higher dimensional boundaries separate MC-MC, MS-MS, MSC-MSC, MC-UC, MS-US, and MSC-USC.

[5]The eleven states for which we have both high quality real estate data and boundaries that uniquely identify one type of jurisdiction are: Arizona, California, Colorado, Illinois, Michigan, Missouri, Oklahoma, Oregon, Texas, and Washington.

Figure 1.10: Boundary discontinuity around municipal boundaries.



Note: Regression-adjusted means of the single-family residence sales price in 50-meter bins near a boundary separating two municipalities with the lower-crime municipality to the right of the boundary, as described in equation (1.12). Both the top and the bottom panel include 1km-boundary-segment by sales-year fixed effects; the bottom panel uses the net-of-taxes sales price and includes hedonic and sociodemographic control variables: age, number of rooms, number of bathrooms, lot size, square feet of building space; total population, population below 17, population above 65, block group size, share hispanic, share black, average household income, and shares of five educational achievement categories.

Table 1.9: House price discontinuities across different.

	SD-SD boundary		Muni-muni boundary		Muni-county boundary	
	(1)	(2)	(3)	(4)	(5)	(6)
House price discontinuity	9,206***	3,791**	20,931***	8,312***	17,261***	7,203***
	(1,654)	(1,495)	(6,015)	(3,163)	(2,730)	(1,779)
Boundary-segment-year FE	X	X	X	X	X	X
LHS net of PDV taxes		X		X		X
Control variables		X		X		X
N of properties	488,277	476,199	296,333	289,177	484,735	470,526
N of jurisdictions	1,851	1,797	867	860	1,864	1,798
R-squared	0.83	0.84	0.80	0.83	0.80	0.81

Note: House price discontinuities as estimated across the "separately identified" boundaries for school districts, municipalities, and between municipalities and counties. Standard errors are clustered on both jurisdictions as well as adjacent jurisdiction pairs.

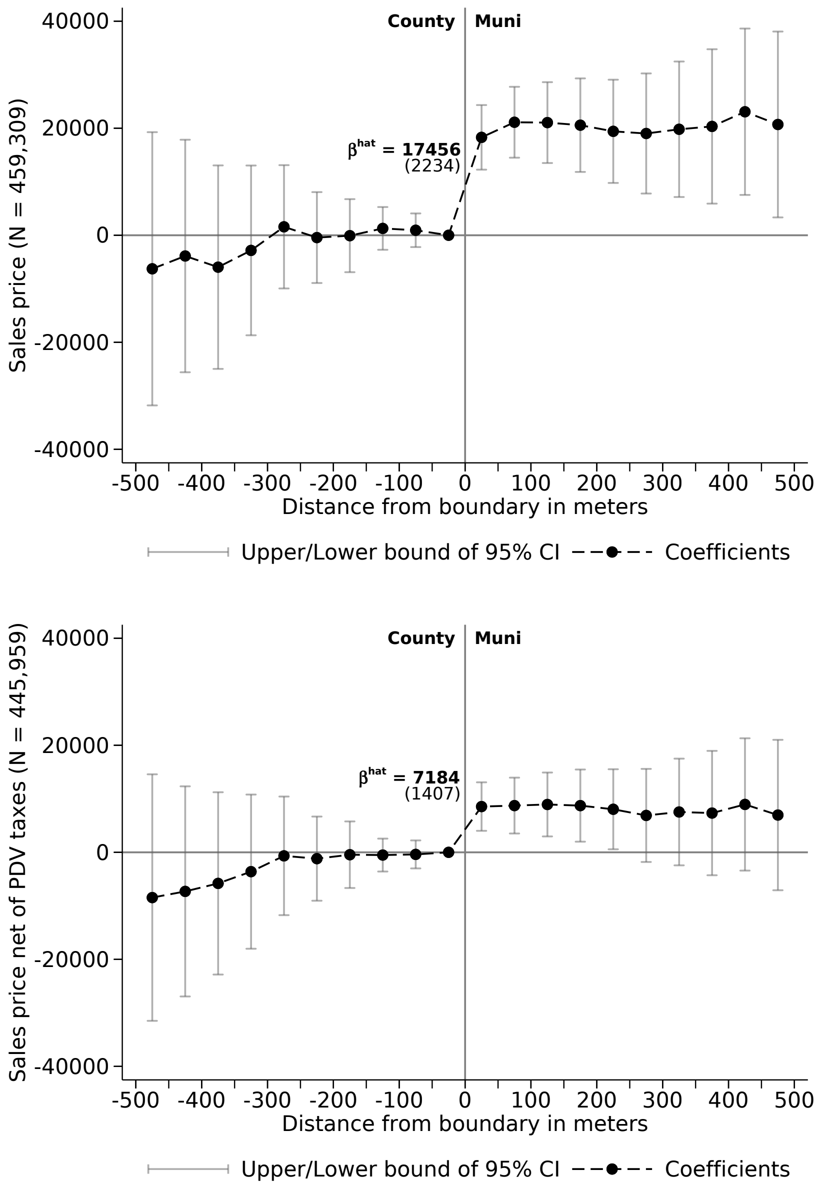
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

crime rate of around 0.7 standard deviations. Remarkably, the price discontinuity is substantially larger than the one between school districts, at $20,931. Even when using net-of-tax sales prices and controlling for housing and demographic characteristics, the discontinuity is still $8,312, about twice as large as the corresponding discontinuity between school districts. If the extent of sorting is similar to school districts, the true average willingness to pay for better municipal services may be in the neighborhood of $4,000. Thus, despite the fact that public goods are more likely to be more non-excludable and non-rival than schooling public goods, we find valuations that seem to be substantially larger than those for school districts.

Having estimated the magnitude of price discontinuities between providers of public goods and those between school districts, we can now compare those magnitudes to those between municipality-county boundaries. In the case of this category of boundary, we do not require to use external data to order it: we can simply put properties in municipalities (presumably providing better public goods) on the right side and those in counties on the left side, such that crossing the boundary from left to right corresponds to entering municipal territory. Price discontinuities for this category of boundaries can be seen in figure 1.11. The estimated discontinuity is of a very similar magnitude to the one between municipalities, at $17,456. After controlling for house characteristics and demographic characteristics, the discontinuity is $7,184, again very similar as to the discontinuity between municipalities.

Finally, we discuss how these cross-sectional estimates of the value of public goods compare to the dynamic estimates from our annexation design. In table 1.3, we esti-

Figure 1.11: Boundary discontinuity around municipal-county boundaries.



Note: Regression-adjusted means of the single-family residence sales price in 50-meter bins near a boundary separating a municipality and a county, as described in equation (1.12). Both the top and the bottom panel include 1km-boundary-segment by sales-year fixed effects as well as linear trends on both sides of the discontinuity; the bottom panel uses the net-of-taxes sales price and includes hedonic and sociodemographic control variables: age, number of rooms, number of bathrooms, lot size, square feet of building space; total population, population below 17, population above 65, block group size, share hispanic, share black, average household income, and shares of five educational achievement categories.

mated long-run average price increases due to annexation of about $9,494. We arrive at similar estimates using the annual price growth estimates and multiplying them by the average number of years we observe areas after annexation (which also roughly corresponds to the number of years before estimates stabilize in our event study in figure 1.5). Comparing these estimates to estimates with controls in table 1.9, we see that the magnitudes from our dynamic estimation procedure is remarkably similar to the ones from our boundary discontinuity approach using municipal boundaries and municipal-county boundaries.[6] Taken together, these results provides suggestive evidence that estimates of the value of local public goods using boundary discontinuities are broadly comparable to those using annexations.

1.9 Conclusion

In this paper, we use municipal annexations to estimate the value of local public goods other than schooling. We find that these local public goods matter to households, possibly as much or even more than schooling. This stands in contrast to the overwhelming attention that schooling receives both in the public perception as well as in the economics literature, relative to other local public goods. The weight of our evidence suggests that public goods are slightly underprovided, possibly due to distributional frictions: unless efficient transfer mechanisms exist, some stakeholders (such as households or local government officials) may oppose welfare-increasing expansions of local public goods.

We arrive at these findings by building a residential choice model with heterogeneous agents that allows us to study the response in optimal residential location, house prices and public goods to changes in jurisdictional arrangements. We show how both house price and housing supply responses inform the value of public goods and estimate these parameters using an event study design. To this end, we collect new data data on administrative boundary changes and combine them with the universe of real estate transaction data in 189 Californian municipalities and their surrounding unincorporated areas. This "big data" approach to public goods highlights the possibility of new insights generated by using the multi-layered spatial structure of local governments together with administrative assessment records on individual properties.

[6] We can also restrict our boundary discontinuity estimates to California to confirm that the presence of Proposition 13 does not lead to substantially different estimates across municipality-county boundaries – despite the fact that unlike in other states, taxes are statistically insignificant across municipality-county boundaries.

These findings raise a number of important questions that should be explored in future research. First, while public goods as a whole are shown to matter, it remains to be seen *which* public goods among the scope of services offered are particularly valued. We have focused on policing and crime since it makes up the largest single expenditure item among public goods, but it is not obvious that other services, such as utilities or roads, are not similarly important or even more important. Annexations may be used directly to identify different changes in magnitudes across different public goods and how these changes are capitalized.

Second, as our interest in this project was on the average marginal valuation of local public goods brought about through municipal annexation, we have not explored the implications of heterogeneous preferences for public goods (as opposed to residential choice). This type of heterogeneity would allow for a more direct connection to the traditional Tiebout literature and address issues of sorting more explicitly.

Third, the efficiency of providing local public goods crucially depends on the extent that economies of scale and spillovers exist. In the work presented here, we focus on the simplest case with constant returns to scale and no spillovers. However, given the richness of the data used here, it may be possible to explore the presence of these forces in this setting in the future as well.

Chapter 2

The Origin of the State: Incentive Compatible Extraction under Environmental Circumscription

2.1 Introduction

Modern human social organization underwent a major transition from stateless societies to states in the course of the last 6,000 years. This transition marks the origin of taxation, law, organized warfare, bureaucratic rule, property rights, the division of labor, and other novel social phenomena, paving the way for modern economic development (Finer, 1997; Bockstette et al., 2002).

Theories of early state formation are as old as the social sciences themselves, with classic writers such as Marx and Smith speculating about the origin of the state. Research in history, sociology and political science in the last forty years has identified extraction by emerging elites in exchange for protection from predation as a key dynamic underlying the process of state formation (Mann, 1986; Tilly, 1990; Olson, 1993). A burgeoning literature in economics formalizes and empirically assesses these central forces and finds supporting evidence for their role in the creation and development of the state (Dal Bó et al., 2016; Mayshar et al., 2017; Sanchez de la Sierra, 2017).

Despite these advances in our understanding of state formation, two major challenges remain: first, while sedentism was widespread among stateless societies, they were typically mobile and frequently relocated in the face of threats or in search for better land. This is evidenced by migratory episodes of groups throughout the Holocene in Africa (Berniell-Lee et al., 2009), Europe (Hofmanová et al., 2016), Asia (Larson et al., 2010) and the Americas (Benson et al., 2007). Mobility limits an emerging elite's ability to extract resources from groups by imposing an incentive compatibility constraint: life under the yoke of the state can be no worse than evasion through migration. This constraint poses a fundamental theoretical challenge to all models of extraction as drivers of state formation.

Second, it remains unclear whether modern state formation theory can empirically account for the geographic pattern of early state formation. Archaeological evidence of the earliest instances of state formation are concentrated in seven places around the world: Mesoamerica, the Andes, West Africa, Egypt, Mesopotamia, the Indus Valley and China (Trigger, 2003). Whether the geographic environment of these early state nurseries is consistent with predictions from state formation models is unknown. This is true despite recent empirical advances because, while the object of interest is early states that arose thousands of years ago, almost all empirical evidence on early state formation in economics rests on data primarily from the 19$^{\text{th}}$ century, at a time when early state development has been severely polluted or completely disrupted by Europeans.[1] Thus, the potential of archaeological data to advance our understanding of the role of extraction in the creation of the early state is immense.

This paper seeks to address both of these challenges in a unified framework. To this end, we first develop a model of extractive state formation with the possibility of evasion through migration. The model identifies conditions of a region and its surrounding land for which both the extractees accept domination by a state and the extractors consider state formation a worthwhile investment. The key intuition of the model invokes an old idea developed in cultural anthropology (Carneiro, 1970): the larger the difference between the quality of the land in a region and its surroundings – that is, the more environmentally circumscribed it is – the more likely is the formation of the state. Since evasion through migration is more costly in circumscribed areas, potential extractors could more easily "cage" (Mann, 1986:p. 38-40) potential extractees. While the argument is essentially cross-sectional, the idea is that the high returns to agriculture and sedentism that emerged in the course of the Holocene opened the door to incentive compatible extraction in some regions sooner than others; the development of pristine states in other regions was cut short by the expansion of early states or their descendants.

The structural parameters of the model are then estimated via probit and logit in a global dataset of archaeological excavation sites related to early state formation and global land quality data from the FAO on a grid with cell size ¼ degrees (about 28km at the equator). We find supporting evidence for the role of circumscription

[1] For example, Fenske (2013:p. 1366) observes that only eight out of 1,267 societies in Murdock's (1967) *Ethnographic Atlas* are observed before 1500.

in predicting the location of early state sites: a one standard deviation increase in circumscription is associated with a 45% increase in the probability of a state site. This effect is as large or even larger than the corresponding effect of land quality itself.

To study the robustness of the effect, we introduce various sets of control variables related to alternative theories of state formation on the role of ecological diversity, regional climate, and other environmental features, as well as $5° \times 5°$ virtual country fixed effects (about 556km squared at the equator). The effect of circumscription becomes larger and more precise with the introduction of more control variables, speaking for important complementarities between circumscription and alternative theories of the early state. Conceptually, this is intuitive: theories of extractive state formation focus on various regional aspects conducive to state formation, many of which are largely orthogonal to the extent of circumscription. Controlling for outside options in the form of circumscription adds precision to alternative explanations of the rise of the state.

To deal with potential endogeneity of circumscription in the specifications so far, we use the interaction of large rivers and arid regions in an instrumental variable strategy. Desert rivers create large differences in land quality between the river valley and the barren land in its vicinity, which turns out to be an important driver of circumscription: the first stage using the interaction of river flow accumulation and the deserts is statistically powerful and robust to a large set of control variables, including rivers and deserts themselves. IV estimates are substantially larger than OLS and precisely estimated, suggesting the presence of attenuation bias in baseline estimates.

We next turn to the heterogeneity of circumscription across civilizations involved in early state formation. Confirming qualitative descriptions of circumscription, we find that the Andean, Egyptian, Mesopotamian and Aztec regions stand out as heavily circumscribed, while others, such as the Maya, are much less so. However, contrary to these qualitative assessments, we find that archaeological sites associated with early states in China are also tightly circumscribed, offering a new avenue for inquiry in early Shang and Zhou state formation.

Finally, the paper analyzes how the strategic environment across different civilizations may contribute to the heterogeneity across civilizations and thereby highlight regions in which extraction was a particularly important motive. To this end, we compare the extent of circumscription across civilizations to the amount of rugged terrain in the vicinity, offering a strategic advantage to challengers. There is substantial variation in the extent to which early state sites of different civilizations are circumscribed and surrounded by rugged terrain, giving rise to the possibility that extraction may have been much more important in some regions (such as Egypt and

China) while other motives such as protection drove state formation in others (such as the Aztec and Maya).

This paper connects to three literatures in economics and the social sciences more broadly. First, it speaks to the literature on early state formation by providing, for the first time, quantitative evidence for the role of circumscription in early state formation. Thereby, it offers the first piece of evidence that modern theories of extractive state formation find empirical support in archaeological data, particularly when considering the outside options of potential state subjects. This finding supports earlier qualitative research by Carneiro (1970) and R. C. Allen (1997) on the role of circumscription in state formation, and more broadly on how opportunities to evade state power diminish state formation and integration (Mann, 1986; Scott, 2009). The role of mobility in state formation also features prominently in Olson (1993). In Olson's terms, this paper cautions that for bandits to become stationary, victims need to be sufficiently immobile too.

Within the economics debate on early state formation, this paper sheds light on the role of the surrounding geographic environment for state formation in a given region. In this sense, it complements various other economic theories of state formation and institutional development. Geography has already been found to be an important factor in early state development, such as Fenske's (2014) research on states providing security for trade across ecologically diverse zones in the African context, as originally suggested by Bates (1987). More broadly, the findings here provide an important mechanism through which geography shapes institutions: extractive institutions are more lucrative in highly productive, constrained places, which may explain why the places that experienced a reversal of fortune after the arrival of Europeans (Acemoglu et al., 2002) acquired these fortunes in the first place, and why they already had extractive institutions in place for Europeans to take over (Dell, 2010).

Extraction plays a key role in economic theories of state formation. In the work by Mayshar et al. (2017), the ease with which agricultural surpluses can be extracted by elites is the driving factor of state formation. This paper comes to a similar conclusion but instead of focusing on the type of crops that are grown, it argues that the extent to which groups could escape into similarly productive land generates variation in early state formation. In this sense, the work presented here is complementary to their findings. Sanchez de la Sierra (2017) finds empirical evidence for the role of both extraction and protection in state formation using Eastern Congo as a quasi pre-state environment. He argues compellingly for the difficulty of using archaeological data from early states directly to study state formation quantitatively due to the absence of systematic disaggregated data. This paper argues that, despite these difficulties, direct archaeological evidence can be embedded in topographic, climatic and other environmental data to produce a coherent picture of early state formation. There is no perfect substitute for the pristine pre-state environment of the mid-Holocene to study the rise of the state due to the looming presence of international markets in raw materials and advanced weaponry in modern stateless environments.

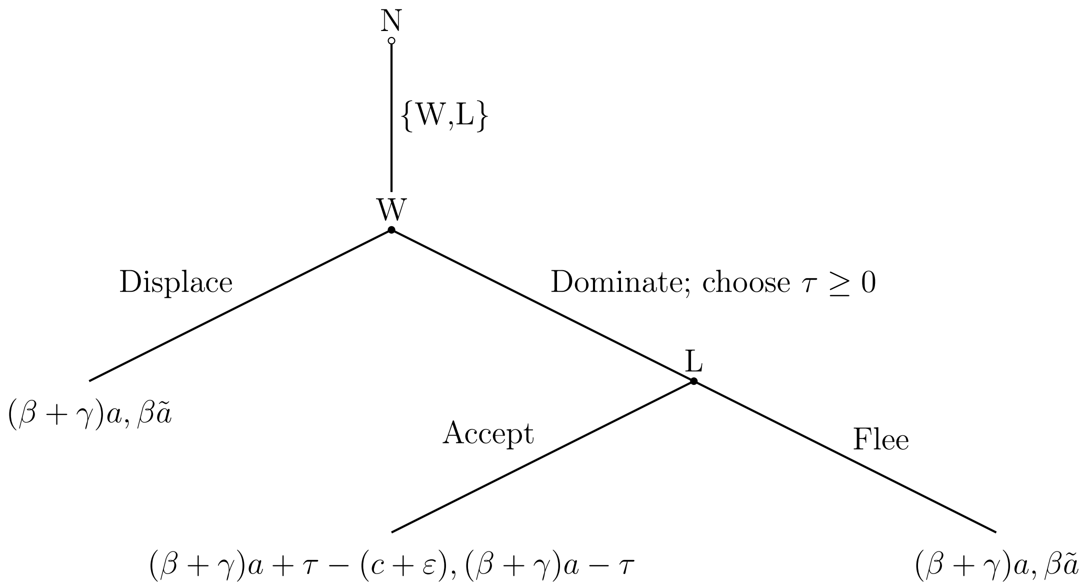
While the focus in economics and related fields is typically on extraction, the study of protection and surplus generation as motives for state formation has a long tradition in archaeology (Childe, 1936; Service, 1975; Haas, 1982; A. W. Johnson and Earle, 2000). Tilly (1990) also famously argued for the role of military capabilities in the development of the state. This paper finds suggestive evidence for the enhanced role of circumscription in regions with more rugged terrain. Dal Bó et al. (2016) deal with the problem of producing and defending surplus in a complementary model to the one presented here, providing an explanation of how states can invest into productive assets if their strategic environment is sufficiently secure.

This paper is also connected to the literature on state capacity in economics, initiated by Besley and Persson (2009); Besley and Persson (2010). Similar to their model, agents choose to invest into state capacity, should the environmental and strategic environment afford it. Incentives to invest in state capacity may also depend on the extent of external threats, as in Gennaioli and Voth (2015).

Historical state capacity and institutions more broadly have been found to be important predictors of long-run development (Bockstette et al., 2002; Acemoglu et al., 2005; Michalopoulos and Papaioannou, 2013). The work presented here provides an explanation of how prehistoric institutions arose. The starting point for these institutions may be found in the spatial configuration of land quality and opportunities for challengers.

The rest of this paper is organized as follows. Section 2.2 develops a simple game of extractive state formation based on circumscription (Carneiro, 1970). Section 2.3 describes the environmental and archaeological data and how circumscription is computed. Next, section 2.4 introduces the baseline estimation strategy derived from the model prediction, presents estimation results and robustness to various alternative theories of state formation. Then, section 2.5 lays out the instrumental variable strategy to deal with the potential endogeneity of circumscription. Section 2.6 studies the heterogeneity of circumscription across civilizations and presents suggestive evidence on the relative importance of extraction and other motives such as protection. Finally, section 2.7 concludes.

Figure 2.1: A dynamic state formation game based on Carneiro (1970).



Note: Dominate the loser of conflict (that is, construct the extractive state) if the agricultural gradient between the shared land and the outside option is large enough. The payoffs listed first are the winner's.

2.2 Model

This section describes a dynamic game of extractive state formation under varying outside options. It is heavily indebted to Carneiro (1970), whose qualitative theory of state formation finds its formal structure here. The purpose of the model is to outline conditions under which both the loser of a conflict is willing to accept political subordination as well as the winner is willing to invest into state capacity, the latter component mirroring ideas first developed in Besley and Persson (2009).

The game is illustrated in figure 2.1. An incumbent group engaged in mixed subsistence farming inhabits a region with agricultural land quality a. The region is surrounded by land of quality $\tilde{a}$. Groups derive a general rate β from any land they work, no matter whether it is their home region or the surrounding region. They additionally derive a home rate γ from their home land, reflecting either a premium: for instance, for their knowledge of the peculiarities of the land, or for fixed infrastructure they put in place to make the land more amenable to farming; or a penalty: for instance, for overworking the land and exhausting its productivity, or for facing off threats particular to the home region.

Separating the general rate β from the home rate γ allows us to study how state formation depends on the land quality calculus of a region versus its surrounding land: if β is zero, extraction capacity only depends on the quality of land in a given region; if γ is zero, it depends on the relative quality in the home region versus the surrounding region; and in the extreme, if γ is negative, it may depend only on the quality of the surrounding land.

As populations grow in the course of the favorable climate of the mid-Holocene, a challenger (internal or external) arrives to compete over arable land, acquiring the skills to benefit from the home premium in the process. When the two groups clash to determine who gets to farm the best land, nature determines the winner and the loser of the conflict. The victory in conflict elevates the winning group to the status of an emerging elite with the potential for extraction.

The winner then chooses between displacing the losing group or attempting to dominate the loser, exacting a tribute of chosen magnitude $\tau \geq 0$ and paying a fixed cost c plus a stochastic component ε to build and maintain an infrastructure to monitor and enforce the payment of the tribute. The stochastic component is mean zero and distributed according to the distribution function $F(\cdot)$. It is known to the winner when choosing whether to displace or dominate; it captures idiosyncratic aspects of domination left outside the model and will be discussed further in the next section. The infrastructure for tribute extraction – consisting of capacities such as keeping track of tribute paid through a specialized bureaucracy or maintaining a monopoly of violence over the losing group – constitute the origins of the extractive state. This is similar to Besley and Persson (2009); Besley and Persson (2010) whose interest is the state's taxation capacity.

If the winning group attempts to dominate the losing group, the latter can either accept domination, staying in the winner's home region and sharing in the home premium γ with the winner, but paying the tribute τ to the winner. Alternatively, the group can flee from the area to the surrounding area with quality $\tilde{a}$, which amounts to the same outcome as being displaced.

We solve for the subgame-perfect equilibrium of the game by backward induction. For a given $\tau \geq 0$, the loser will accept domination if and only if $(\beta + \gamma)a - \tau \geq \beta\tilde{a}$. Otherwise, it will flee to the surrounding region. This is an incentive compatibility constraint: since there is no way to physically constrain the losing group, no more than the surplus from the home region relative to the surrounding region can be extracted for the losing group to stay. If this incentive compatibility condition holds, it determines the optimal tribute chosen by the winner in case of domination by making the loser indifferent between domination and flight: $\tau^* = (a - \tilde{a})\beta + a\gamma$. Given this optimal tribute τ^*, the winner will attempt domination if and only if the return from tribute extraction is greater than the investment cost into state capacity:

$(\beta+\gamma)a+\tau^*-(c+\varepsilon)\geq(\beta+\gamma)a$, which yields a simple participation condition for the winner to form the extractive state:

$$(a-\tilde{a})\beta+a\gamma\geq c+\varepsilon \tag{2.1}$$

We have thus proven the following proposition:

Proposition 1 (Extraction under Circumscription)**.** Assume the setup of the game along the lines of the game in figure 2.1, and let Y be a binary random variable evaluating at unity if the extractive state is formed. There are three subgame-perfect equilibria:

1. {Dominate with $\tau^*=(a-\tilde{a})\beta+a\gamma$, Accept} if $c+\varepsilon\leq(a-\tilde{a})\beta+a\gamma$,
2. {Displace, Accept} if $c+\varepsilon>(a-\tilde{a})\beta+a\gamma\geq 0$,
3. {Displace, Flee} if $(a-\tilde{a})\beta+a\gamma<0$.

The probability of state formation is $\Pr(Y=1|a,\tilde{a})=F((a-\tilde{a})\beta+a\gamma-c)$. The size of the extractive state τ^* and the probability of state formation increase with higher land quality gradient between home region and surrounding region $a-\tilde{a}$ and higher land quality a in the home region; and the state formation probability falls with higher investment costs of extractive state capacity c.

The core intuition of the model is that a high gradient in agricultural land quality (i.e. the region is "circumscribed" by lower quality land) means that the losing group faces a relatively worse outside option from fleeing. It will thus be more willing to accept domination through a state if it resides in a place with high quality land if it is surrounded by low quality land. This captures the essence of Carneiro's (1970) circumscription theory.[2]

While the basic intuition of the model is fairly straightforward, the model clarifies the relationship between environmental circumscription – which is a function of land quality – and land quality itself: controlling for land quality, circumscription captures the extent to which evasion through migration matters for early state formation. In addition to the gradient in land quality affecting the probability of state creation, land quality also has a direct effect through the size of the home rate γa. More

[2]Interestingly, the model suggests that for some parameter values (the second equilibrium in the list), the loser would be better off if the state were formed; however, the surplus produced in the home region is not large enough to pay the fixed cost of state formation, and so the winner prefers to displace the loser. It seems this possibility has been overlooked so far in the literature on circumscription.

states should form if groups are more strongly attached to their land due to specific knowledge or infrastructure. On the other hand, if regions with high land quality face frequent challenges from outsiders (meaning that γ is negative), less states may form there.

While state formation is modeled as a binary outcome, the underlying latent magnitude of extraction τ^* quantifies the extent to which extraction is a motive for state formation. In section 2.6, we discuss the intensity of extraction across civilizations in more detail.

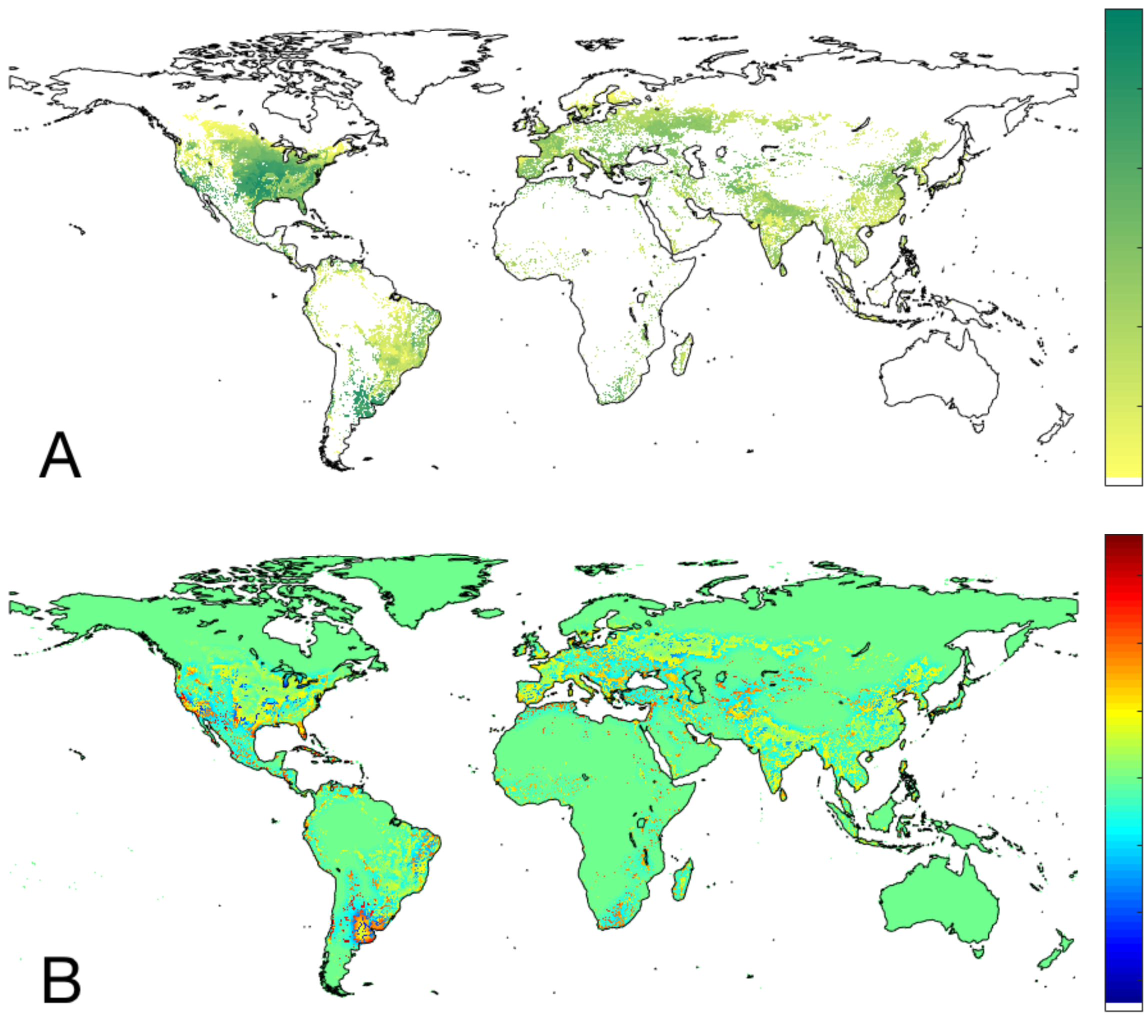
The focus on competition over arable land of semi-nomadic groups engaged in some agricultural production in the model deserves further discussion. The groups in the model conduct some small-scale agriculture, but they may still be engaged in hunting and gathering, either as part of their nutritional intake throughout the year or for part of the season. The transition to full-scale agriculture took several thousand years in some regions (Arranz-Otaegui et al., 2016), during which groups remained semi-nomadic; they often brought their domesticated crops with them across large migratory distances (Erickson et al., 2005; Crowther et al., 2016). After the initial domestication period, agriculture spread and became available almost everywhere on the globe (J. Diamond and Bellwood, 2003), often displacing pure hunter-gatherer groups (Hofmanová et al., 2016).

2.3 Data

To empirically assess the importance of environmental circumscription for early state formation, we combine data on agricultural land quality, archaeological data on excavation sites associated with early state formation, as well as a number of other data describing the climatic and topographic environment. We describe each of these in turn.

As a proxy for agricultural land quality, we use data on the maximal potential production capacity in t/ha over seventeen crops from the FAO's Global Agroecological Zones (GAEZ) database (Fischer et al., 2008), scaled by historical calories per ton for each crop by the FAO (Chatfield, 1953) and corrected for the Columbian Exchange (Nunn and Qian, 2010). This proxy is very similar to the Caloric Suitability Index developed by Galor and Özak (2016), with one important difference: since we are interested in identifying steep gradients in land quality, we want to allow for alluvial agriculture based on rivers in addition to rainfed agriculture. The GAEZ database and covers total crop production value for all important crops across the globe on a 5-minute resolution grid, totaling 9,331,200 cells. Figure 2.2 shows a map of the resulting agricultural land quality data; details about data construction are described in the appendix.

Figure 2.2: Map of (A) land quality and (B) circumscription.



Note: Panel A: Land quality proxy: maximal potential production capacity (FAO GAEZ database) over the 18 most important crops, scaled by calories of each crop and corrected for the Columbian Exchange, in standard deviation units, with cell size ¼. Panel B: Circumscription with cell size ¼ and $r = 10$.

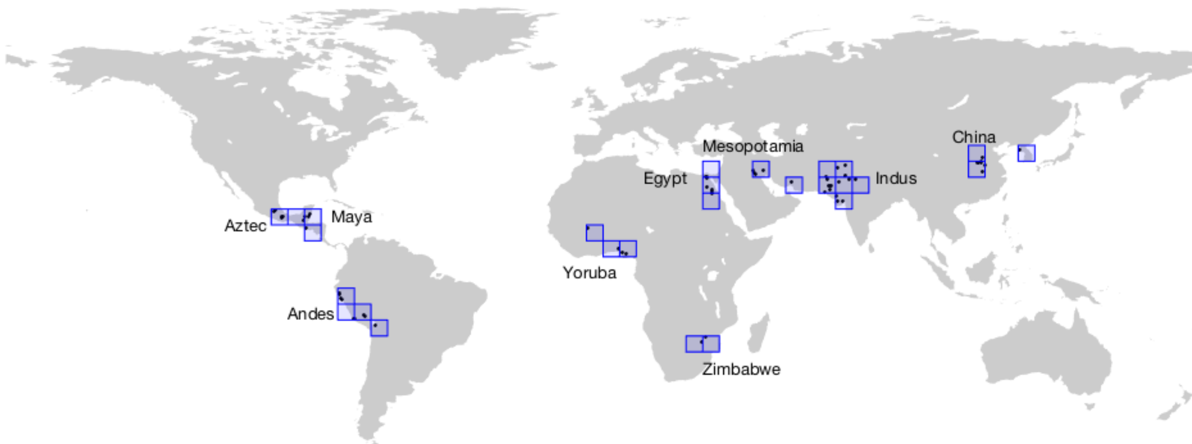
To compute the intensity of circumscription between a location and its surrounding land, we compute the difference between the the quality in a given cell and the average quality of all surrounding cells at various radii (see section 2.4). In contrast to land quality per se, circumscription intensity may be relatively low in large fertile expanses such as the American Midwest or Western Russia since most cells are surrounded by equally productive cells; on the other hand, circumscription intensity may be high in areas that are agriculturally productive but confined in otherwise barren land or surrounded by oceans, such as the river valleys of the Nile or the Indus, or the islands of Cuba or Sicily. More details about the construction of circumscription intensity are discussed in the appendix.

The data on the location of early states comes from the archaeological record presented in Bogucki (1999) and shown in figure 2.3. States are defined as "powerful, complex, institutionalized hierarchies of public decision-making and control" based on (Brumfiel, 1994). In his book, several regional maps show key sites involved in the creation of the first states across Africa, Asia and the Americas. This selection of early state sites captures the broad consensus of archaeological sites in what are traditionally considered the geographic centers of early state creation: Mesopotamia, Mesoamerica, the Andes, Egypt, the Indus valley, and China; this collection thus includes both groups of city states and territorial states (Trigger, 2003). He also includes sites in West Africa (Yoruba) and Great Zimbabwe. The civilizations in these regions are frequently referred to as single territorial states, but archaeological and historical evidence points towards systems of city states or small territorial states, at least at the beginning of their development (Finer, 1997; Trigger, 2003). While this selection of early state sites is by no means exhaustive and some of these archaeological sites play a larger role in each region's formation of early states than others, they are the most prominent sites to provide key evidence of public architecture, kingship, or urban elites typically associated with early states, and as such each of these sites' locations are indicative of the environment in which early states formed.

In addition to agricultural land quality, a number of other climatic, topographic and environmental factors may be more or less conducive to state formation. Included are annual mean and standard deviation of temperature from the from the WorldClim global climate database (Hijmans et al., 2005); river flow freshwater accumulation from the HydroSHEDS database (Lehner et al., 2008); potential vegetation from SAGE (Ramankutty and Foley, 1999); and topographic slope and ruggedness (Riley et al., 1999; Burchfield et al., 2006; Nunn and Puga, 2012). We address in section 2.4 how each of these relate to alternative theories of early state formation.

Table 2.1 summarizes descriptive statistics of all the data. First, it can be seen that given the rarity of early state formation, only very few cells (0.03%, a total of 60 cells) have a state site in them. Noteworthy is also that the median cell is negatively circumscribed, meaning it is surrounded by land on average of slightly better quality than itself; but mean circumscription is positive with a magnitude of about 9% of

Figure 2.3: Map of archaeological sites from (Bogucki, 1999).



Note: Archaeological sites of early states (Bogucki, 1999) are shown as black dots. The blue rectangles show the twenty-seven 5° × 5° virtual countries which have at least one early state site. The names of civilizations are common shorthands used for their respective regions.

land quality. This skew is due to the fact that most cells have at least some cell in their vicinity with some positive land quality, although the majority of cells has land quality zero. Naturally, the correlation between land quality and circumscription is high at 0.74.

2.4 Baseline Estimation

In this section, we take proposition 1 to the data. To this end, we divide the entire landmass of the globe into cells with size ¼ degrees (28km at the equator) and index them by $i = 1, ..., N$, which results in 194,102 cells. The outcome variable is an indicator Y_i taking on unity if cell i includes a site from Bogucki's sample of early state locations. From the FAO data, we know agricultural land quality a_i and circumscription $a_i - \tilde{a}_{ir}$, where

$$\tilde{a}_{ir} = \frac{1}{N_r}\left[\sum_{j \neq i : \text{dist}(i,j) \leq r} a_j\right]$$

for radius r in cell units. N_r is the number of cells (minus the cell at i) in the circle with radius r, as explained in the data section. When we write $\tilde{a}_i$ without a radius

Table 2.1: Summary statistics for global grid dataset.

Variable	Mean	Median	Std.dev.	Min	Max
State site indicator	0.00033	0	0.018	0	1
Circumscription in MCal/ha	370.2	-120.6	5348.5	-31280.6	32676.2
Land quality in MCal/ha	4034.4	0	7818.9	0	40722.8
S.d. land quality in MCal/ha	4155.7	3870.0	3489.7	0	16668.8
Ecological diversity index	0.29	0.20	0.28	0	0.87
Ecological polarization index	0.37	0.36	0.31	0	0.99
Mean of annual temperature in C°	14.5	17.0	10.7	-15.4	31.4
Std.dev. of annual temperature in C°	6.95	6.45	4.69	0.12	22.1
abs(latitude)	31.1	30.9	16.7	0.12	62.1
Slope in percent	2.89	0.86	4.95	0.0090	92.8
Ruggedness in meters	98.0	29.9	166.0	1	2660.0
log(Riverflow accumulation in km^2 cells)	4.10	3.80	2.07	-15.4	12.9
Desert indicator	0.12	0	0.33	0	1
N	177,299				

Note: State site indicator is one if there is at least one archaeological site relevant to state formation in it. Land quality is measured as maximal megacalories (MCal) per hectare (ha) across the seventeen crops in the FAO database, using historical calorie tables to scale them and correcting for the Columbian Exchange. See text on how circumscription is computed. The radius of $r = 10$ applies to circumscription, std.dev. land quality, the ecological diversity index, and the ecological polarization index. The number of observations is lower than the number of land cells with land quailty data (194,102) because riverflow accumulation does not cover some parts north of 60°. More details in the appendix.

subscript, we are using our default radius at $r = 10$, resulting in a radius of about 280km at cell size ¼.

If we assume that the stochastic component ε is distributed as $\varepsilon \sim \mathcal{N}(0, 1)$, i.e. standard normal, or that $\varepsilon \sim$ logistic$(0, 1)$, then we can directly estimate the structural parameters (β, γ, c) from the model. That is, we run

$$\Pr(Y_i = 1|a_i, \tilde{a}_i) = F(-c + (a_i - \tilde{a}_i)\beta + a_i\gamma) \tag{2.2}$$

with $F(\cdot) = \Phi(\cdot)$ or $F(\cdot) = \Lambda(\cdot) \equiv \exp(\cdot)/(1 + \exp(\cdot))$. This is the key empirical prediction from proposition 1. Results are shown in table 2.2. We run three variations of the model with both probit and logit: in column (1) and (4), we force the home rate γ to be zero, studying the role of circumscription in isolation; in column (2) and (5), we force the general rate β to be zero, focusing only on the land quality of the home cell; in column (3) and (6) we estimate both of them simultaneously.

In terms of the model structure, the estimates in table 2.2 show that both the general rate $\hat{\beta}$ and the home rate $\hat{\gamma}$ are strongly associated with the presence of a

Table 2.2: Probit/Logit estimation of structural parameters (β, γ, c).

	(1)	(2)	(3)	(4)	(5)	(6)
Circumscription: $\hat{\beta}$	0.16***		0.13***	0.56***		0.46***
	(0.03)		(0.05)	(0.08)		(0.16)
Land quality: $\hat{\gamma}$		0.14***	0.04		0.49***	0.12
		(0.02)	(0.04)		(0.06)	(0.13)
investment cost: $-\hat{c}$	-3.48***	-3.53***	-3.50***	-8.28***	-8.44***	-8.33***
	(0.04)	(0.04)	(0.04)	(0.15)	(0.15)	(0.15)
Estimator	Probit	Probit	Probit	Logit	Logit	Logit
Mean dependent variable	0.00033	0.00033	0.00033	0.00033	0.00033	0.00033
Marginal effect of circumscription	0.00019		0.00015	0.00018		0.00015
Marginal effect of land quality		0.00017	0.00005		0.00016	0.00004
Test $\hat{\beta} = \hat{\gamma}$			0.28			0.22
Log-Likelihood	-525.08	-528.85	-524.73	-525.16	-529.43	-524.94
N	184,523	184,523	184,523	184,523	184,523	184,523
Pseudo-R^2	0.031	0.024	0.032	0.031	0.023	0.031

Note: Logit models in columns (1) to (3), and Probit models in (4) to (6). Dependent variable: indicator for early state site in cell. Circumscription and land quality are in standard deviation units.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

state site. The fact that the general rate $\hat{\beta}$ is significantly different from zero provides evidence for the relevance of circumscription; and the significance of $\hat{\gamma}$ confirms the importance of agricultural land quality for the formation of early states, as proposed in much of the early state formation literature. Interestingly, including both of them in columns (3) and (6) suggests that circumscription may actually be more important than land quality; or, in terms of the model, the general yield rate is more important than a home premium. However, the main point here is that, according to these estimates, the difference in the quality of the home land and the land surrounding a particular location may be just as important as the land at the location itself (and in fact, we cannot reject that $\hat{\beta} = \hat{\gamma}$).

It is also useful to recall that since $(a - \tilde{a})\beta + a\gamma = (\beta + \gamma)a - \beta\tilde{a}$, the sum of the coefficients on circumscription and land quality give the gross effect of land quality and $-\beta$ gives the effect of a change in the quality of surrounding land. Interpreting the coefficients in this way speaks for a substantial effect of the quality of surrounding land, almost as large as the effect of the own land.

The magnitudes of the coefficients can also be interpreted in terms of an underlying latent index of state formation. The rarity of state formation is reflected in a very high investment cost estimate $\hat{c}$, relative to the stochastic error distribution. For example, in case of the probit models, the standard deviation of the error is unity, and in the unconstrained model (3) we have that $\hat{c} = 3.51$, setting a very high bar for any cell to feature a state site. The magnitude of $\hat{\beta}$ of around 0.12 suggests that a one standard deviation unit increase in circumscription increases the latent index value by 12% relative to the distribution of the error. Of course, this magnitude will be different in the logit models, given that the standard deviation of the error is assumed to be higher (if ε is logistic$(0, 1)$, then $\mathrm{Var}(\varepsilon) = \pi^2/3$, meaning the standard deviation of the error is about 1.81).

These magnitudes translate into economically relevant changes in the probability of a state site in a given cell. As shown in the model summary section of the table, starting from a base of 0.033%, the marginal effect of circumscription is 0.015%, which represents a 45% increase in the probability of a state. Assuringly, marginal effects are almost the same across probit and logit models.

Having established the basic relationship between state formation, circumscription, and land quality, let us turn to the robustness of this relationship to alternative theories of state formation and early development. To this end, we allow the investment cost into extraction capacity to vary from cell to cell according to $c = -\mathbf{X}_i'\eta$, where $\mathbf{X}_i$ includes a constant and groups of control variables accounting for these alternative theories. First, measures of ecological diversity are introduced, specifically, the standard deviation of land quality of cells within the radius, the ecological diversity index and the ecological polarization index (Fenske, 2014) using potential vegetation. This group of controls addresses the significance of trade across ecological boundaries as proposed by Bates (1987).

Second, we include climatic conditions: the annual mean and standard deviation temperature, and absolute latitude (which is a measure of seasonal daylight fluctuations). This group of controls mainly speaks to the extent to which an early Neolithic transition, driven by a favorable climatic environment, may have given some regions a head start in the formation of a state (Olsson and Hibbs Jr, 2005; Putterman, 2008; Ashraf and Michalopoulos, 2015; Matranga, 2017).

Third, topographic, biogeographic and other environmental variables are included: average slope, ruggedness, freshwater accumulation and fixed effects for potential vegetation. These controls address theories about early development related to the ease of control and extraction by state institutions (Nunn and Puga, 2012; Michalopoulos and Papaioannou, 2014). Finally, including virtual country fixed effects takes into account various regional unobservables that may be important for early state formation. Each virtual country is $5^\circ \times 5^\circ$ large, including up to 400 cells $(5/(1/4)$ squared). Differences in unobserved geographic features, population density, and regional culture may all have contributed to state formation, and virtual country fixed effects helps to account for them.

With these groups of control variables, we can now test the robustness of the results presented in table 2.2. To this end, we run the following linear probability model (LPM):

$$Y_i = (a_i - \tilde{a}_i)\beta + a_i\gamma + \mathbf{X}_i'\eta + u_i \tag{2.3}$$

where u_i is the CEF error (i.e. $u_i \equiv Y_i - E[Y_i|a_i, \tilde{a}_i, \mathbf{X}_i]$). While the LPM no longer allows for a structural interpretation of the parameters in the model (Horrace and Oaxaca, 2006), it approximates the CEF and thereby provides reasonable approximations of the marginal effects (Angrist and Pischke, 2008). Similar estimates using logit can be found in the appendix.

Table 2.3 shows eight increasingly demanding specifications of the association of early state sites with circumscription and land quality. In addition to introducing groups of control variables, these estimates also show three different types of standard errors for each point estimate: robust standard errors in parentheses, as in table 2.2; standard errors clustered on the virtual country level, forming 712 clusters, in brackets; and standard errors taking into account the spatial correlation of the error structure (Conley, 1999) in curly brackets.

The first three models (1)-(3) replicate the results from table 2.2 in an LPM framework and find similar results. Again we first see both circumscription and land quality having a significant and positive effect on the probability of a state site. Due to the LPM, marginal effects are about 50% higher in model (3), meaning that controlling for land quality, a one standard deviation increase in circumscription leads to an increase in state site probability of around 70%.

The introduction of ecological diversity controls in model (4) substantially changes these conclusions: first, we see that the coefficient on circumscription more than doubles to 0.052 percentage points, implying a one standard deviation increase in circumscription being associated with a 157% increase in the probability of finding an early state site in the cell. Interestingly, we now see the coefficient on land quality turning significantly negative, with a fairly large magnitude. The implication is that, controlling for circumscription, a one standard deviation increase in land quality actually leads to a decrease in the probability of a state site of around 109%.

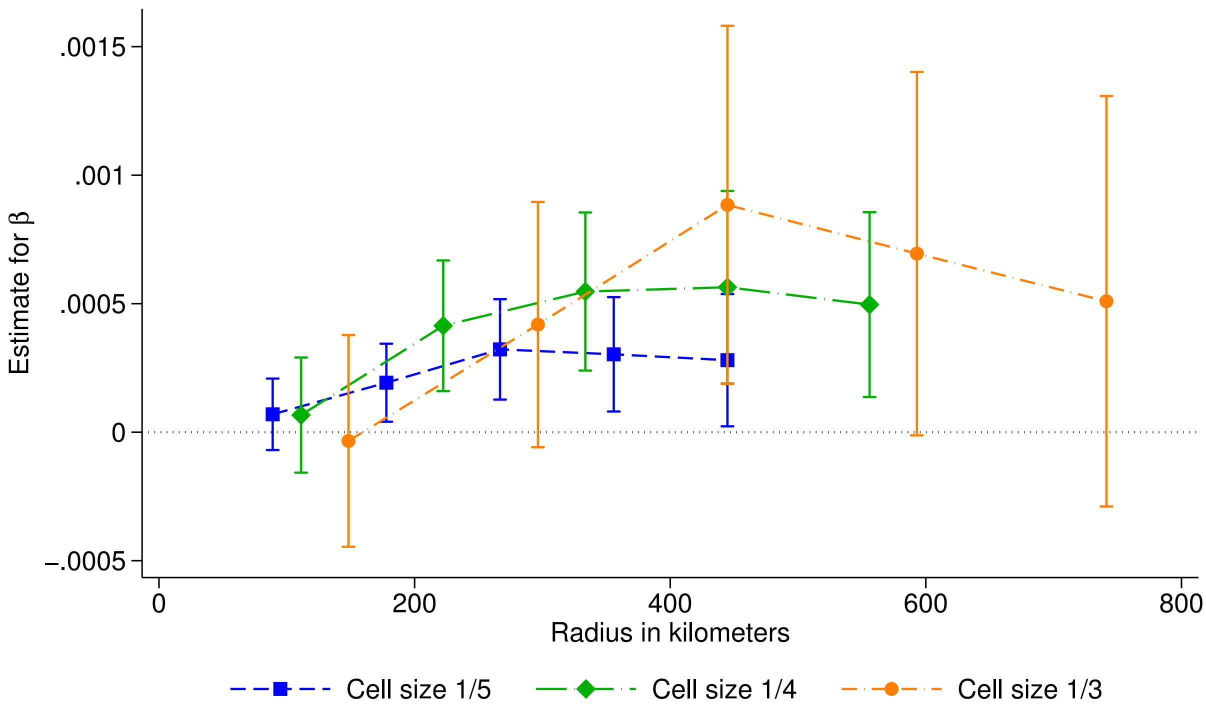
Given the negative point estimate of $\hat{\gamma}$, it is useful to again recall that $(a - \tilde{a})\beta + a\gamma = (\beta + \gamma)a - \beta\tilde{a}$. That is, we can also interpret the coefficients in terms of a gross land quality effect and a surrounding-land effect. The estimates in model (4) suggest that a one standard deviation increase in gross land quality ($\hat{\beta} + \hat{\gamma}$) increases the probability of a state site by about 0.016 percentage point, or 48%, and an increase in the quality of the surrounding land by one standard deviation *decreases*

Table 2.3: Robustness estimates using OLS.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Circumscription: $\hat{\beta}$	0.00028		0.00023	0.00052	0.00051	0.00051	0.00047	0.00054
	(0.00007)		(0.00009)	(0.00010)	(0.00010)	(0.00010)	(0.00010)	(0.00014)
	[0.00009]		[0.00011]	[0.00015]	[0.00015]	[0.00015]	[0.00015]	[0.00021]
	{0.00011}		{0.00013}	{0.00017}	{0.00017}	{0.00017}	{0.00016}	{0.00021}
Land quality: $\hat{\gamma}$		0.00025	0.00008	-0.00036	-0.00037	-0.00035	-0.00027	-0.00038
		(0.00005)	(0.00005)	(0.00007)	(0.00008)	(0.00008)	(0.00008)	(0.00014)
		[0.00008]	[0.00009]	[0.00012]	[0.00013]	[0.00013]	[0.00012]	[0.00020]
		{0.00009}	{0.00009}	{0.00012}	{0.00013}	{0.00013}	{0.00012}	{0.00019}
Mean dep. variable	0.00033	0.00033	0.00033	0.00033	0.00033	0.00033	0.00033	0.00033
Ecol. div. controls				X	X	X	X	X
Climate controls					X	X	X	X
Topog. & env. controls						X	X	X
Pot. vegetation FEs							X	X
Virtual country FEs								X
N	184,523	184,523	184,523	184,523	184,523	177,299	177,299	177,289
Adjusted R^2	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.007

Note: Linear probability models (via OLS). Dependent variable: indicator for early state site in cell. Circumscription and land quality are in standard deviation units. Robust standard errors in parentheses; standard errors clustered by $5° \times 5°$ virtual countries (712 clusters) in brackets; spatially robust (Conley) standard errors with radius 278km (standard cell size of 1/4 (28km at equator) times standard radius $r = 10$ in cell units) in curly brackets. Significance stars omitted. Ecological diversity controls includes the standard deviation of land quality, the ecological diversity index and the ecological polarization index; climate controls includes annual mean and standard deviation temperature and absolute latitude; topographic and environmental controls includes slope, ruggedness and river flow accumulation.

Figure 2.4: Coefficients for various cell sizes and radii.



Note: Estimates for β from equation (2.3) and specification (8) from table 2.3 for various radii and cell sizes. Cell sizes are in degrees; radii are $r = 4, 8, 12, 16, 20$.

the probability by −109%. This interpretation implies the quality of surrounding land is more important than the quality of a land in a given place.

Moving on to models (5)-(8), estimates are largely stable and lend themselves to a similar interpretation as model (4). This suggests that ecological diversity and the trade theory of state formation typically associated with it has merit in the data, but is largely complementary to circumscription. Adding further explanatory variables from other theories hardly moves the needle. Most estimates across all models are significant at a one percent level using all three types of standard errors, except land quality in models (3), (7) and (8) (significant at five and ten percent level, respectively) and circumscription in model (8) (significant at five percent level).

So far, all results presented were using a grid with cell size ¼ and circumscription radius of $r = 10$. Figure 2.4 presents results varying the cell size between ⅕ and ⅓ degrees as well as radii between 4 and 20 cell units. Each point in the graph represents an estimate from a separate regression, with corresponding confidence intervals extending vertically. We use the most demanding specification from table 2.3 (model (8), with robust standard errors). Since this is a very demanding specification, we also show results for model (7) from table 2.3 in the appendix.

Looking across radii, we see that the point estimates show an inverted-U shape as the radius increases: coefficients are small and insignificant at a small radius, then they rise at intermediate ranges, and finally they drop off slightly. This suggests there may be a distance at which circumscription is most relevant, for a given cell size. At ¼, the "Goldilocks" zone for circumscription seems to be at a radius of around 400km, showing the strongest association with early state formation.

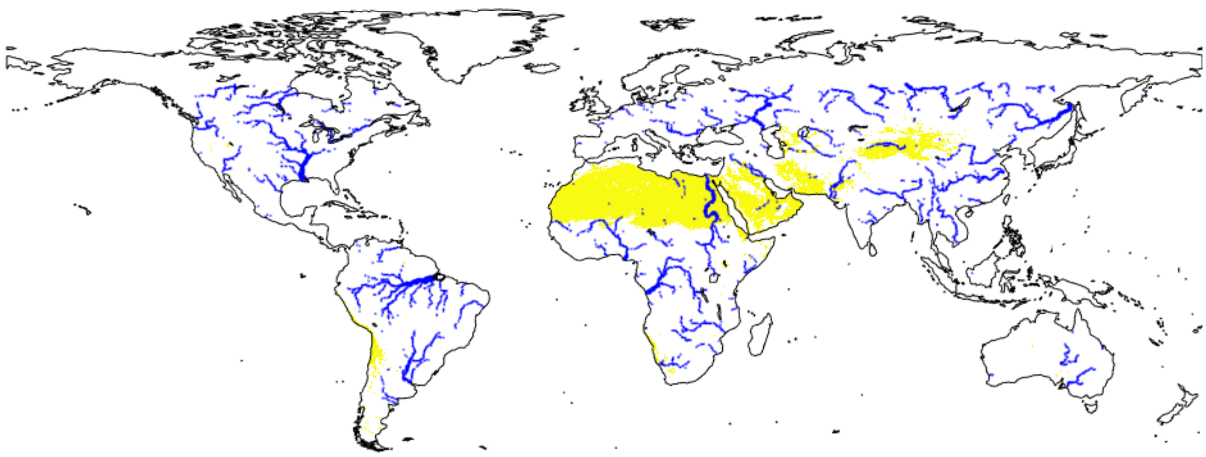
Comparing across cell sizes, we can see that there is a level effect as cell size decreases from ⅓ to ⅕. This is to be expected due to the fact that the mean of the dependent variable falls with with cell size as well (since a smaller share of cells has any state site in them at smaller cell sizes). In fact, looking at the marginal effect of state creation for a one standard deviation unit increase in circumscription on $r = 10$, it is 119% for ⅓, 166% for ¼ and 144% for ⅕. We can also see that the arcs formed across larger radii are shorter for smaller cell sizes. This implies that the relevant magnitude is the number of cell units for a given cell size. In this sense, circumscription seems most relevant within a radius of about twelve cell size units from a given location.

2.5 Instrumental Variable Estimation

So far, we assumed that the degree of circumscription is exogenous with respect to the process of state formation. This assumption relies on the fact that many features determining the distribution of land quality are immutable: daylight time and topographic features are unaffected by human activity; and things like regional climate, average cloud cover, potential vegetation or freshwater accumulation change only slowly over time. However, there remain two sources of bias that may affect a causal interpretation of the estimation. First, the land quality measure may be subject to measurement error, for example due to assuming domesticated crops on either side of the Atlantic were available in all places on their respective continents. Second, it is possible that early states induced a very particular form of land degradation that may have produced a pattern of circumscription: peripheral erosion and preservation at the center. Although it is unlikely that degradation would be systematically higher outside the core region of a state than inside, this type of reverse causation would lead to biased estimates of circumscription.

To deal with these potential concerns, we use an instrumental variable strategy.

Figure 2.5: Map of large rivers and deserts.



Note: Areas with potential vegetation being deserts in yellow; the top one percent of river flow accumulation (rivers) in blue, with river flow accumulation scaled to magnitude.

Specifically, we instrument circumscription with the interaction of two geographic features that, together, are an important source of steep land quality gradients but are unlikely to directly affect the cost of state formation: large rivers intersecting arid land. We employ this instrument while controlling for the direct effects of rivers and potential vegetation. This means any threat to identification has to rely on an explanation of how the simultaneous presence of rivers and deserts affects state formation other than through circumscription; the identification strategy controls for any explanation involving rivers or deserts in isolation.

Figure 2.5 shows a map of the top one percent of river flow accumulation and areas whose potential vegetation has been classified as desert by SAGE. This highlights that particularly four regions would exhibit circumscription driven by the intersection of these two geographic phenomena: Egypt, Mesopotamia, the Indus Valley and Western China.

The instrumental variable strategy is formalized in the following first stage regression equation:

$$a_i - \tilde{a}_i = Z_i\psi + a_i\xi + \mathbf{X}_i'\delta + e_i \tag{2.4}$$

where $Z_i = \log(\text{RiverAccum}_i) \times \mathbf{1}[\text{PotVeg}_i = \text{Desert}]$ and e_i is assumed to be an uncorrelated error term. Both $\log(\text{RiverAccum}_i)$ and $\mathbf{1}[\text{PotVeg}_i = \text{Desert}]$ are always included in the set of control variables $\mathbf{X}_i$. Similar to table 2.3, we introduce consec-

Table 2.4: Instrumenting using desert rivers.

Panel A:	First stage			Reduced form		
	(1)	(2)	(3)	(4)	(5)	(6)
RiverAccum × Desert	0.0161***	0.0160***	0.0046***	0.0009***	0.0009***	0.0010***
	(0.0012)	(0.0012)	(0.0005)	(0.0003)	(0.0003)	(0.0003)
Control variables	X	X	X	X	X	X
Pot. Vegetation FEs		X	X		X	X
Virtual Country FEs			X			X
N	177,299	177,299	177,289	177,299	177,299	177,289
R^2	0.764	0.771	0.952	0.001	0.002	0.012
Panel B:	IV estimates					
	(1)	(2)	(3)	(4)	(5)	(6)
Circumscription	0.059***	0.057***	0.056***	0.055***	0.056***	0.210***
	(0.022)	(0.020)	(0.019)	(0.019)	(0.020)	(0.073)
Ecol. div. controls		X	X	X	X	X
Climate controls			X	X	X	X
Topo. & env. controls				X	X	X
Pot. vegetation FEs					X	X
Virtual country FEs						X
Partial F of instr.	53.2	162.7	175.1	176.5	180.0	93.0
N	177,299	177,299	177,299	177,299	177,299	177,289
Within R^2	0.008	0.095	0.110	0.112	0.114	0.001

Note: IV/2SLS estimates. Panel A: first stage (1)-(3) and reduced form (4)-(6). Panel B: IV estimates. In the first stage, the dependent variable is circumscription; in the reduced form and the IV estimates, it is an indicator for whether the cell has a state site. All regressions include land quality, river flow accumulation and a desert indicator as controls. Kleinbergen-Paap rk statistic (2006) as Partial F for instrument.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

utively larger sets of additional control variables controlling for alternative theories of state formation and early development. Equation (2.3) forms the second stage.

Results are presented in table 2.4. Panel A shows the reduced form and the first stage of the regression. Across all specifications, we see very stable coefficients, except when virtual country fixed effects are introduced. The first stage coefficient on the instrument is highly significant, yielding partial F statistics substantially above common thresholds for weak instruments. In columns (1) and (2), since river flow accumulation is in logs, the coefficient is interpreted as a ten percent increase in river flow accumulation in deserts leads to an increase of circumscription of 0.161

standard deviation units. This coefficient drops by about a factor of three after controlling for virtual country fixed effects: now that we are only comparing within virtual countries, the fact that only a minority of virtual countries have desert rivers sucks up a lot of the variation. However, the coefficient is still highly significant.

In the reduced form models, a one percent increase in desert river flow accumulation has a large (considering the small base) effect of around 0.1% on the probability of a state site in a cell in all specifications. This is around twice as large as the effect of a standard deviation increase in circumscription reported in table 2.3, confirming qualitative assessments about the importance of desert rivers such as the Nile in accelerating the formation of early states (R. C. Allen, 1997). Interestingly, the coefficient is virtually unaffected by the introduction of virtual country fixed effects, suggesting that the result is just as strong comparing cells only within the virtual countries that have state sites.

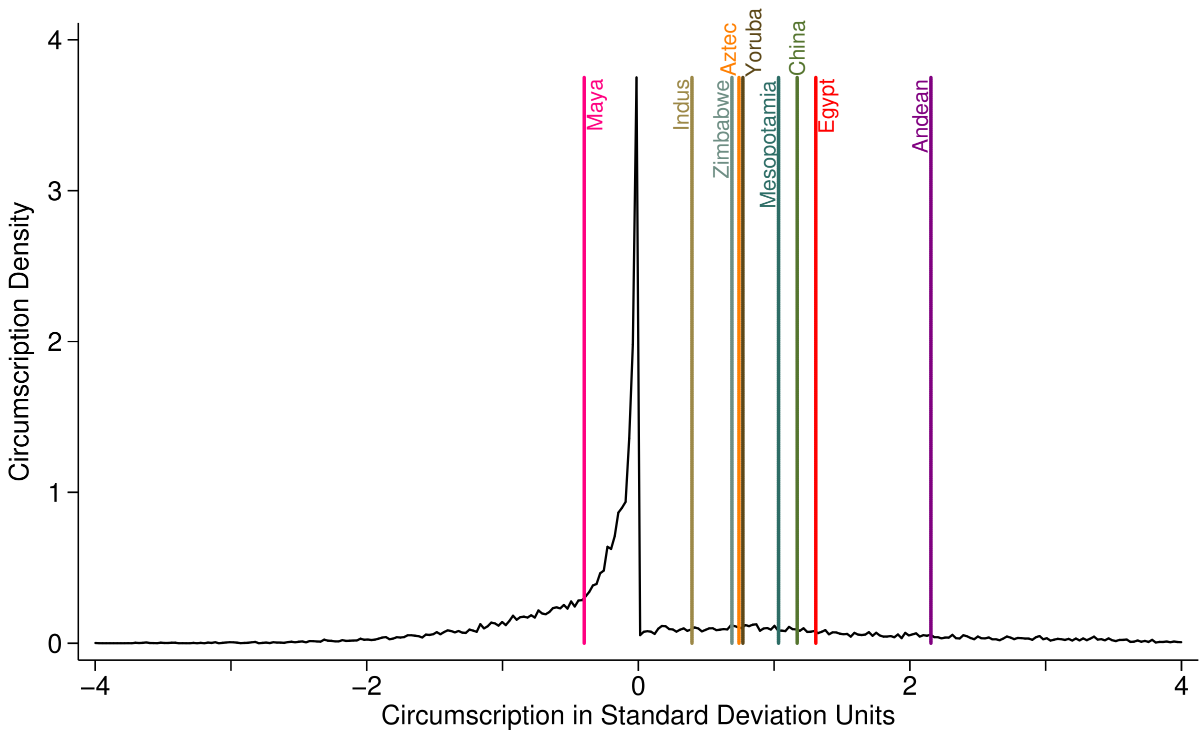
Panel B shows the instrumental variable estimates. We find large, precise and stable coefficients across all specifications except when introducing virtual country fixed effects in model (6), which, while maintaining precision, lead to a four-fold increase in the coefficient. The magnitude in models (1)-(5) implies a one standard deviation increase in circumscription leads to an increase in probability of finding a state site in a cell of around 5.7%. Off of the minuscule base of 0.03%, this is obviously an enormous effect that needs to be interpreted with caution. However, the stability and precision of the results speaks for the possibility of substantial attenuation bias due to measurement error in the OLS specifications, and for the importance of circumscription as a driver of early state formation.

2.6 The Intensity of Extraction across Civilizations

The extent to which early state formation was driven by circumscription may differ from one civilization to the next. To assess heterogeneity in a straightforward way, we simply plot average circumscription by civilization in figure 2.6. The skew in the distribution of circumscription is due to the fact that most cells have at least one cell in their vicinity with positive land quality, even if their own land quality is zero.

We see that circumscription indeed differs substantially from one civilization to the next, in ways that are largely consistent with the qualitative literature on circumscription: the original examination by Carneiro (1970) begins with contrasting the extent of circumscription of agricultural groups in the Andes with those in the Amazon basin. The evidence in figure 2.6 confirms his assessment of the Andes being a strongly circumscribed region, with average circumscription of Andean state sites being more than two standard deviations higher than the mean. Egypt is the

Figure 2.6: Density of circumscription and circumscription per civilization.

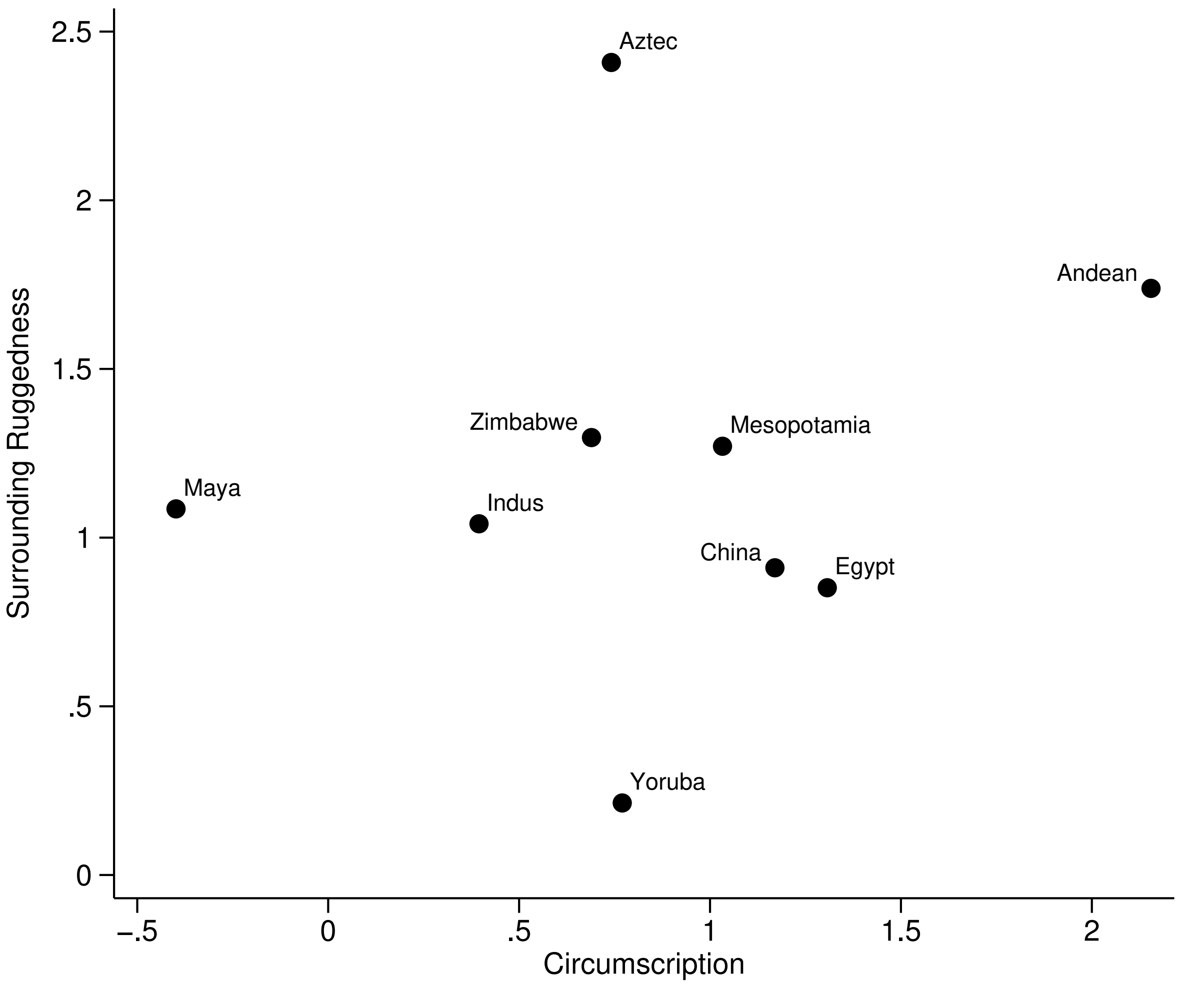


Note: Estimated kernel density function of circumscription in black; average circumscription per civilization as vertical lines in various colors. The skew in the distribution of circumscription is due to the fact that most cells have at least one cell in their vicinity with positive land quality, even if their own land quality is zero.

other classic example frequently employed (R. C. Allen, 1997), and it also shows up as heavily circumscribed. On the other hand, the formation of states in China is much less closely associated with circumscription, although the data here speaks for circumscription being an important source of state formation in China as well.

Not all civilizations have circumscribed early state sites. At the bottom of the list rank the Maya, whose sites are actually on average negatively circumscribed, meaning they are on average surrounded by better land than their own. The fact that the Maya do not adhere to the pattern of circumscription is consistent with the detailed archaeological assessment by Trigger (2003), who concludes that the Mayan region is not circumscribed. McAnany (2014) finds that the Maya could have expanded further South into Central America, offering an escape route to state evaders. The Indus Valley civilization, while about 0.4 standard deviations above the mean, also ranks low in terms of circumscription. This stands contrary to the qualitative archaeological assessment. However, the aridification of the Indus region due to the weakening of the Monsoon is well documented in the archaeological and paleoclimatic record (Giosan et al., 2012), which may explain why it looks much less circumscribed today.

Figure 2.7: Scatter plot of ruggedness against circumscription by civilization.



Note: Both variables are in standard deviation units.

One possible source of this heterogeneity across civilizations may arise from the extent to which other drivers of early state formation were present. The strategic environment for outside groups to stage a challenge to incumbents is a particularly relevant alternative, as it directly connects to the model components of conflict, displacement and domination. The ruggedness of the surrounding terrain has been found to present a favorable strategic environment for conflict in studies of insurgency (Fearon and Laitin, 2003) and slave capture (Nunn and Puga, 2012). Figure 2.7 plots average circumscription and average surrounding ruggedness by civilization. We interpret higher circumscription relative to surrounding ruggedness primarily as evidence for extraction, while the reverse points towards other motives such as protection.

It is interesting to note that the three civilizations that eventually formed the first territorial early states – China, Egypt and the Andes – all show relatively higher circumscription, while civilizations that maintained city-state systems for a long time, such as the Aztec, the Maya and Mesopotamia, are relatively more surrounded by rugged terrain. Also noteworthy is that the three civilizations that brought about early territorial states were also arguably the most extractive, with substantial agricultural shares and corvée labor collected by the state (Finer, 1997).

The Andean civilization stands testament to the massive extent of the extractive state that was possible in areas that were both circumscribed and surrounded by rugged terrain. The conquistador Pedro Cieza de León noted that the population in today's Peru was easy to control because of the lack of refuge for dissidents (Salomon, 1986). The extractive institution of the *Mita* maintained by the Spanish in Peru until 1812 was a remnant of the Inca state designed as a means to extract agricultural labor (see Dell (2010) and references therein). This example illustrates the long-lasting impact of the geographic environment on economic development through the extractive institutions of early states.

2.7 Conclusion

This paper shows how the lack of outside options for mobile groups enhances the capacity for extraction through states. Based on this argument, it then provides evidence that the quality of land surrounding an area is an important cause of state formation, as important or even more important than the quality of land of the area itself. Thereby, it introduces the notion of circumscription as a driving force of political economic dynamics.

The strong relationship of archaeological excavation sites related to state formation with circumscription masks substantial heterogeneity across civilizations. We explore the possibility of ruggedness enhancing the role of circumscription by offering strategic opportunities for challengers: as more conflicts take place due to the favorable strategic environment offered by mountains, more trials of state creation are undergone, leading to the formation of more early states. Another possible explanation for the heterogeneity is that the migratory capacity of groups may have depended on other factors besides the configuration of land quality. For example, J. M. Diamond (1998) suggests that it is harder to transport crops and agricultural techniques along a meridian (e.g. from North to South) than along circles of latitude (e.g. East to West) due to the relatively higher ecological similarity along the East-West axis. This would make an escape out of the narrow, ecologically diverse isthmus connecting the Americas harder than other in other environments.

It is also noteworthy that circumscription may be a necessary condition for early state formation, but it is not sufficient: there are several regions across the globe that show substantial circumscription but apparently never brought about early states. This is particularly true for coastal regions of Argentina and Uruguay. While there is a lively about the role of population density and more complex societies (Powell et al., 2009; Bettinger, 2016), one explanation for the absence of states in this region is that there was simply not enough time for human societies to grow large enough in this region for the dynamics of circumscription to kick in. This possibility underlines the issue that circumscription may favor the development of different societies at different scales of circumscribed areas, which may be depend on the overall size of human societies in a given region.

The debate in economics about the origins of states is centered on the theme of extraction, with some research on the role of protection. These are the classic motives identified in political science and economics and to which this paper contributes. However, a more comprehensive organization may include the distinction from evolutionary archaeology into voluntaristic and coercive theories. The former focus on the state's ability to generate a surplus, while the latter focus on the state's ability to extract resources from its subjects. Putting together both traditions yields a conceptual triad of protection, extraction and surplus production that may prove useful in future research.

More generally, this paper adds to the growing economics literature highlighting the role of spatial proximity for social phenomena in a given place. This notion has been fruitfully exploited in such diverse fields as international trade (T. Allen and Arkolakis, 2014) or real estate and crime Linden and Rockoff (2008). In this sense, we contribute to the intellectual development of what has been called geography's "second law": "the phenomenon external to an area of interest affects what goes on in the inside" (Tobler, 1999).

Finally, the results presented here also point towards the dearth of globally consistent archaeological data currently available to researchers to study the rise of complex societies, and state societies in particular. Bogucki's collection of sites satisfies some basic conditions of a global, geocoded dataset of archaeological sites relevant to state creation, but larger and more ambitious data collection and standardization exercises may bring about much more precise insights into the mechanics of early state formation. In this sense, efforts such as Wright's Atlas of Chiefdoms and Early States Wright (2006) promise to open up new approaches to study this question through the collaboration of economists and archaeologists.

Chapter 3

State Power and Urban Growth: Evidence from the Universe of Boundary Changes in Europe 1000-1850

3.1 Introduction

European states were formidable before European economies. Numerous conquests across the globe – either directly led by European state forces or sanctioned and supported by them – stand testament to the reach of the European state (J. M. Diamond, 1998). At home as well, European states proved increasingly capable, extracting larger shares of the economy and investing more resources into defending and expanding their territorial possessions (Hoffman, 2015). Meanwhile, while there are early signs of a "Great Divergence" in the 16th century, European economies and urbanization rates were still broadly comparable with those in Asia (Broadberry and Gupta, 2006; Fouquet and Broadberry, 2015).

If European state power preceded European economic power, it is natural to wonder whether the increasingly powerful states in Europe directly contributed to economic development on the continent. The increasingly stable rule and the ability to defend and conquer territory may have provided an environment more conducive to economic development than in places where states were weak and competitive theft of roving bandits was rampant (Olson, 1993).

In this paper, we document the importance of stable state rule for economic development. To this end, we use newly available data on the universe of *de facto* boundary changes between all European states over almost one thousand years starting in

AD 1000, assembled in the *Centennia Historical Atlas* (Reed, 2016). We combine these boundary changes with data on urban growth from 2,181 cities across Europe (Bairoch et al., 1988). While we refer to our outcome of interest simply as city population, this is typically used as a proxy for economic growth more broadly (Acemoglu et al., 2005).

Our boundary data from the *Centennia Historical Atlas* allows us to study the evolution of European states in new detail. Specifically, we can identify all changes in *de facto* territorial possessions of every significant European state since the Middle Ages on nearly a monthly basis across more than 10,000 time periods.

We begin by documenting several new facts about state power and urban growth. First, in each century since AD 1000, 47-71% of cities switched at least once from one state to another, affecting similar shares of urban population. Second, the average number of switches per city in a given decade was around 10% up to the 17$^{\text{th}}$ century, after which it skyrocketed to more than 50% – more than 1,000 city switches occurred between 1750-1800 – before falling back to around 40% in the early 19$^{\text{th}}$ century. Overall, 82% of cities switched states at least once since their first appearance in the data, and on average they switch 9.29 times. Third, we document substantial high-frequency fluctuations in the number and the set of existing states. Of the 192 significant European states (i.e. with at least one city) in the data, the average lifespan is only 156 years.

To move towards a causal estimate of the impact of protective state capacity on urban growth, we estimate the effect of switching from one state to another on city population, exploiting the dynamic nature of the data to simultaneously control for both unobserved fixed city characteristics as well as either unobserved fixed state characteristics or unobserved time-varying state characteristics. In this way, we can account for both geographic drivers of urban development and factors related to institutional quality.

We establish two key results. First, we show that switching from one state to another has a strong negative effect on city population of around 12%. This result is robust to various sets of additional fixed effects, including period fixed effects, state-by-period fixed effects, and country-by-period fixed effects. Second, we show that there are negative spillovers from nearby cities switching to another state of around 6%. This seems to be at least partially compensated by positive spillovers of allied cities subject to switches.

Using these estimates, we then simulate counterfactual urban population growth in Europe over the period 1000-1850. We find that in the absence of the negative spillovers, European city population in 1850 would have been about 4.2% higher. In the absence of both direct switching and spillover effects, Europe's urban population would have been about 9.2% higher.

We contribute to the literatures on the rise of Europe, the causes of long-run economic growth, and the role of state capacity in development. Broadly, the rise of Europe has been attested to at least four categories of causes: culture (Clark, 2008), geography (J. M. Diamond, 1998; Nunn and Qian, 2011), technology (McNeill, 1982), and institutions (Acemoglu et al., 2005). Our explanation is institutional as well, but instead of focusing on the inclusivity of institutions, we advance the complementary hypothesis that the ability to secure and hold territory is conducive to development.[1]

Military competition is often seen as beneficial for European state development (Tilly, 1990; Voigtländer and Voth, 2013; Hoffman, 2015). Our work highlights that this may have come at a substantial price in terms of economic development, *ceteris paribus*.

A recent literature studying the long-run impact of state capacity on economic development finds mixed results on how powerful state institutions may affect economic outcomes. Dell (2010) documents persistent negative effects of pre-colonial state institutions on outcomes in Peru. Michalopoulos and Papaioannou (2014) find no effect of externally imposed boundaries on economic development of, but they find positive effects for pre-colonial states (Michalopoulos and Papaioannou, 2013). Bockstette et al. (2002) document how the age of the state correlates with modern economic development. Dell et al. (2017) show how a strong, centralized historical state contributes to higher living standards. The dynamic nature of state borders over the long-run in our data uniquely allows us to distinguish geography from states, identifying each of their roles separately.

Finally, our work connects to the literature on state capacity and weak states (Mann, 1984; Tilly, 1990; Besley and Persson, 2010; Gennaioli and Voth, 2015). Protection against external threats is a key driver of the development of state capacity and lays the foundation of economic development. We provide evidence for this link from security to economic growth. Relatedly, Dal Bó et al. (2016) show theoretically how states can achieve both security and prosperity, a link that we establish empirically for Europe over the second millennium.

The rest of this paper is organized as follows. We briefly discuss the underlying logic and mechanisms relating state power and urban growth in section 3.2. In section 3.3, we describe the data in detail. We then set up the empirical design in section 3.4, with results presented in section 3.5. Section 3.6 concludes.

[1]Noteworthy here is also the case of modern China. Chinese economic growth over the last thirty years made up a significant fraction of global economic development. This happened under a state that was hardly inclusive, but rather provided stability and control.

3.2 Conceptual Framework

As briefly discussed during the literature section in the introduction, scholars have identified a number of ways in which the presence of strong states may contribute to economic development. State power may be good for urban growth because it provided stable rule, minimized uncertainty and guaranteed internal peace. Both public and private actors may conduct themselves in a way that is more in line with economic development: investments into intra- and inter-city infrastructure through taxation is more likely when states build capacity (Besley and Persson, 2010); and entrepreneurs may have more confidence in business investment if they expect a city to adhere to a particular set of rules guaranteed by the state (Knack and Keefer, 1995).

In addition to these upsides of state power, a strong state may also prevent a number of downsides that go hand in hand with its absence. Frequent switches in state rule suggest an environment as described by Olson (1993): the tragedy of the commons in the form of over-extracting roving bandits. In contrast, a strong state has an interest in providing public goods to its domain. The provision of security against external threats also has a direct effect of preventing physical destruction from warfare in the city.

Establishing that the lack of stable state rule leads to lower city population raises the question whether the population was diminished or dispersed in the countryside, or whether it was displaced into nearby cities. It may also be the case that unstable rule in nearby cities has a negative effect on a given city due to the benefits from regional agglomeration and frictionless trade within clusters of cities. These may be disrupted by new rulers setting up tariffs or other barriers to trade and cross-city interaction. Additionally, negative spillovers from nearby cities switching states may also be a sign of broader regional warfare affecting all cities in a region.

Of course, the strength or weakness of a state is relative. We interpret the loss of a city as a sign of relative weakness; but it is likely that in an environment with many weak states, many more opportunities for territorial changes show up than in one with strong states. In this sense, the infrequency of border changes in the developed world is a consequence of powerful states who are able to protect their territory from other state actors.

3.3 Data and Descriptive Statistics

To understand the role of the state in urban growth in Europe in the second millennium, we combine two data sources: the *Centennia Historical Atlas* (henceforth:

Atlas) by Reed (2016) and city population data by Bairoch et al. (1988). The former dataset is newly available and marks a significant improvement in the quality of European state boundary data, as detailed below. The latter dataset has been widely used and recently updated in Voigtländer and Voth (2013), with the only addition from our side being a more precise geocoding of all cities, which we illustrate in a series of maps.

The Atlas consists of boundary data covering Europe, Western Asia, the Middle East and North Africa, although most changes take place in Europe. Instead of a single static map, the Atlas covers ten sets of boundaries each year for every year between AD 1000 and AD 2003, resulting in a total of 10,030 periods (that is, a period is 5.2 weeks).

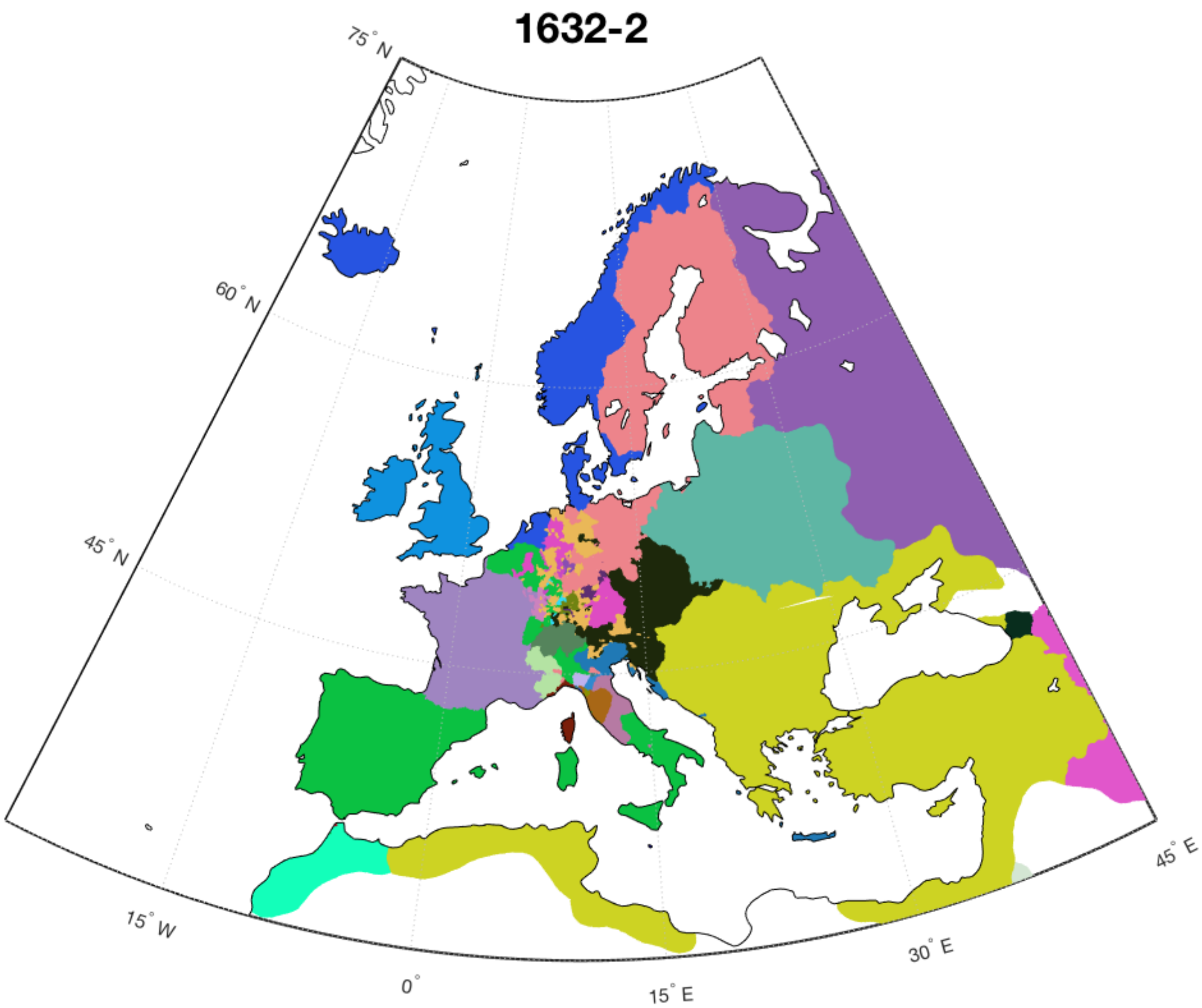
The data has been collected and processed over several years by Clockwork Mapping, a small cartography business with a focus on historical mapping. It is the basis of the view-only dynamic map-based guide of the history of Europe and the Middle East, which is available at `http://www.clockwk.com/`. The goal of the Atlas is to depict *de facto* territorial control by states, as opposed to claims to territory. Most of the work consisted of assembling various historical maps according to the consensus of the historical cartography community, with some discretion if the consensus reflected claims rather than actual territorial control. In case there was no consensus, an attempt was made to provide a consistent judgement across regions and time with the goal of depicting "boots-on-the-ground" power.[2]

To ensure the data was of high quality, we consulted several historians with knowledge of various periods of European history and tested the boundaries shown in the Atlas against their historical timeline of individual places, with special attention to the accuracy of the state in control and the timing of boundary changes. In the two test regions, the Low Countries and the central Holy Roman Empire, territorial control and timing of changes were consistent with the historical narrative. Of course, while the data may accurately depict which state had de facto power over a given territory, some states had a much stronger grip on the territory they claimed sovereignty over than others, and some boundaries were more porous than others. Additionally, European states also took on increasingly large territorial possessions overseas, which we cannot observe.

Figure 3.1-3.3 show examples of the high-frequency data available in the Atlas. In figure 3.1, we see the active states during the second period of the year 1632. At the height of the Thirty Years' War, King Gustavus of Sweden (purple) invaded Pomerania to retake the lands lost to the Catholic armies of the Austrian Habsburgs

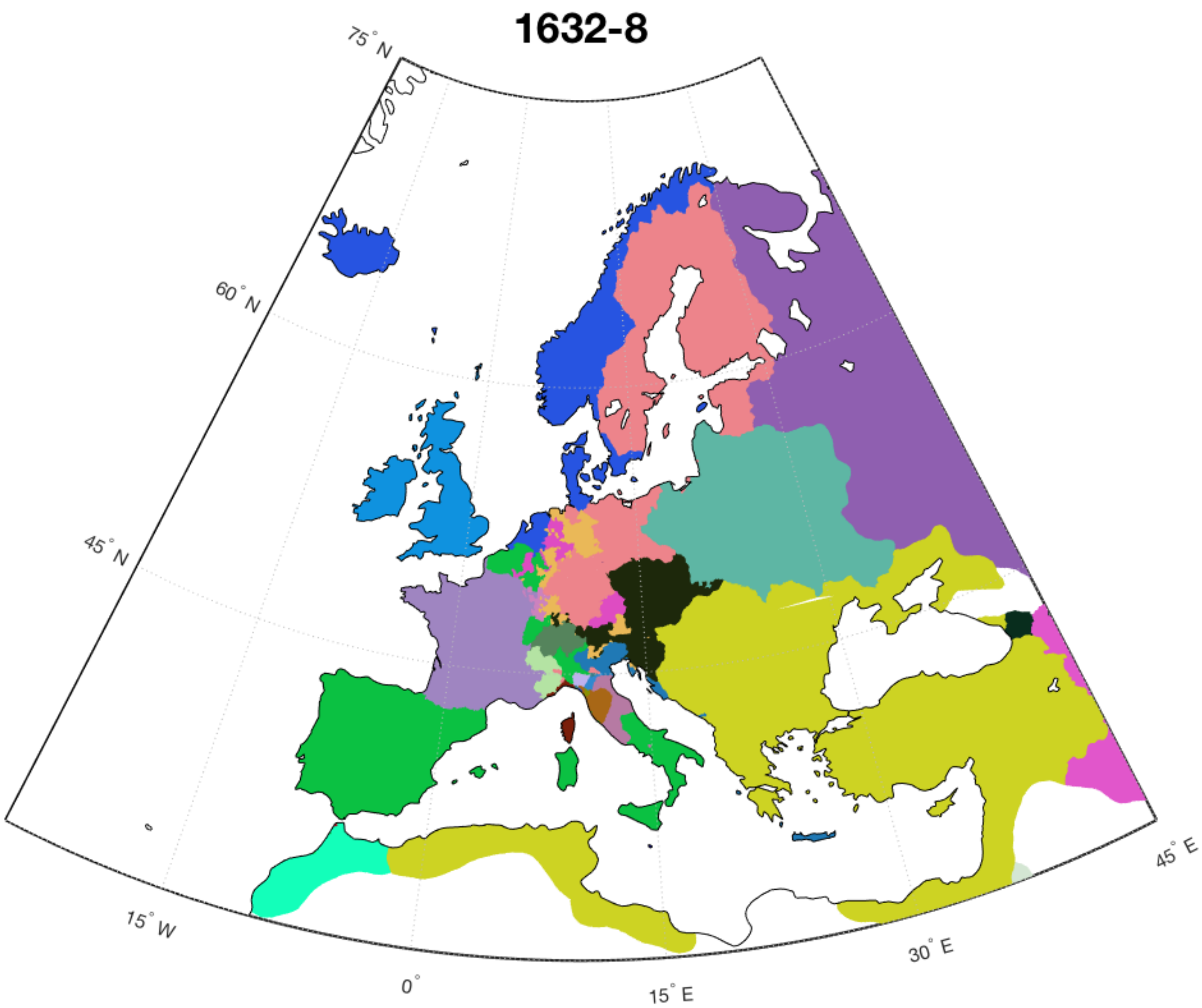
[2]This information is based on personal communication with the head cartographer of Clockwork Mapping, Frank Reed.

Figure 3.1: Example of the *Centennia Historical Atlas*: year 1632, period 2



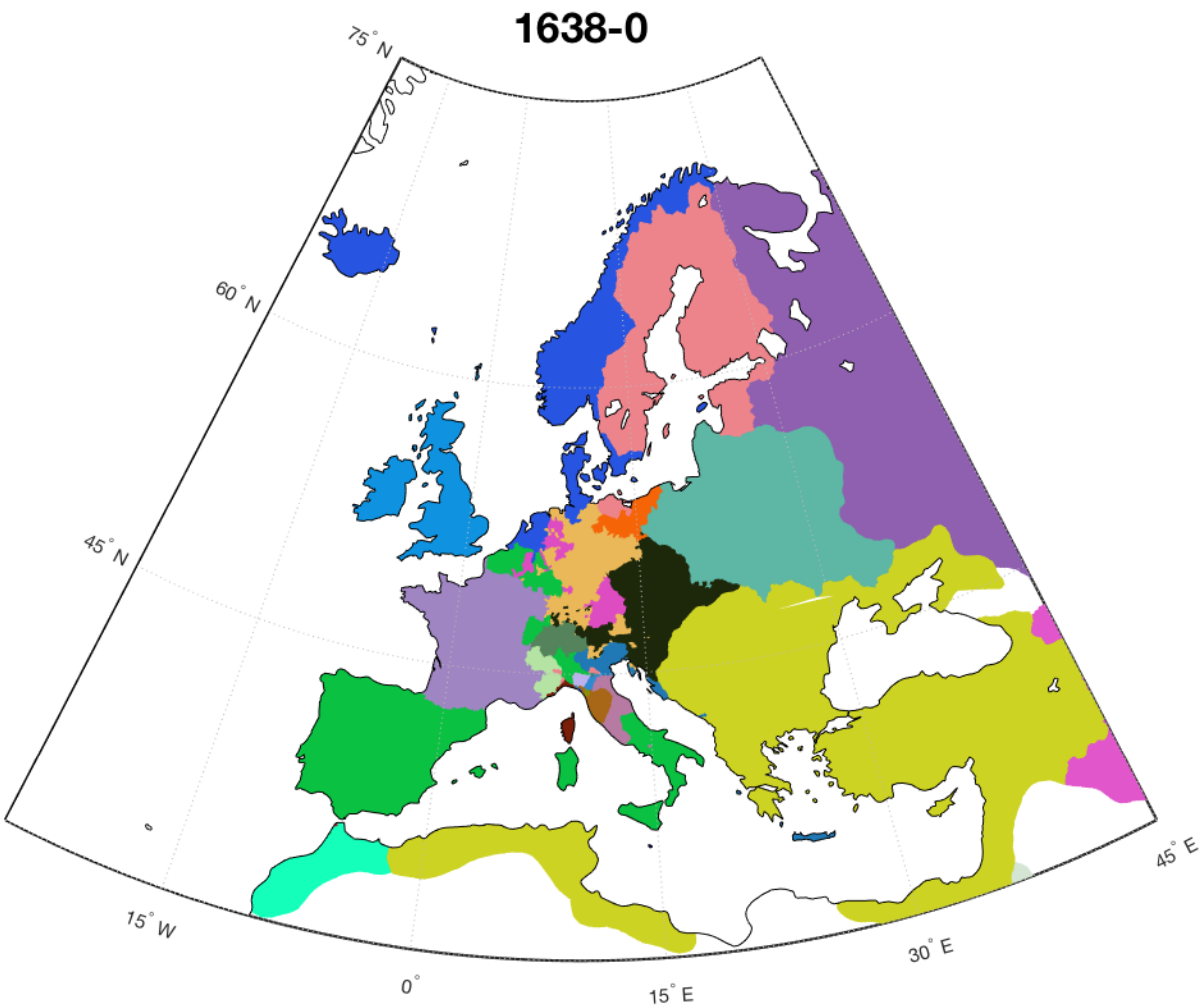
Note: Territorial control of european states in early 1632, period 2 out of $\{0, 1, ..., 9\}$.

Figure 3.2: Example of the *Centennia Historical Atlas*: year 1632, period 8



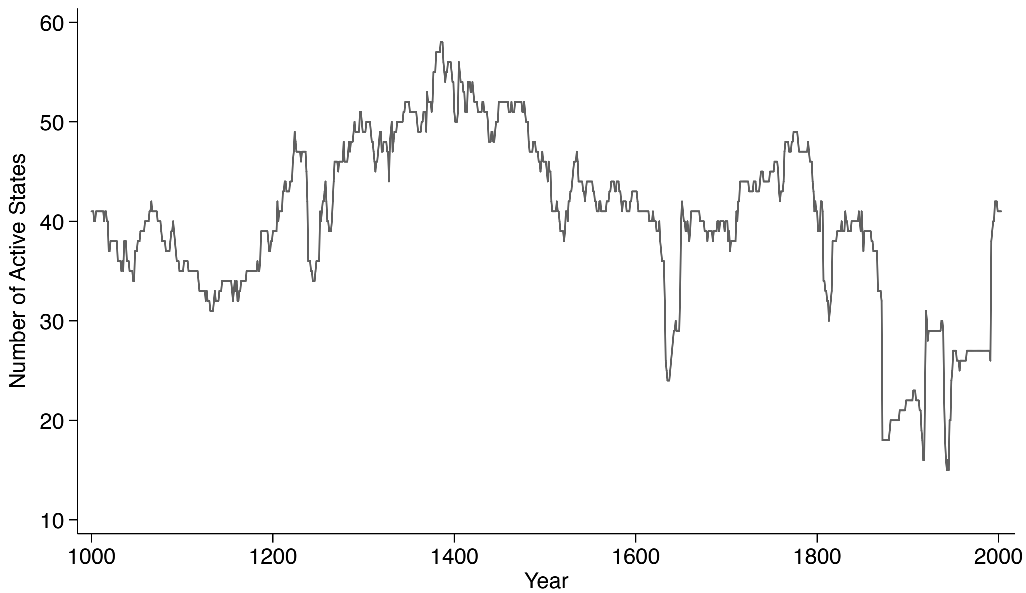
Note: Territorial control of european states in late 1632, period 8 out of $\{0, 1, ..., 9\}$.

Figure 3.3: Example of the *Centennia Historical Atlas*: year 1638, period 0



Note: Territorial control of european states in late 1637, period 0 out of $\{0, 1, ..., 9\}$.

Figure 3.4: Number of states with cities over time



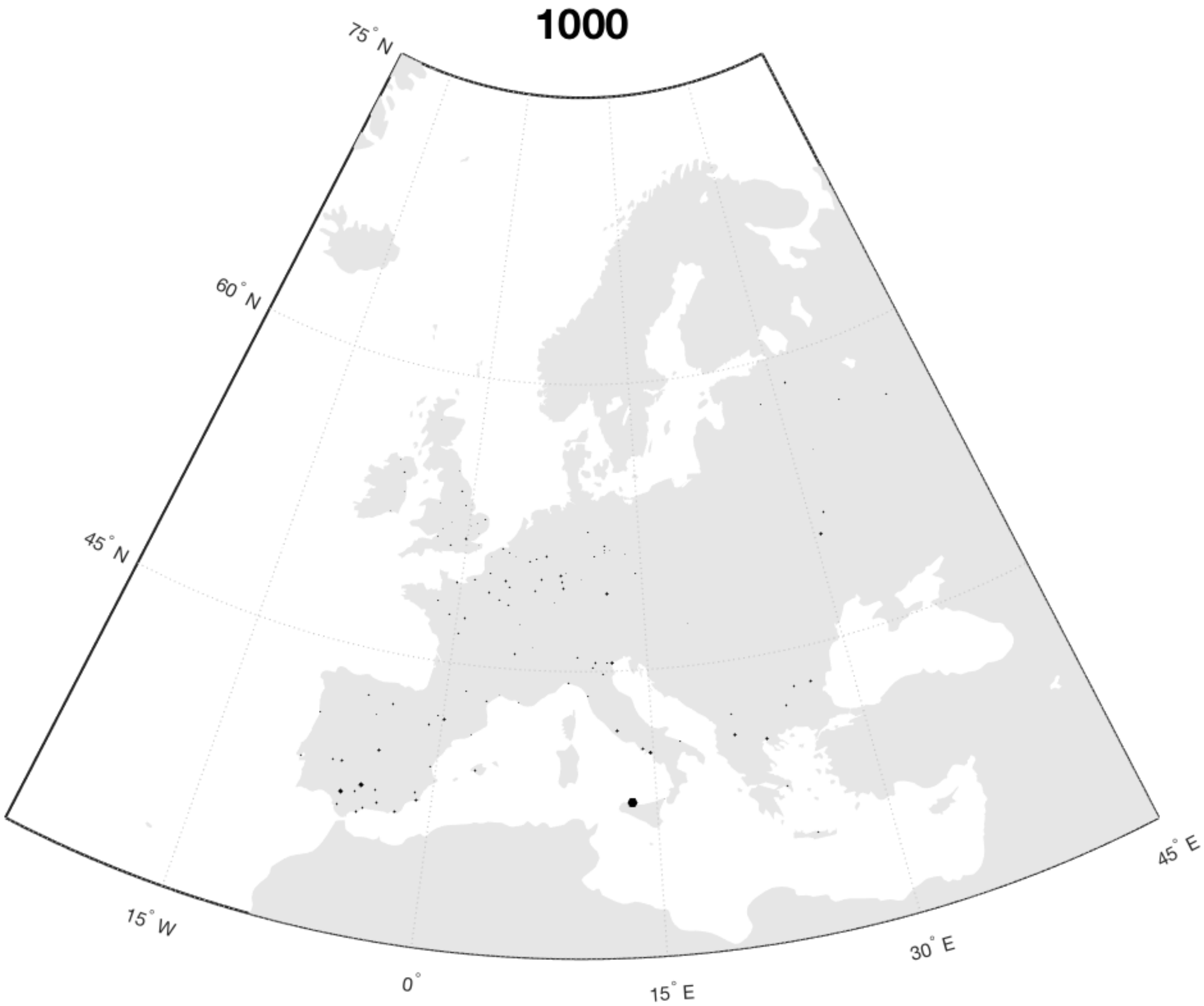
Note: Number of active states in each year with at least one city in the data by Bairoch et al. (1988).

(pink). In early 1632, Sweden had conquered significant portions of central Germany, and by late 1632 (shown in figure 3.2), they also took Silesia and most of Bavaria. About five years later (figure 3.3), the Holy Roman Empire had retaken control of most of the German lands, removing both the Swedes as well as smaller independent principalities in the Southwest.

The Atlas shows substantial variation over time in the number of active states and the extent of territory they control. Figure 3.4 shows the number of states over time that have at least one city in their territory. In 1100, around 30 states held *de facto* power in Europe, but by 1400 it was almost 60. Episodes of conflict with one state taking over territory from many small states are visible in the sudden drops in the number of states. For example, the number drops precipitously during our earlier example of the Thirty Years' War due to the conquest of many small German principalities by either the Habsburgs or the Swedes.

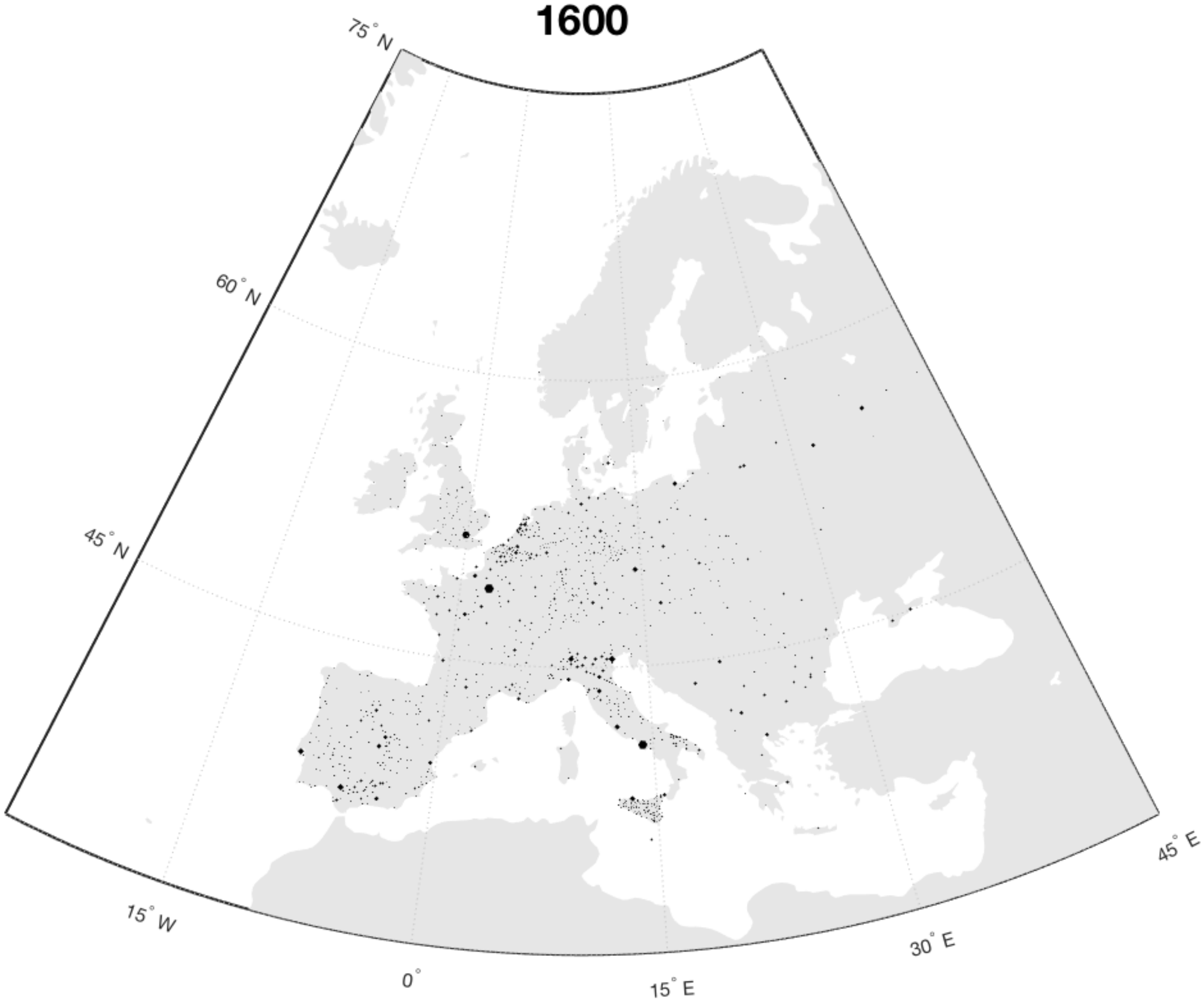
We combine the data on boundary changes from Reed (2016) with the city population data assembled by Bairoch et al. (1988). These show estimated population data for 2,182 cities across Europe for centuries in 800-1700 and fifty-year interval in 1750-1850 (that is, in total each city has up to 13 data points).

Figure 3.5: Geocoded city population data: year 1000



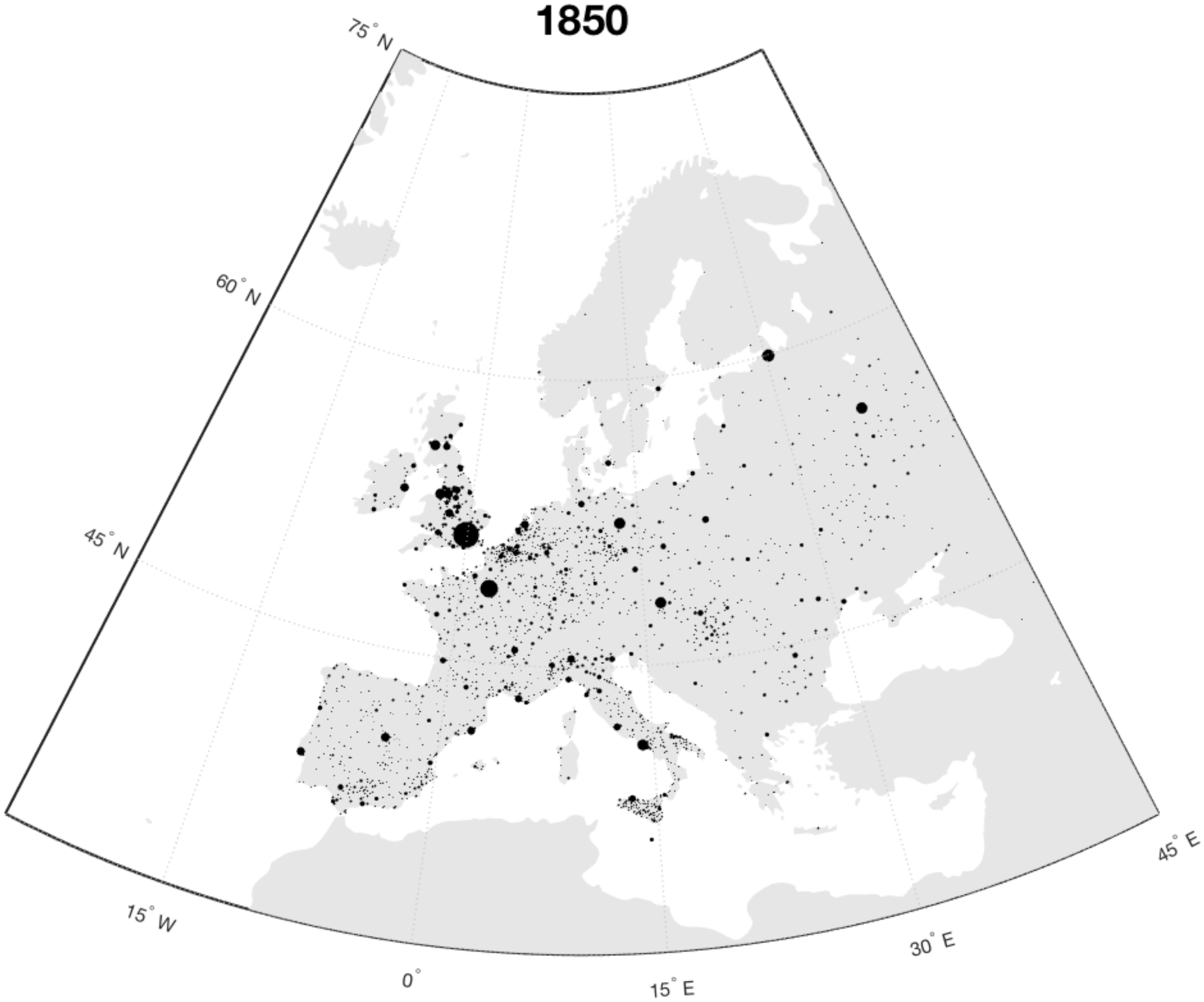
Note: European cities in 1000 based on data from Bairoch et al. (1988). Markers scaled to population size.

Figure 3.6: Geocoded city population data: year 1600



Note: European cities in 1600 based on data from Bairoch et al. (1988). Markers scaled to population size.

Figure 3.7: Geocoded city population data: year 1850



Note: European cities in 1850 based on data from Bairoch et al. (1988). 1850 is the most recent period coded. Markers scaled to population size.

Figures 3.5-3.7 show the newly geocoded city data for years 1000, 1600 and 1850, with markers scaled to population size.[3] At the turn of the millenium, the largest city in Europe was Palermo ruled at the time by the Emirate of Sicily. Other relatively large cities for this time period were found in Moorish Spain. Much of Central and Northern Europe had only minor cities and was largely rural. London and Paris both had less than 30,000 inhabitants, and Berlin, Moscow and Amsterdam have yet to show up in the data. By 1600, while Italy and Spain still had a number of sizable cities, the center of European urbanization has decisively shifted towards Northwestern Europe, especially the Low Countries. In 1850, Europe is sprawling with cities, and London has surpassed all others with more than 2 million inhabitants.

Combining high-frequency boundary data and city population data highlights a number of interesting facts about the frequency of changes in state control of cities. First, as shown in the top panel of figure 3.8, throughout most of the second millennium of European history, more than half of existing cities switched from one state to another in any given period (centuries up to 1700 and fifty-year intervals for 1750-1850).[4] Second, in the bottom panel of the figure we show that this affected a substantial amount of urban population: for example, in the period 1750-1800, 17 million European city dwellers out of a total of 25.4 million experienced at least one change in state control.

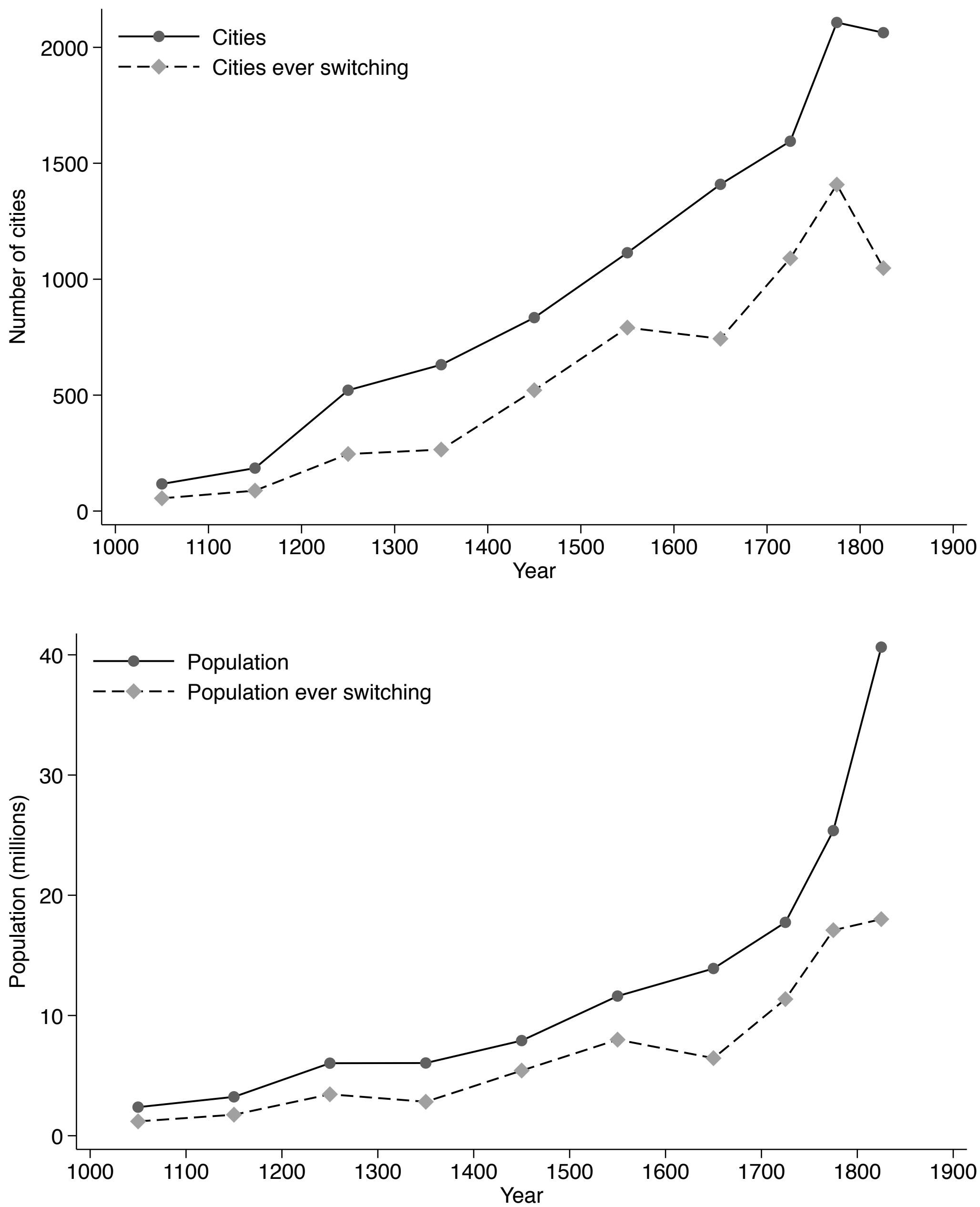
Third, as shown in figure 3.9, switching intensity rapidly increased and peaked in the beginning of the 18$^{\text{th}}$ century. While the probability of a city switching was between 10-20% per decade before, it now rose to more than 50%. In the second half of the 18$^{\text{th}}$ century, more than 1,000 city switches occurred across Europe. Switching intensity dropped off during the latter half of the 18$^{\text{th}}$ century and the early 19$^{\text{th}}$ century but stayed at a much higher level than before. This is consistent with the historical changes brought about by the French Revolution and the Napoleonic Wars (Acemoglu, Cantoni, et al., 2011). Overall, these facts illustrate how the association of cities with any state was not particularly strong.

We present an overview of further relevant city and state characteristics in table 3.1. In Panel A, we show summary statistics for cities. Noteworthy are in particular that the median city is subject to five different states over the course of their existence (i.e. when their population estimates are positive), switching eight times from one state to another, and only 18% of cities are controlled by the same state for the whole

[3] The original data is already geocoded, but with substantial error. We use the google maps API to geocode them with higher precision, with attention to city name changes after the fall of the Soviet Union.

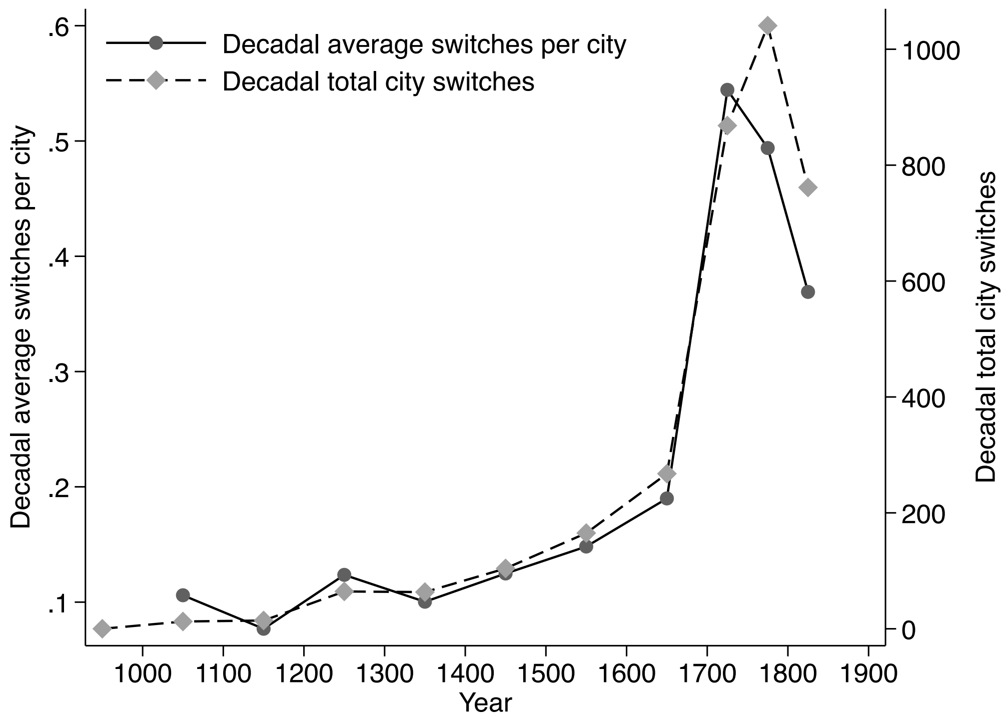
[4] Naturally, given the high frequency of the boundary data, many of these switches last only for a short period. However, we exclude switches that do not last past a new calendar year.

Figure 3.8: Cities and population switching state



Note: Top panel: Number of cities in each period, and number of cities that switch sovereign at least once during the period. Bottom panel: Total urban population, and urban population that switches sovereign at least once during the period. Periods are centuries for years 1000-1700 and fifty-year intervals for 1750-1850.

Figure 3.9: City exposure to switching



Note: Average number of sovereign switches per period (left axis), and total number of switching cities each period (right axis). Periods are centuries for years 1000-1700 and fifty-year intervals for 1750-1850.

period. In Panel B, we show statistics for states. It is interesting to note that the median lifespan of a European state is only 99.5 years.

3.4 Econometric Setup

We proceed as follows to investigate the effect of state power on urban growth. We study the absence of a powerful state that can provide security from foreign powers taking over the city by first estimating the effect of switching from one state to another on the size of the population. Second, to test for population displacement due cities switching states, we then include variables on whether nearby cities (within 100km) switch from one state to another.

Let pop_{it} be the population of cities $i = 1, ..., N$ for each time period t in 1000-

Table 3.1: Summary statistics for cities and states with cities

Panel A: Cities	Mean	Median	SD	Min	Max
First year in data	1567.6	1600	234.6	1000	1850
Highest population	20.3	10	62.7	1	2236
Average population	10.3	6.89	17.0	1	439.9
Average population of nearby cities	13.1	11.9	5.64	2.75	74.8
Average population of nearby allied cities	13.7	11.8	6.99	1.67	92.2
Number of states	4.90	5	2.70	1	11
City ever switches per period	0.59	0.62	0.35	0	1
City ever switches over 1000-1850	0.82	1	0.38	0	1
Number of switches per period	9.29	8	8.78	0	48
Any nearby cities switches per period	0.57	0.57	0.36	0	1
Any nearby allied cities switches per period	0.49	0.50	0.34	0	1
N	2,182				
Panel B: States with cities	Mean	Median	SD	Min	Max
First year	1390.6	1365	277.8	1000	1849
Last year	1587	1649	251.6	1002	1850
Years active	155.5	99.5	174.7	1	851
Most cities controlled	34.1	5	85.4	1	793
Least cities controlled	7.11	1	22.6	1	228
Most population controlled	426.6	81.5	1159.6	2	9260
Least population controlled	78.8	20	224.0	1	2268
N	192				

Note: *Note*: Population counts are in thousands. Only states with at least one city included. Close cities are those within 100 km. Allied cities are those that were in the same state in the preceding period. Both datasets are limited to the period 1000-1850 (which is when both are available).

1700 by century and 1750-1850 by 50-year interval. For the first step, we run the following regressions:

$$\log\left(\text{pop}_{it}\right) = \gamma_i + \phi_t + \psi_{\mathbf{J}(i,t)} + \text{Switch}_{it}\beta + \varepsilon_{it} \tag{3.1}$$

where γ_i is a city fixed effect, accounting for time-invariant city (and location) specific drivers of population growth; ϕ_t is a period fixed effect, controlling for continent-wide trends in population; and $\psi_{\mathbf{J}(i,t)}$ is a state fixed effect for the state $j = \mathbf{J}(i,t)$ that is in control of i in t, which accounts for state institutions that may be more or less conducive for growth. Notice that $\psi_{\mathbf{J}(i,t)}$ is separately identified from γ_i within a connected set of states that are linked by cities switching between them, which is akin to workers moving between firms to identify worker and firm effects separately

(Abowd et al., 1999). We also use variants of these fixed effects that include state-by-year fixed effects and current-country-by-year fixed effects, which account for varying state-specific capacity over time as well as time-varying region-wide factors affecting growth.[5]

The variable Switch_{it} is either an indicator for whether i switched at all between t and $t-1$ (called EverSwitch_{it}, see table 3.2). Essentially, we compare city populations of those that went through at least one change in state control to those that were controlled by the same state during the entire preceding period. Alternatively, Switch_{it} is the number of times i switched between t and $t-1$ (called NumSwitch_{it}, see table 3.3). β is our coefficient of interest and estimates the effect of switching on (the logarithm of) population. Finally, ε_{it} is an error term.

It is possible that population is merely displaced and may show up in nearby cities as a result of switching. To estimate these spillovers from nearby cities, we assess whether there were any switching cities within a radius of 100km of each city. That is, we calculate $\text{NearbySwitch}_{it} = \max_{\text{dist}(i,k)\leq 100\text{km}} \mathbf{1}[\text{EverSwitch}_{kt} = 1]$. We also separately measure the presence of nearby switching cities that were allied with the city of interest. That is, we first compute an indicator whether city i and city k were allied: $\text{Allied}_{ikt} = \mathbf{1}[\mathbf{J}(i, t-1) = \mathbf{J}(k, t-1)]$. We then use this to construct an indicator for whether any nearby allied city switched: $\text{NearbyAlliedSwitch}_{it} = \max_{\text{dist}(i,k)\leq 100\text{km}} \mathbf{1}[\text{Allied}_{ikt} \times \text{Switch}_{kt} > 0]$. Using these variables, we regress:

$$\begin{aligned}\log\left(\text{pop}_{it}\right) = \gamma_i + \phi_t + \psi_{\mathbf{J}(i,t)} + \text{EverSwitch}_{it}\beta + \text{NearbySwitch}_{it}\lambda \\ + \text{NeverSwitch}_{it} \times \text{NearbySwitch}_{it}\chi + \varepsilon_{it} \quad (3.2)\end{aligned}$$

where of course $\text{NeverSwitch}_{it} = 1 - \text{EverSwitch}_{it}$, and a similar equation in the case of nearby allied cities. The parameter λ represents the effect of nearby cities switching from one state to another on population in a given city. χ parameterizes the effect of city i having not switched state since the last period, but a nearby city has.[6]

3.5 Results

We organize results as follows: first, we present results from estimates of the direct effect according to equation (3.1). Second, we show spillovers from nearby cities

[5] Current countries are from the original data in Bairoch et al. (1988), meaning they reflect countries as of the mid-1980s.

[6] Naturally, this is algebraically equivalent with including EverSwitch_{kt} in the interaction; we use NeverSwitch_{it} because we are more interested in the effect of nearby cities switching holding the state for city i fixed.

Table 3.2: Effect of ever switching in period

Dependent variable:	log(population)				
	(1)	(2)	(3)	(4)	(5)
EverSwitch	-0.062***	-0.124***	-0.129***	-0.085***	-0.111***
	(0.016)	(0.017)	(0.025)	(0.021)	(0.021)
City FE	X	X	X	X	X
Period FE	X	X			
State FE		X			X
State-Period FE			X		
Country-Period FE				X	X
City N	2,063	2,060	2,060	2,061	2,058
Total N	10,576	10,555	10,485	10,553	10,533
R-squared	0.73	0.77	0.79	0.79	0.80

Note: OLS regression. Dependent variable: log(population). EverSwitch is an indicator variable for a city having switched states at least once in a period. Periods: 1000-1700 in centuries and 1750-1850 in fifty-year intervals. The country in country-period fixed effects are the countries as of 1988, as coded by Bairoch et al. (1988). Standard errors clustered on city level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

switching from one state to another, as in equation (3.2). And finally, we quantify the effect of both the direct effect and the spatial spillovers on the aggregate urban population in Europe.

Direct Effect of Switching State

We begin by looking at the estimates of the effect of whether a city has switched at all in the preceding period compared to stable rule, as formalized in equation (3.1) and reported in table 3.2. Each column varies the fixed effects included in the regression: all specifications use city fixed effects; in columns (1) and (2) we include year fixed effects; in columns (2) and (5) we include state fixed effects; in column (3) we include state-by-period fixed effects; and in columns (4) and (5) we include country-by-period fixed effects. Since our outcome variable is in logs, we can interpret the coefficient as the effect of switching on a change in population growth. We cluster standard errors on the city level throughout, although inference for the main results is largely unchanged with standard errors clustered at the state level, at the current-country level, using standard errors robust to spatial and serial autocorrelation, or standard errors clustered both on the city and the state level.

Our first key result is that we find substantial and stable estimates of switching

Table 3.3: Effect of number of switches per period

Dependent variable:	log(population)				
	(1)	(2)	(3)	(4)	(5)
NumSwitch	-0.016***	-0.019***	-0.018***	-0.009**	-0.015***
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
City FE	X	X	X	X	X
Period FE	X	X			
State FE		X			X
State-Period FE			X		
Country-Period FE				X	X
City N	2,063	2,060	2,060	2,061	2,058
Total N	10,458	10,437	10,377	10,438	10,418
R-squared	0.74	0.78	0.80	0.79	0.80

Note: OLS regression. Dependent variable: log(population). NumSwitch is the number of switches per period. Periods: 1000-1700 in centuries and 1750-1850 in fifty-year intervals. The country in country-period fixed effects are the countries as of 1988, as coded by Bairoch et al. (1988). Standard errors clustered on city level.

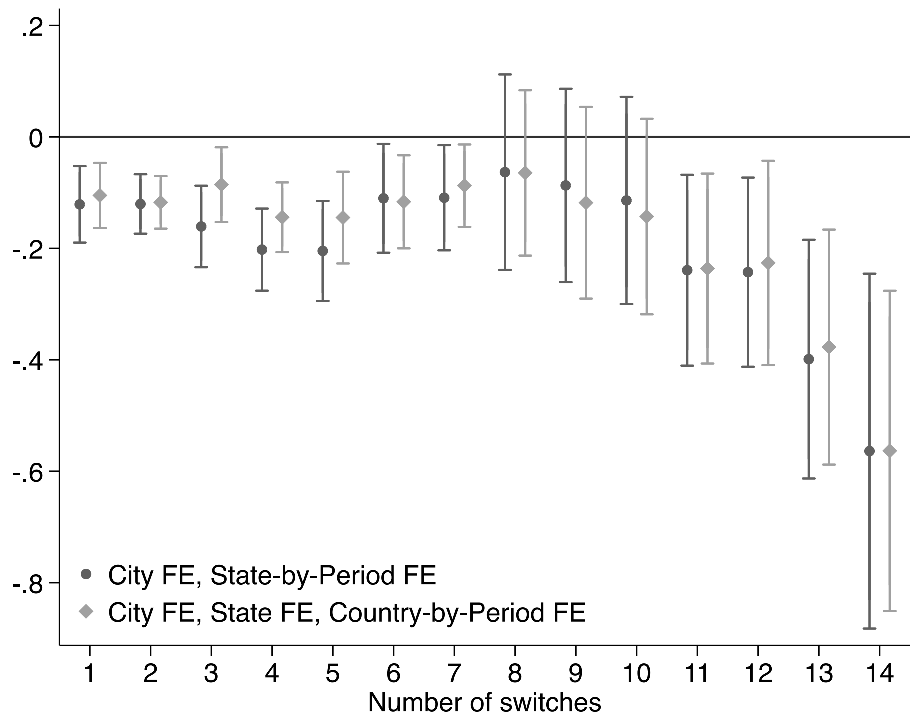
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

on population of around 6-13%. While these growth rate differentials are relatively small relative to the length of the periods (100 or 50 years), the frequency of switching across all cities and time periods implies considerable population losses due to unstable state rule over cities. It is interesting to note that the estimates get generally larger when controlling for state or state-by-period fixed effects, which suggests that heterogeneity in the effect of states to urban growth masks part of the effect of switching between states.

Table 3.3 shows the effect of an additional switch in a given period on urban population growth, with the same setup of fixed effects as before. We estimate that an additional switch leads to a decrease in population growth of around 0.9-1.9%. Recalling from table 3.1 that the average city undergoes 9.2 switches, the cumulative effect on city growth is substantial.

The estimates in table 3.3 impose linearity of the number of switches on population. In figure 3.10, we relax this assumption by nonparametrically estimating the effect of each number of switches on population. Specifically, we include a set of indicators for each number of switches in the same regression. We do so with the fixed effects specifications from columns (3) and (5) in table 3.2 and 3.3. We see that a negative effect of switching of around 10% kicks in after a single switch. This number is stable through most switching counts per period up to around ten, although

Figure 3.10: Nonparametric effect of switching



Note: Each dot is the coefficient on an indicator for a number of switches. 14 switches is 14 or more (up to 16). 95% confidence intervals. Standard errors clustered on city level. Regression models are the ones from column (3) and (5) in tables 3.2 and 3.3. 95% confidence intervals extend from the coefficient estimates.

imprecisely estimated for numbers of switches between 8 and 10. If a city has gone through even more changes in state control, the negative effect rapidly intensifies, dropping to almost 60% for 14 or more switches.

This nonlinearity is especially noteworthy because it suggests that not all of the negative effect of switching is due to physical destruction of the city. If it were, each additional change of hands from one state to another would be associated with a loss of population due to living conditions worsening with each military takeover. Instead, the fact that a single switch leads to a similar drop in population as several switches strengthens the case for uncertainty about the institutional settings and the rules associated with it as the main mechanism of the effect of changes in institutions on urban growth: a single switch can disrupt the status quo in a way that substantially changes the incentives for investments into economic growth.

Spillovers from Nearby City Switches

We now consider the possibility that the effect of a city switching from one state to another may lead to significant changes in the population of surrounding cities, as described in the context of equation (3.2). Table 3.4 shows coefficient estimates for this specification. All columns include city fixed effects; odd-numbered columns include state-by-period fixed effects; and even-numbered columns include state fixed effects and country-by-period fixed effects.

Columns (1) and (2) show our second key result: nearby switches further decrease population growth in a given city, controlling for whether the city itself switches. The direct effect is almost as high as in our baseline specification in table 3.2 at around 10%. The negative spillover effect is just over half that size at around 6%. As discussed in section 3.2, this may be due to either positive agglomeration effects between neighboring cities being disrupted by the changes in borders, and/or the fact that cities nearby also switched reflecting broader state instability in the region, such as in an extended territorial war.

In column (3) and (4), we show that spillovers are smaller and statistically insignificant for cities that were formerly in the same state as the city of interest. This is confirmed by including both all nearby cities as well as nearby allied cities in the regression, as in column (5) and (6): the negative spillovers from non-allied cities are large, but close to zero for allied cities. One interpretation of the lack of spillovers from allied cities is that the negative spillovers that come into effect generally are counteracted by a positive effect of population displacement from nearby cities to the city of interest.

Columns (6)-(10) show the specifications with interaction effects of the city of interest not switching but some of the neighboring (allied) cities do. While these estimates are not significant, three out of the four point estimates are positive, with the median estimate being 3.2%. This estimate strengthens the case for non-switching cities offer some refuge to populations displaced by switching cities.

Counterfactual Simulations

Finally, let us evaluate the magnitude of these effects in the context of aggregate urban population growth in Europe over this period. This thought experiment gives us a sense of the importance of strong states that are able to provide security to cities for their growth. To this end, we compute fitted values of equation (3.2), forcing different sets of coefficients to zero. We include city, state, and year fixed effects throughout.

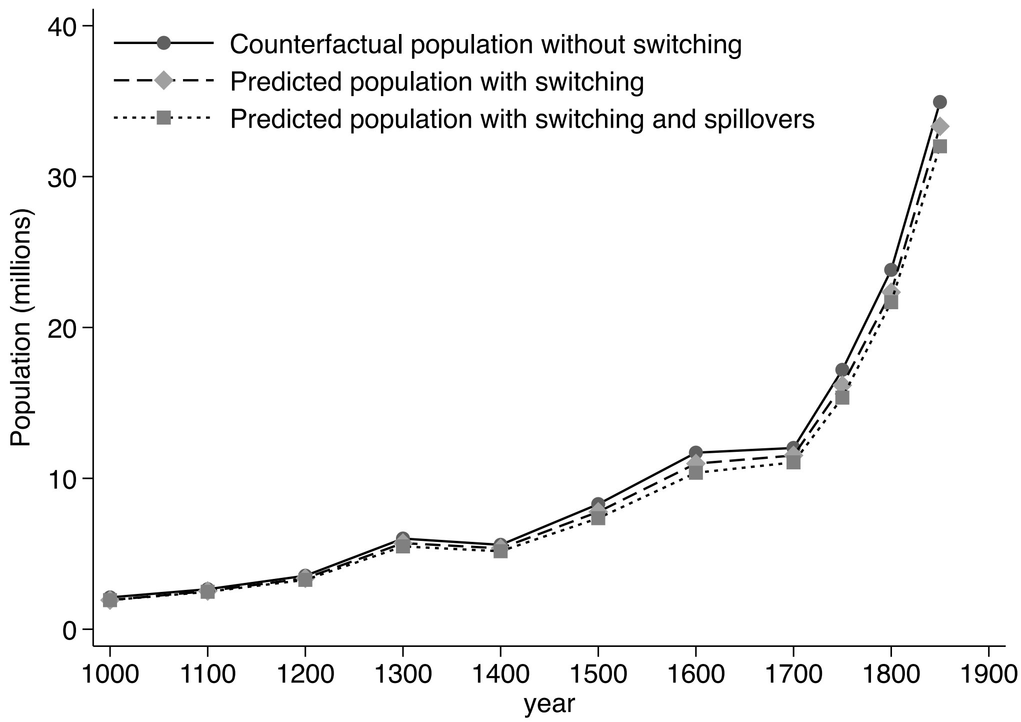
Table 3.4: Spillover effects of nearby switching

Dependent variable:	log(population)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
EverSwitch	-0.107***	-0.093***	-0.113***	-0.097***	-0.109***	-0.096***	-0.073*	-0.100**	-0.089**	-0.081***
	(0.027)	(0.023)	(0.028)	(0.024)	(0.028)	(0.024)	(0.043)	(0.039)	(0.035)	(0.031)
NearbySwitch	-0.058**	-0.057**			-0.064**	-0.063**	-0.091**	-0.050		
	(0.023)	(0.023)			(0.028)	(0.027)	(0.040)	(0.034)		
NearbyAlliedSwitch			-0.034	-0.028	0.008	0.010			-0.058*	-0.043
			(0.023)	(0.020)	(0.028)	(0.023)			(0.031)	(0.027)
NeverSwitch × NearbySwitch							0.049	-0.010		
							(0.047)	(0.041)		
NeverSwitch × NearbyAlliedSwitch									0.051	0.032
									(0.042)	(0.040)
City FE	X	X	X	X	X	X	X	X	X	X
State FE		X		X		X		X		X
State-Period FE	X		X		X		X		X	
Country-Period FE		X		X		X		X		X
City N	2,060	2,058	2,060	2,058	2,060	2,058	2,060	2,058	2,060	2,058
Total N	10,485	10,533	10,485	10,533	10,485	10,533	10,485	10,533	10,485	10,533
R-squared	0.79	0.80	0.79	0.80	0.79	0.80	0.79	0.80	0.79	0.80

Note: OLS regression. Dependent variable: log(population). EverSwitch is an indicator variable for a city having switched states at least once in a period; naturally, NeverSwitch is the opposite. NearbySwitch is an indicator for whether any city within 100km switched at least once; NearbyAlliedSwitch is the same but only for cities that were in the same state in the last period. Periods: 1000-1700 in centuries and 1750-1850 in fifty-year intervals. The country in country-period fixed effects are the countries as of 1988, as coded by Bairoch et al. (1988). Standard errors clustered on city level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Figure 3.11: Counterfactual aggregate population



Note: Predicted urban population in Europe with and without switching.

As a benchmark, we first compute fitted values for the unrestricted model, such that both the direct effect and the spillover effect are in place; we then impose that $\widehat{\lambda} = 0$, such that only the direct effect is present; and finally, we set $\widehat{\lambda} = \widehat{\beta} = 0$, shutting down both the direct effect and spillovers.[7]

Results are presented in figure 3.11. Predicted urban population by 1850 shutting down both direct and spillover effects – that is, in the absence of any switching – is 34.9 million.[8] Allowing only for the direct effect from switches, we predict population to be 4.2% lower in 1850. Allowing for both direct and spillover effects leads to a prediction of 31.9 million, which is 9.2% lower than in the absence of switching.

An important caveat to note is that we are not explicitly modeling the cumulative

[7] We set the interaction effect $\widehat{\chi} = 0$ throughout, but results are virtually identical without this restriction.

[8] This is substantially lower than the actual value of 40.6 million. Despite the inclusion of all the fixed effects, we systematically underpredict the strong urban growth during this period.

nature of these effects. A more sophisticated counterfactual prediction would explicitly take into account that earlier disruption through switching reverberates through the entire growth series of a given city. Also, the effect may be heterogeneous across different regions and time periods.

3.6 Conclusion

We show that European cities subject to switching states experienced lower population growth over the period 1000-1850. We interpret this effect as evidence for weak states being detrimental for urban growth, and conversely, strong states being an important cause of the rise of European economies. We also provide evidence that we interpret as agglomeration economies being disrupted through boundary changes, exacerbating the negative effect of weak states.

We conclude with a few avenues for future research using these data. First, it may be interesting to investigate further the city and state fixed effects in terms of their contribution to the variance in population growth over time. This type of exercise may illuminate how the role of the state has changed relative to the importance of location over time, similar to studies observing changes in the role of firms in the rise of inequality (Card et al., 2013).

Second, given that we directly observe the relative strength of a state based on the territorial changes with other states, we can use this for a jackknife instrumental variable strategy: leaving out all territorial changes between state j and state ℓ, we can estimate the (relative) power of j by estimating territorial changes of ℓ with all other states except j. Preliminary results suggest that we get a sufficiently strong first stage to explore this further. These estimates can form the foundation of a structural estimate of state power that may be more broadly applicable to other settings.

Bibliography

[1] John M. Abowd, Francis Kramarz, and David N. Margolis. “High Wage Workers and High Wage Firms”. In: *Econometrica* 67.2 (Feb. 1999), pp. 251–333.

[2] Daron Acemoglu, Davide Cantoni, Simon Johnson, and James A. Robinson. “The Consequences of Radical Reform: The French Revolution”. In: *American Economic Review* 101.7 (Dec. 2011), pp. 3286–3307.

[3] Daron Acemoglu, Simon Johnson, and James Robinson. “The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth”. In: *American Economic Review* 95.3 (Sept. 2005), pp. 546–579.

[4] Daron Acemoglu, Simon Johnson, and James A. Robinson. “Institutions as a Fundamental Cause of Long-Run Growth”. In: *Handbook of Economic Growth* 1, Part A (2005). Ed. by Philippe Aghion and Steven N. Durlauf, pp. 385–472.

[5] Daron Acemoglu, Simon Johnson, and James A. Robinson. “Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution*”. In: *The Quarterly Journal of Economics* 117.4 (2002), p. 1231.

[6] Robert C Allen. “Agriculture and the Origins of the State in Ancient Egypt”. In: *Explorations in Economic History* 34.2 (1997), pp. 135–154.

[7] Treb Allen and Costas Arkolakis. “Trade and the Topography of the Spatial Economy”. In: *The Quarterly Journal of Economics* 129.3 (2014), pp. 1085–1140.

[8] Michelle Wilde Anderson. “Cities Inside Out: Race, Poverty, and Exclusion at the Urban Fringe”. In: *UCLA Law Review* 55.1095-1160 (2008).

[9] Joshua D Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist’s Companion*. Princeton university press, 2008.

[10] Amaia Arranz-Otaegui, Sue Colledge, Lydia Zapata, Luis Cesar Teira-Mayolini, and Juan José Ibáñez. “Regional diversity on the timing for the initial appearance of cereal cultivation and domestication in southwest Asia”. In: *Proceedings of the National Academy of Sciences* 113.49 (2016), pp. 14001–14006.

[11] Quamrul Ashraf and Stelios Michalopoulos. "Climatic Fluctuations and the Diffusion of Agriculture". In: *Review of Economics and Statistics* 97.3 (2015), pp. 589–609.

[12] D.Andrew Austin. "Politics vs Economics: Evidence from Municipal Annexation". In: *Journal of Urban Economics* 45.3 (1999), pp. 501–532.

[13] Paul Bairoch, Jean Batou, and Pierre Chevre. *The Population of European Cities. Data Bank and Short Summary of Results: 800-1850.* Geneva Switzerland Librairie Droz 1988., 1988.

[14] Lisa Barrow and Cecilia Elena Rouse. "Using Market Valuation to Assess Public School Spending". In: *Journal of Public Economics* 88.9–10 (2004), pp. 1747–1769.

[15] Robert H Bates. *Essays on the Political Economy of Rural Africa.* Vol. 38. Univ of California Press, 1987.

[16] Patrick Bayer, Fernando Ferreira, and Robert McMillan. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods". English. In: *Journal of Political Economy* 115.4 (2007), pp. 588–638.

[17] Patrick Bayer, Robert McMillan, Alvin Murphy, and Christopher Timmins. "A Dynamic Model of Demand for Houses and Neighborhoods". In: *Econometrica* 84.3 (May 2016), pp. 893–942.

[18] Larry Benson, Kenneth Petersen, and John Stein. "Anasazi (Pre-Columbian Native-American) Migrations During The Middle-12Th and Late-13th Centuries – Were they Drought Induced?" In: *Climatic Change* 83.1 (2007), pp. 187–213.

[19] Gemma Berniell-Lee et al. "Genetic and Demographic Implications of the Bantu Expansion: Insights from Human Paternal Lineages". In: *Molecular Biology and Evolution* 26.7 (2009), p. 1581.

[20] Timothy Besley and Torsten Persson. "State Capacity, Conflict, and Development". In: *Econometrica* 78.1 (2010), pp. 1–34.

[21] Timothy Besley and Torsten Persson. "The Origins of State Capacity: Property Rights, Taxation, and Politics". In: *American Economic Review* 99.4 (Sept. 2009), pp. 1218–44.

[22] Robert L. Bettinger. "Prehistoric hunter–gatherer population growth rates rival those of agriculturalists". In: *Proceedings of the National Academy of Sciences* 113.4 (2016), pp. 812–814.

[23] Sandra E. Black. “Do Better Schools Matter? Parental Valuation of Elementary Education”. In: *The Quarterly Journal of Economics* 114.2 (1999), pp. 577–599.

[24] Valerie Bockstette, Areendam Chanda, and Louis Putterman. “States and Markets: The Advantage of an Early Start”. English. In: *Journal of Economic Growth* 7.4 (2002),

[25] Peter Bogucki. *The Origins of Human Society*. Blackwell Publishers, 1999.

[26] Leah Platt Boustan. “Local Public Goods and the Demand for High-Income Municipalities”. In: *Journal of Urban Economics* 76 (2013), pp. 71–82.

[27] Katharine L. Bradbury, Christopher J. Mayer, and Karl E. Case. “Property tax limits, local fiscal behavior, and property values: evidence from Massachusetts under Proposition 212”. In: *Journal of Public Economics* 80.2 (2001), pp. 287–311.

[28] Geoffrey Brennan and James M Buchanan. *The Power to Tax: Analytic Foundations of a Fiscal Constitution*. Cambridge University Press, 1980.

[29] Stephen Broadberry and Bishnupriya Gupta. “The Early Modern Great Divergence: Wages, Prices and Economic Development in Europe and Asia, 1500–1800”. In: *The Economic History Review* 59.1 (Feb. 2006), pp. 2–31.

[30] Jan K. Brueckner. “Property Values, Local Public Expenditure and Economic Efficiency”. In: *Journal of Public Economics* 11.2 (1979), pp. 223–245.

[31] Elizabeth M Brumfiel. *The Economic Anthropology of the State*. 11. Lanham: University Press of America, 1994.

[32] Marcy Burchfield, Henry G Overman, Diego Puga, and Matthew A Turner. “Causes of Sprawl: A Portrait from Space”. In: *The Quarterly Journal of Economics* 121.2 (2006), pp. 587–633.

[33] Matias Busso, Jesse Gregory, and Patrick Kline. “Assessing the Incidence and Efficiency of a Prominent Place Based Policy”. In: *American Economic Review* 103.2 (2013), pp. 897–947.

[34] Anna M. Caballero. *Guide to the Cortese-Knox-Hertzberg Local Government Reorganization Act of 2000*. Tech. rep. Assembly Committee on Local Government, California Legislature, 2009.

[35] Stephen Calabrese, Dennis Epple, Thomas Romer, and Holger Sieg. “Local Public Good Provision: Voting, Peer Effects, and Mobility”. In: *Journal of Public Economics* 90.6–7 (2006), pp. 959–981.

[36] David Card, Jörg Heining, and Patrick Kline. “Workplace Heterogeneity and the Rise of West German Wage Inequality*”. In: *The Quarterly Journal of Economics* 128.3 (2013), pp. 967–1015.

[37] David Card and Alan B. Krueger. “School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina”. In: *Journal of Economic Perspectives* 10.4 (Dec. 1996), pp. 31–50.

[38] Robert L. Carneiro. “A Theory of the Origin of the State”. In: *Science* 169.3947 (1970), pp. 733–738.

[39] Richard T. Carson. “Contingent Valuation: A Practical Alternative When Prices Aren’t Available”. In: *Journal of Economic Perspectives* 26.4 (2012), pp. 27–42.

[40] Stephanie Riegg Cellini, Fernando Ferreira, and Jesse Rothstein. “The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design”. In: *The Quarterly Journal of Economics* 125.1 (2010), pp. 215–261.

[41] Charlotte Chatfield. “Food Composition Tables for International Use”. In: *FAO Nutritional Studies No. 3.* Nutrition Division, FAO UN, Washington, DC, 1953, 56 pp.

[42] V Gordon Childe. *Man Makes Himself*. London, Watts & Co, 1936.

[43] Gregory Clark. *A Farewell to Alms: a Brief Economic History of the World.* Princeton University Press, 2008.

[44] Michael Coleman. *Rethinking the Property Tax*. Tech. rep. The California Local Government Almanac, 2015.

[45] T.G. Conley. “GMM Estimation with Cross Sectional Dependence”. In: *Journal of Econometrics* 92.1 (1999), pp. 1–45.

[46] Alison Crowther et al. “Ancient crops provide first archaeological signature of the westward Austronesian expansion”. In: *Proceedings of the National Academy of Sciences* 113.24 (2016), pp. 6635–6640.

[47] Ernesto Dal Bó, Pablo Hernández, and Sebastián Mazzuca. “The Paradox of Civilization: Pre-Institutional Sources of Security and Prosperity”. In: *Working Paper* (2016).

[48] Melissa Dell. “The Persistent Effects of Peru’s Mining “Mita””. English. In: *Econometrica* 78.6 (2010),

[49] Melissa Dell, Nathaniel Lane, and Pablo Querubin. *The Historical State, Local Collective Action, and Economic Development in Vietnam*. Working Paper 23208. National Bureau of Economic Research, Mar. 2017.

[50] Jared M Diamond. *Guns, Germs, and Steel: A Short History of Everybody for the Last 13,000 Years*. Random House, 1998.

[51] Jared Diamond and Peter Bellwood. "Farmers and Their Languages: The First Expansions". In: *Science* 300.5619 (2003), pp. 597–603.

[52] Rebecca Diamond. "Housing Supply Elasticity and Rent Extraction by State and Local Governments". In: *American Economic Journal: Economic Policy* 9.1 (Feb. 2017), pp. 74–111.

[53] Mary M Edwards. "Understanding the Complexities of Annexation". In: *Journal of Planning Literature* 23.2 (2008), pp. 119–135.

[54] Dennis Epple and Thomas Romer. "On the Rlexibility of Municipal Boundaries". In: *Journal of Urban Economics* 26.3 (1989), pp. 307–319.

[55] Dennis Epple, Thomas Romer, and Holger Sieg. "Interjurisdictional Sorting and Majority Rule: An Empirical Analysis". English. In: *Econometrica* 69.6 (2001),

[56] Dennis Epple and Holger Sieg. "Estimating Equilibrium Models of Local Jurisdictions". In: *Journal of Political Economy* 107.4 (1999), pp. 645–681.

[57] David L. Erickson, Bruce D. Smith, Andrew C. Clarke, Daniel H. Sandweiss, and Noreen Tuross. "An Asian origin for a 10,000-year-old domesticated plant in the Americas". In: *Proceedings of the National Academy of Sciences* 102.51 (2005), pp. 18315–18320.

[58] Rex L. Facer. "Annexation Activity and State Law in the United States". In: *Urban Affairs Review* 41.5 (2006), pp. 697–709.

[59] James D. Fearon and David D. Laitin. "Ethnicity, Insurgency, and Civil War". In: *American Political Science Review* 97.1 (2003), pp. 75–90.

[60] James Fenske. "Does Land Abundance Explain African Institutions?" In: *The Economic Journal* 123.573 (Dec. 2013), pp. 1363–1390.

[61] James Fenske. "Ecology, Trade, and States in Pre-Colonial Africa". In: *Journal of the European Economic Association* (2014), n/a–n/a.

[62] Samuel Edward Finer. *The History of Government from the Earliest Times: Ancient Monarchies and Empires*. Vol. 1. Oxford University Press, 1997.

[63] Guenther Fischer et al. *Global Agro-Ecological Zones Assessment for Agriculture (GAEZ 2008).* Tech. rep. IIASA, Laxenburg, Austria and FAO, Rome, Italy, 2008.

[64] Roger Fouquet and Stephen Broadberry. "Seven Centuries of European Economic Growth and Decline". In: *Journal of Economic Perspectives* 29.4 (2015), pp. 227–44.

[65] William T Fujioka. *Unincorporated Area Patrol Service Levels.* Tech. rep. County of Los Angeles, 2014.

[66] Oded Galor and Ömer Özak. "The Agricultural Origins of Time Preference". In: *American Economic Review* 106.10 (Oct. 2016), pp. 3064–3103.

[67] Nicola Gennaioli and Hans-Joachim Voth. "State Capacity and Military Conflict". In: *The Review of Economic Studies* 82.4 (2015), pp. 1409–1448.

[68] Liviu Giosan et al. "Fluvial Landscapes of the Harappan Civilization". In: *Proceedings of the National Academy of Sciences* 109.26 (2012), E1688–E1694.

[69] Jonathan Haas. *The Evolution of the Prehistoric State.* Columbia University Press, 1982.

[70] Jerry Hausman. "Contingent Valuation: From Dubious to Hopeless". In: *Journal of Economic Perspectives* 26.4 (2012), pp. 43–56.

[71] Robert J. Hijmans, Susan E. Cameron, Juan L. Parra, Peter G. Jones, and Andy Jarvis. "Very high resolution interpolated climate surfaces for global land areas". In: *International Journal of Climatology* 25.15 (Dec. 2005), pp. 1965–1978.

[72] Philip T Hoffman. *Why did Europe Conquer the World?* Princeton University Press, 2015.

[73] Zuzana Hofmanová et al. "Early farmers from across Europe directly descended from Neolithic Aegeans". In: *Proceedings of the National Academy of Sciences* 113.25 (2016), pp. 6886–6891.

[74] William C. Horrace and Ronald L. Oaxaca. "Results on the bias and inconsistency of ordinary least squares for the linear probability model". In: *Economics Letters* 90.3 (2006), pp. 321–327.

[75] C. Kirabo Jackson, Rucker C. Johnson, and Claudia Persico. "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms". In: *The Quarterly Journal of Economics* 131.1 (2016), pp. 157–218.

[76] Allen W Johnson and Timothy K Earle. *The Evolution of Human Societies: From Foraging Group to Agrarian State*. Stanford University Press, 2000.

[77] Frank Kleibergen and Richard Paap. “Generalized Reduced Rank Tests using the Singular Value Decomposition”. In: *Journal of Econometrics* 133.1 (2006), pp. 97–126.

[78] Patrick Kline and Enrico Moretti. “People, places, and public policy: Some simple welfare economics of local economic development programs”. In: *Annual Review of Economics* 6.1 (2014), pp. 629–662.

[79] Stephen Knack and Philip Keefer. “Institutions and Econoic Performance: CrossCountry Tests using Alternative Institutional Measures”. In: *Economics & Politics* 7.3 (Nov. 1995), pp. 207–227.

[80] Julien Lafortune, Jesse Rothstein, and Diane Whitmore Schanzenbach. “School Finance Reform and the Distribution of Student Achievement”. In: *American Economic Journal: Applied Economics* Forthcoming (2017).

[81] Greger Larson et al. “Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA”. In: *Proceedings of the National Academy of Sciences* 107.17 (2010), pp. 7686–7691.

[82] B. Lehner, K. Verdin, and A. Jarvis. “New Global Hydrography Derived From Spaceborne Elevation Data”. In: *Eos, Transactions American Geophysical Union* 89.10 (Mar. 2008).

[83] Leigh Linden and Jonah E. Rockoff. “Estimates of the Impact of Crime Risk on Property Values from Megan’s Laws”. In: *American Economic Review* 98.3 (June 2008), pp. 1103–27.

[84] Michael Mann. “The Autonomous Power of the State: Its origins, Mechanisms and Results”. In: *European Journal of Sociology / Archives Européennes de Sociologie / Europäisches Archiv für Soziologie* 25.2 (1984), pp. 185–213.

[85] Michael Mann. *The Sources of Social Power: A History of Power from the Beginning to AD 1760, volume I*. 1986.

[86] Andrea Matranga. “The Ant and the Grasshopper: Seasonality and the Invention of Agriculture”. In: *Working Paper* (2017).

[87] Joram Mayshar, Omer Moav, Zvika Neeman, and Luigi Pascali. “Cereals, Appropriability and Hierarchy”. In: *Working Paper* (2017).

[88] Patricia A McAnany. *Living with the Ancestors: Kinship and Kingship in Ancient Maya Society*. Cambridge University Press, 2014.

[89] Howard Lee McBain. *The Law and the Practice of Municipal Home Rule.* Columbia University Press, 1916.

[90] Daniel McFadden. "Conditional Logit Analysis of Qualitative Choice Behavior". In: *Institute of Urban and Regional Development, University of California* (1973).

[91] William H McNeill. *The Pursuit of Power: Technology, Armed force, and Society.* 1982.

[92] Stelios Michalopoulos and Elias Papaioannou. "National Institutions and Subnational Development in Africa". In: *The Quarterly Journal of Economics* 129.1 (2014), pp. 151–213.

[93] Stelios Michalopoulos and Elias Papaioannou. "Pre-Colonial Ethnic Institutions and Contemporary African Development". In: *Econometrica* 81.1 (2013), pp. 113–152.

[94] George Peter Murdock. *Ethnographic Atlas.* Pittsburgh, PA: University of Pittsburgh Press, 1967.

[95] Nathan Nunn and Diego Puga. "Ruggedness: The Blessing of Bad Geography in Africa". In: *Review of Economics and Statistics* 94.1 (2012), pp. 20–36.

[96] Nathan Nunn and Nancy Qian. "The Columbian Exchange: A History of Disease, Food, and Ideas". In: *Journal of Economic Perspectives* 24.2 (2010), pp. 163–88.

[97] Nathan Nunn and Nancy Qian. "The Potato's Contribution to Population and Urbanization: Evidence From A Historical Experiment". In: *The Quarterly Journal of Economics* 126.2 (2011), pp. 593–650.

[98] Mancur Olson. "Dictatorship, Democracy, and Development". In: *American Political Science Review* (1993), pp. 567–576.

[99] Ola Olsson and Douglas A Hibbs Jr. "Biogeography and Long-run Economic Development". In: *European Economic Review* 49.4 (2005), pp. 909–938.

[100] Adam Powell, Stephen Shennan, and Mark G. Thomas. "Late Pleistocene Demography and the Appearance of Modern Human Behavior". In: *Science* 324.5932 (2009), pp. 1298–1301.

[101] Louis Putterman. "Agriculture, Diffusion and Development: Ripple Effects of the Neolithic Revolution". English. In: *Economica.* New Series 75.300 (2008),

[102] Navin Ramankutty and Jonathan A. Foley. "Estimating historical changes in global land cover: Croplands from 1700 to 1992". In: *Global Biogeochemical Cycles* 13.4 (1999), pp. 997–1027.

[103] Frank Reed. *The Centennia Historical Atlas: Academic Research Edition.* Clockwork Mapping: Jamestown RI. 2016.

[104] Shawn J Riley, Stephen J DeGloria, and Robert Elliot. "Index That Quantifies Topographic Heterogeneity". In: *Intermountain Journal of Sciences* 5.1–4 (1999), pp. 23–27.

[105] Jennifer Roback. "Wages, Rents, and the Quality of Life". English. In: *Journal of Political Economy* 90.6 (1982),

[106] Sherwin Rosen. "Wage-based Indexes of Urban Quality of Life". In: *Current Issues in Urban Economics*. Ed. by Peter Mieszkowski and Mahlon Straszheim. Baltimore: Johns Hopkins University Press, 1979, pp. 47–104.

[107] Frank Salomon. *Native Lords of Quito in the Age of the Incas: The Political Economy of North Andean Chiefdoms*. Cambridge University Press Cambridge, 1986.

[108] Paul A. Samuelson. "The Pure Theory of Public Expenditure". In: *The Review of Economics and Statistics* 36.4 (1954), pp. 387–389.

[109] Raul Sanchez de la Sierra. "On the Origin of States: Stationary Bandits and Taxation in Eastern Congo". In: *Working Paper* (2017).

[110] James C Scott. *The Art of Not Being Governed: An Anarchist History of Upland Southeast Asia*. Yale University Press, 2009.

[111] Elman R Service. *Origins of the State and Civilization*. New York. Norton, 1975.

[112] Heather Taylor. *Cost of Constructing a Home*. Tech. rep. NAHB Economics and Housing Policy Group, 2015.

[113] Charles M. Tiebout. "A Pure Theory of Local Expenditures". English. In: *Journal of Political Economy* 64.5 (1956),

[114] Charles Tilly. *Capital, Coercion and European States*. Cambridge, MA: Basil Blackwell, 1990.

[115] Waldo Tobler. "Linear Pycnophylactic Reallocation Comment on a Paper by D. Martin". In: *International Journal of Geographical Information Science* 13.1 (1999), pp. 85–90.

[116] Bruce G Trigger. *Understanding Early Civilizations: A Comparative Study*. Cambridge University Press, 2003.

[117] Matthew A. Turner, Andrew Haughwout, and Wilbert van der Klaauw. "Land Use Regulation and Welfare". In: *Econometrica* 82.4 (July 2014), pp. 1341–1403.

[118] U.S. Census Bureau. *2011-2015 American Community Survey 5-Year Estimates*. 2016.

[119] U.S. Census Bureau. *Boundary and Annexation Survey*. 2015.

[120] U.S. Census Bureau. *TIGER/Line Shapefiles Technical Documentation*. Tech. rep. U.S. Census Bureau Geography Division, 2016.

[121] U.S. Census Bureau Government Division. *2012 Census of Governments: Individual State Descriptions*. U.S. Government Printing Office, Washington, DC, 2013.

[122] Nico Voigtländer and Hans-Joachim Voth. "The Three Horsemen of Riches: Plague, War, and Urbanization in Early Modern Europe". In: *The Review of Economic Studies* 80.2 (2013), pp. 774–811.

[123] Henry T Wright. "Atlas of Chiefdoms and Early States". In: *Structure and Dynamics* 1.4 (2006).

Chapter 4

Appendix

4.1 Appendix for Chapter 1

Data Sources and Preparation

Boundary Data

To obtain the boundary change data, we contacted individual county and municipal offices, who sent us the change shapefile if available. We then standardized all these files by combining them into a single shapefile showing the municipality, the county, the type (annexed or unincorporated), the year of annexation (if annexed), and the acreage. This shapefile provided the basis to which we matched the property locations. Table 4.1 shows a list of the counties from which we obtained data as well as the share of spheres we obtained for each of them.

In total, we obtained 189 spheres from across California. Figure 4.2 shows San Jose as an example of a sphere with the location of properties. Figure 4.3 shows a map of California with the included municipalities.

Public Goods Data

To construct a measure of per capita expenditures in unincorporated county territory, we do the following: we sum county expenditures for police protection, fire protection and library services, all of which are generally targeted at unincorporated areas, and divide them by the population in the unincorporated areas. For per capita expenditures in the municipality, we divide total municipal expenditures by the population in the municipality. Other services provided by counties, such as real estate assessment, elections, administration of state public welfare programs or hospitals are open to all residents of the county, including those in municipalities.

Table 4.1: List of counties in data.

County	Total pop.	Uninc. pop.	# spheres	# in data	Share
Los Angeles	9,980,432	1,064,241	88	88	100%
Riverside	2,253,516	585,784	28	28	100%
Kern	861,646	307,977	14	11	79%
Ventura	836,864	96,716	10	10	100%
Stanislaus	523,707	111,414	9	9	100%
Tulare	455,025	149,820	8	8	100%
Santa Barbara	430,103	162,268	9	8	89%
San Joaquin	701,620	143,548	7	7	100%
San Luis Obispo	270,768	118,536	7	7	100%
Placer	361,733	111,069	6	6	100%
San Bernardino	2,069,806	324,717	24	3	13%
Butte	221,430	101,902	5	2	40%
Sonoma	488,661	164,382	9	2	22%
Santa Clara	1,844,389	105,629	15	2	13%
Fresno	954,040	168,769	14	2	14%
Sacramento	1,442,993	686,414	7	1	14%
Orange	3,087,715	232,900	34	1	3%
San Diego	3,164,818	494,542	18	1	6%

Note: Spheres shows the number of spheres in the data. Share is the share of all spheres that exist in the county.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

We use the adjusted crime clearance rate instead of the crime rate per capita because county sheriff departments often take on additional responsibilities beyond services for unincorporated areas such as running county jails or guarding courts. Crimes reported by county sheriff departments relative to the population in unincorporated areas are thus likely to be inflated. The clearance rate does not require a population denominator and is thus robust to this concern.

To deal with the possibility that the types of crimes faced by county sheriff departments is systematically different (and thus systematically easier or harder to clear), we adjust raw clearance rates in the following way. For types of crime $k = 1, ..., K$ and providers $j = \{\text{County}, \text{Municipality}\}$ across all reporting agencies (both police and sheriff) $c = 1, ..., C$ in California for each year $t = 1985, ..., 2013$, we compute

$$r_{kt} = \frac{1}{K} \sum_{c=1}^{C} \frac{x_{ckt}}{y_{ckt}}$$

where x_{ckt} is the number of cleared cases reported by agency c of crime k in year t, and y_{ckt} is the number of reported crimes. Thus, r_{kt} is an adjusted clearance rate of crime k in year t across all agencies. The adjustment takes care of the fact that some crimes are harder to clear than others, such as theft.

4.2 Appendix for Chapter 2

Data Sources and Preparation

All datasets except the archaeological data from Bogucki (1999) are raster data (i.e. grid-cell level data), usually at resolution of ⅟₁₂ or higher. These datasets have been down-sampled to ¼ (or another cell size) using nearest-neighbor interpolation.

Archaeological Sites Related to State Formation

We use the maps in Bogucki's (1999) chapter 8 ("Early States and Chiefdoms in the Shadow of States"). Since the maps are no longer digitally available, all relevant sites were geocoded by hand using google maps, relying on satellite evidence of the site whenever possible. We exclude chiefdom sites. We also exclude state sites that have been deemed not pristine in the text. This leaves us with 68 sites spread across Mesoamerica, South America, Africa, the Near East, South Asia and East Asia.

FAO Agricultural Data and Derivates

Land Quality Proxy. Agricultural productivity is the maximum potential production capacity in tons per hectare over the seventeen crops (buckwheat, barley, chickpea, foxtail millet, groundnut, maize, oat, pearl millet, wetland rice, rape, rye, sunflower, soybean, sweet potato, sorghum, wheat and white potato) after correcting for the Columbian Exchange (e.g. no wheat in the Americas, no potatoes in Eurasia). See http://gaez.fao.org/Main.html for the database.

We use the earliest data available (baseline period 1961-1990), with intermediate input level and irrigated water supply. While the data is also available for rainfall water supply, it renders areas that were very productive but relied on river flooding and alluvial agriculture completely unproductive if they are in arid environments. Note that most early civilizations substantially relied on irrigation based on freshwater delivered via rivers. As such, the data with rainfall water supply does not represent a good proxy for agricultural land quality when other water sources are available. For example, according to the FAO GAEZ data on potential production capacity using low input level and rainfed agriculture (and data derived from it, such as Galor and Özak's (2016) Caloric Suitability Index), Egypt, including all of the

Nile valley and most of the Delta, are completely unproductive, despite thousands of years of highly productive agriculture. Concerning the input level, the intermediate level is the lowest for irrigated water supply.

Next, we apply restrictions from the Columbian Exchange on which crops are available to a given region. To this end, we divide the crops into Old World crops, applied to Africa, Asia and Europe, and New World Crops, applied to the Americas. New World crops are maize, sweet potato, white potato and sunflower; the others are Old World crops.

In the next step, each crop is scaled using the historical Food Composition Tables by the FAO (Chatfield, 1953), using the following calorie values per 100 grams for each crop:

- buckwheat: 330
- barley: 332
- chickpea: 345
- foxtail millet: 343
- groundnut: 388
- maize: 356
- oat: 385
- pearl millet: 348
- wetland rice: 357
- rape: 26
- rye: 319
- sunflower: 284
- soybean: 335
- sweet potato: 97
- sorghum: 343
- wheat: 334
- white potato: 70

Finally, we take the maximal calorie amount across all crops available in a region as the land quality proxy.

Circumscription. To compute circumscription, we subtract average land quality in all cells within radius r (without the cell of interest) from the value at the cell of interest, as described in the main text. This is illustrated in figure 4.4.

Standard Deviation of Land Quality. The standard deviation in land quality is computed for each cell i as

$$\mathrm{StdDev}(a_i) = \sqrt{\frac{1}{N_r - 1} \sum_{j:\mathrm{dist}(i,j) \leq r} (a_j - \bar{a}_{ir})^2}$$

where N_r is the number of cells within r cell units; and $\bar{a}_{ir}$ is the average land quality wihthin a radius of r. As usual, we use a default radius of $r = 10$.

Potential Vegetation and Derivates

SAGE Potential Vegetation. Global potential vegetation data is from the Center for Sustainability and the Global Environment (SAGE) in the Nelson Institute at the University of Wisconsin-Madison. The data consists of a global map of natural vegetation at a 5 min resolution classified into 15 vegetation types. These are: tropical evergreen forest / woodland, tropical deciduous forest / woodland, temperate broadleaf evergreen forest / woodland, temperate needleleaf evergreen forest / woodland, temperate deciduous forest / woodland, boreal evergreen forest / woodland, boreal deciduous forest / woodland, evergreen / deciduous mixed forest / woodland, savanna, grassland / steppe, dense shrubland, open shrubland, tundra, desert, and polar desert / rock / ice. The data is representative of the world's "potential" vegetation (i.e., vegetation that would most likely exist now in the absence of human activities). See Ramankutty and Foley (1999) for further details. The data is available at https://nelson.wisc.edu/sage/data-and-models/global-potential-vegetation/index.php.

The SAGE data is also being used in section 2.5 to identify desert cells. Note that given that it is based on potential vegetation cells, the desert measure should be unaffected by human activity that may have led to desertification.

Ecological Diversity and Polarization Indices. As described in Fenske (2014), the ecological diversity index is the Herfindahl index constructed from shares s^t of various ecological types $t = 1, ..., T$. We use SAGE's potential vegetation as the ecological types. In the case of our grid cell data set, we compute for each cell i:

$$\text{EcoDiv}_{ir} = 1 - \sum_{t=1}^{T} \left(s_{ir}^{t}\right)^2$$

where

$$s_{ir}^{t} = \frac{1}{N_r} \sum_{j:\text{dist}(i,j) \leq r} n_j^t$$

with n_i^t being an indicator for cell i having ecological type t and N_r as the sum of cells within r as before. As usual, we use $r = 10$ as our default radius. Similarly, ecological polarization is

$$\text{EcoPol}_{ir} = 1 - \sum_{t=1}^{T} \left(1 - 2s_{ir}^{t}\right)^2 s_{ir}^{t}.$$

Mean and Standard Deviation Temperature

Data on the annual mean and standard deviation temperature are from the World-Clim global climate database. We use current condition (1960-1990) 5 minute resolution data for mean monthly temperature. This is averaged over the year to get annual mean temperature for each cell. The standard deviation is taken across months for each grid cell. See http://www.worldclim.org/ for the data source.

Ruggedness and Slope

The Terrain Ruggedness Index was developed in Riley et al. (1999). The data on ruggedness and slope used here the grid-cell level data prepared for Nunn and Puga (2012). The data is described in detail on their data repository at `http://diegopuga.org/data/rugged/`.

River Flow Accumulation

River flow accumulation data is from the Hydrological data and maps based on SHuttle Elevation Derivatives at multiple Scales (HydroSHEDS) project, which offers hydrographic information in a consistent and comprehensive format on a global scale. River flow accumulation is derived primarily from elevation data of the Shuttle Radar Topography Mission (SRTM) at 3 arc-second resolution and measures the number of cells of drainage accumulation due to the elevation data. See http://www.hydrosheds.org/.

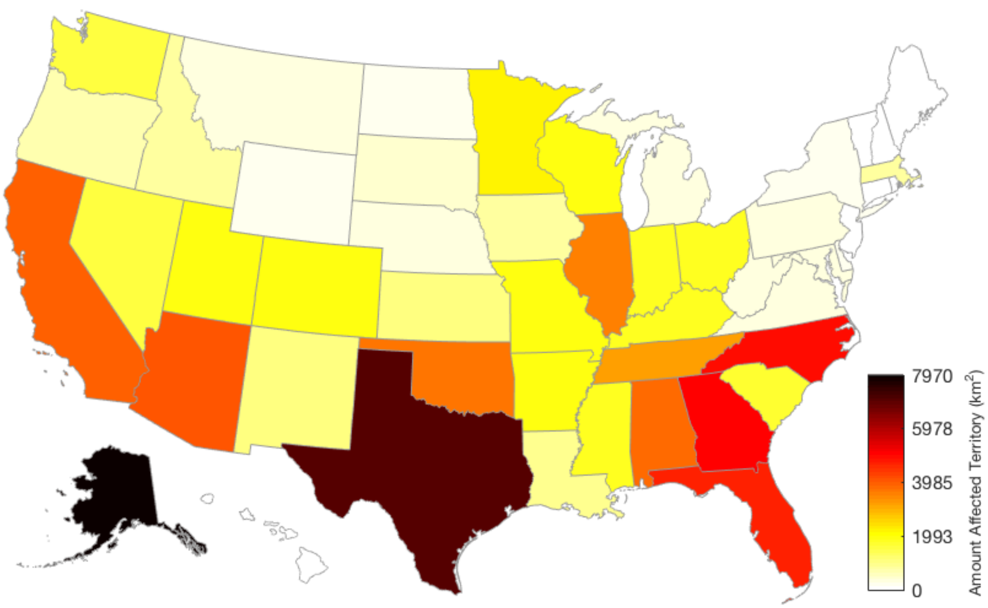
Table 4.2: Robustness using Probit, similar to table 2.3.

	(1)	(2)	(3)	(4)	(5)	(6)
Circumscription: $\hat{\beta}$	0.160		0.126	0.483	0.472	0.442
	(0.026)		(0.046)	(0.132)	(0.145)	(0.138)
	[0.030]		[0.064]	[0.240]	[0.255]	[0.239]
Land quality: $\hat{\gamma}$		0.142	0.041	-0.534	-0.526	-0.475
		(0.019)	(0.035)	(0.187)	(0.206)	(0.195)
		[0.027]	[0.061]	[0.346]	[0.368]	[0.346]
Mean dependent variable	0.00033	0.00033	0.00033	0.00033	0.00033	0.00033
Ecological diversity controls				X	X	X
Climate controls					X	X
Topog. and environm. controls						X
N	184,523	184,523	184,523	184,523	184,523	177,299
Adjusted R^2						

Note: Probit model. Dependent variable: indicator for early state site in cell. Circumscription and land quality are in standard deviation units. Robust standard errors in parentheses; standard errors clustered by $5° \times 5°$ virtual countries (712 clusters) in brackets. Significance stars omitted. Ecological diversity controls includes the standard deviation of land quality, the ecological diversity index and the ecological polarization index; climate controls includes annual mean and standard deviation temperature and absolute latitude; topographic and environmental controls includes slope, ruggedness and river flow accumulation.

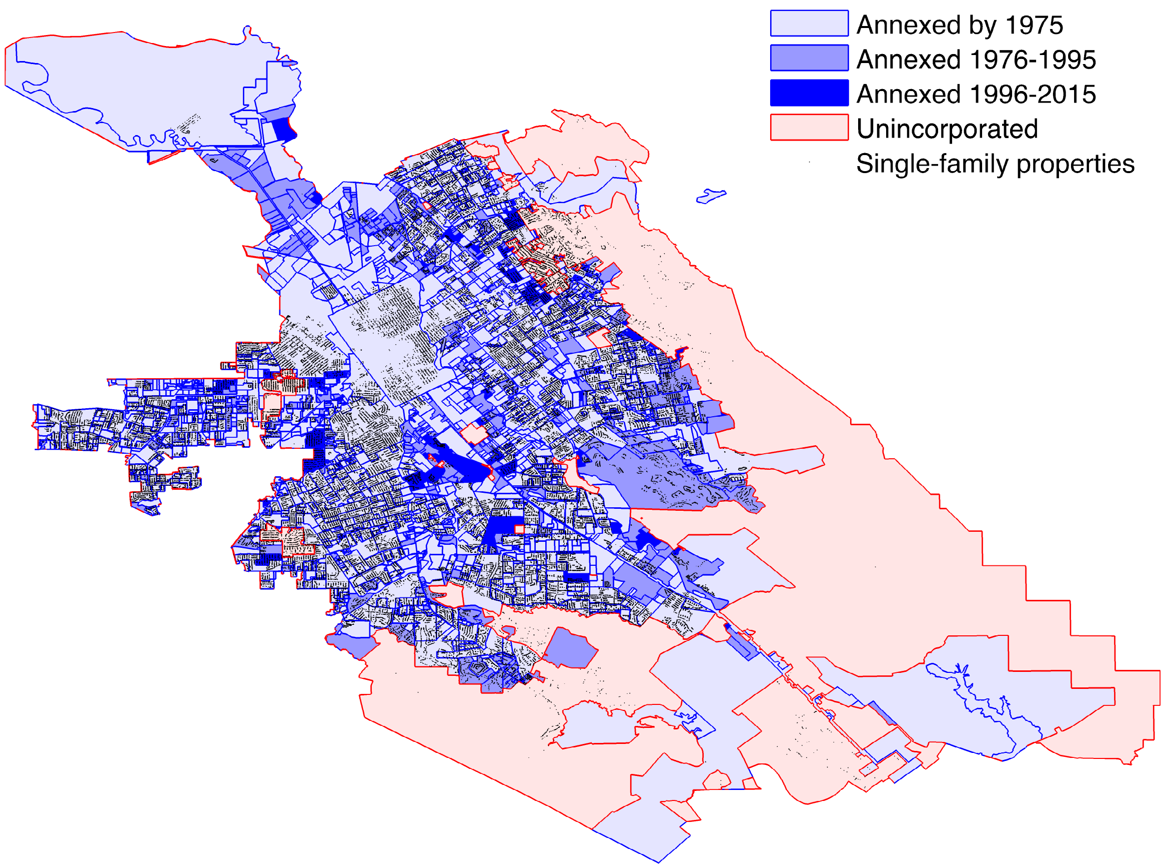
Further Tables and Figures

Figure 4.1: Annexed territory by state.



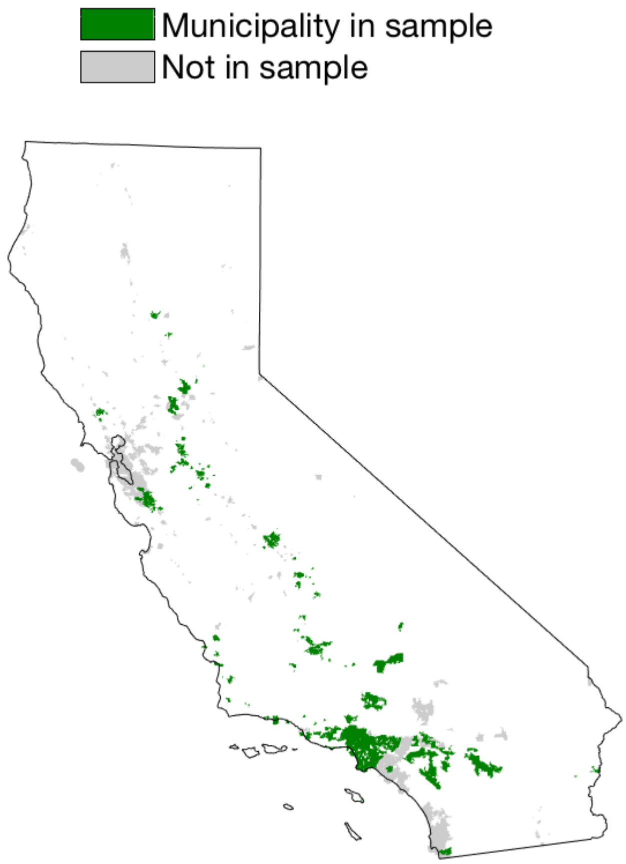
Note: Amount of territory added to municipalities (i.e. annexed) over the period 1990-2010 by state. These annexations included both sparsely built up areas as well as densely populated areas.

Figure 4.2: Example of municipal annexations: San Jose.



Note: Map of San Jose with areas (neighborhoods) and their respective annexation date. Shades of blue denote the era in which an area was annexed: from light blue – before 1975 – to dark blue – after 1996; red are areas that continue to be unincorporated. Black dots are the locations of observed property sales.

Figure 4.3: Spheres included in dataset.



Note: This map shows the 189 municipalities in California for which we have the complete administrative boundary change history. In gray are municipalities for which no boundary change data is available.

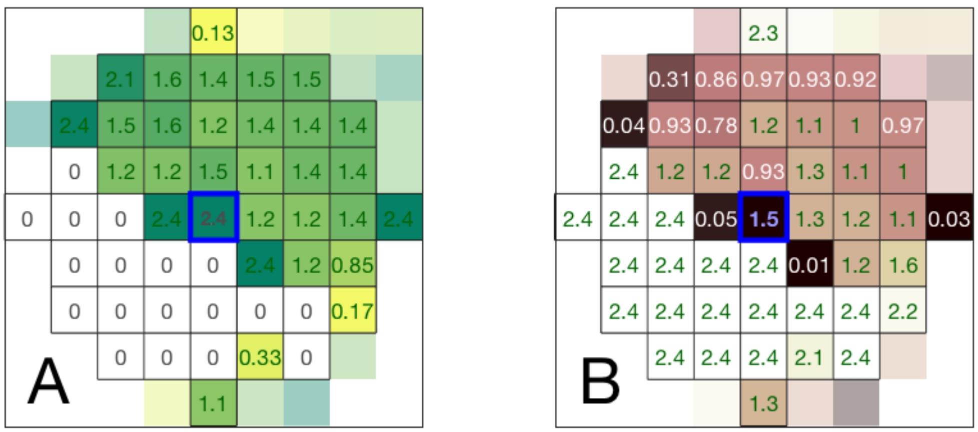


Figure 4.4: Example map computing circumscription.

Panel A shows land quality in units of standard deviation using data from the Nile delta. How circumscribed is the cell with the blue frame? In panel B, we subtract the average value of all surrounding cells (0.9) from the value of the framed cell (2.4) to arrive at 1.5. In other words, moving away from the framed cell into a cell within 5 cell units results in land quality that is on average 1.5 standard deviation units lower.

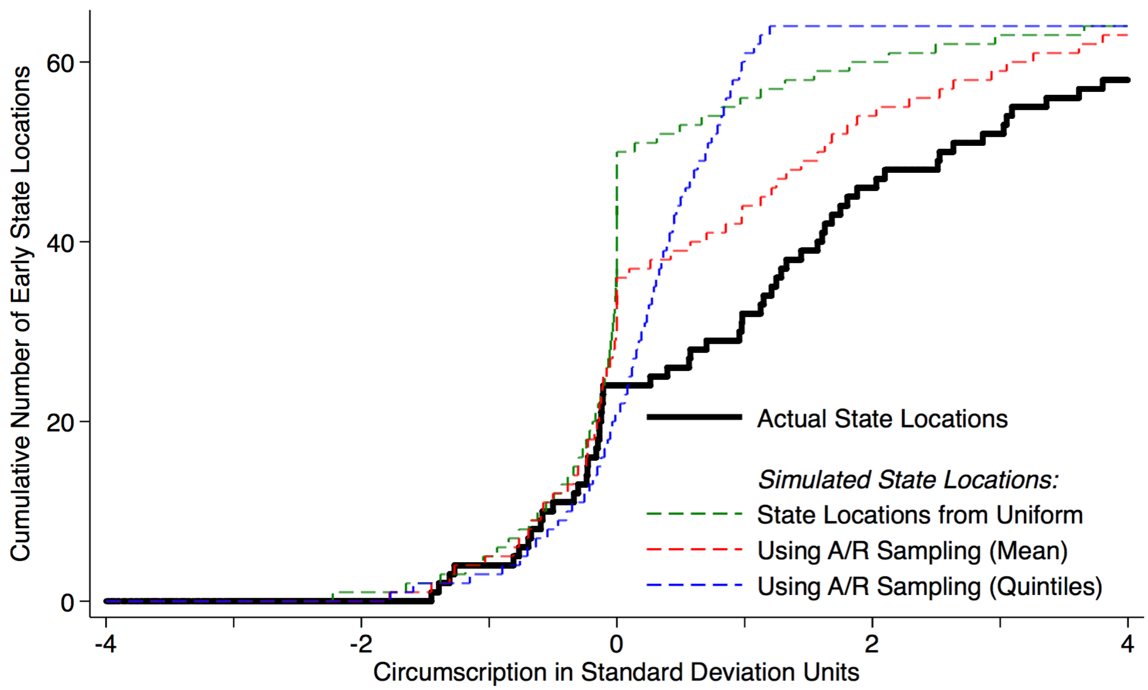


Figure 4.5: Empirical CDFs of actual and simulated state locations.

Empirical CDFs for actual state locations and simulated state locations using 1,000 iterations. Actual state locations in black (five sites are beyond the right edge of the figure). Simulated state locations are distributed according to three different assumptions. In green, state locations are uniformly drawn from all cells; in red, accept/reject sampling to match mean land quality of state cells (i.e. the sample is drawn from cells with the same mean land quality as state cells); in blue, accept/reject sampling to match the quintiles in land quality distribution of state cells.

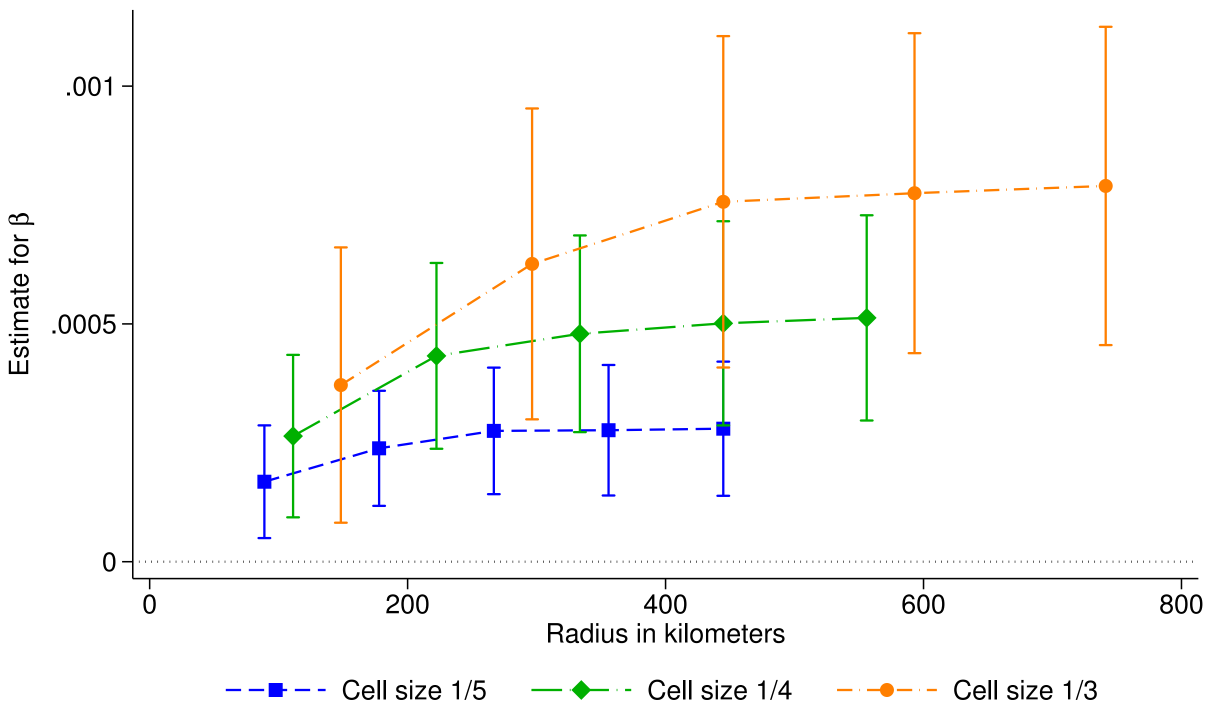


Figure 4.6: Coefficients for various cell sizes and radii, alternative specification.

Estimates for β from equation (2.3) and specification (7) from table 2.3 for various radii and cell sizes. Cell sizes are in degrees; radii are $r = 4, 8, 12, 16, 20$.

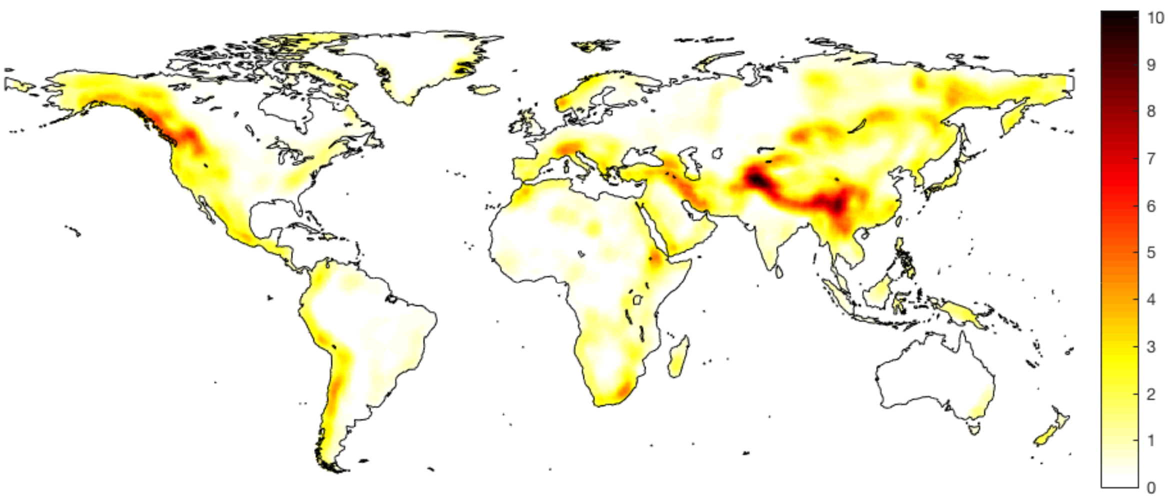


Figure 4.7: Map of surrounding ruggedness.

Surrounding ruggedness in standard deviation units. Standard radius of $r = 10$.

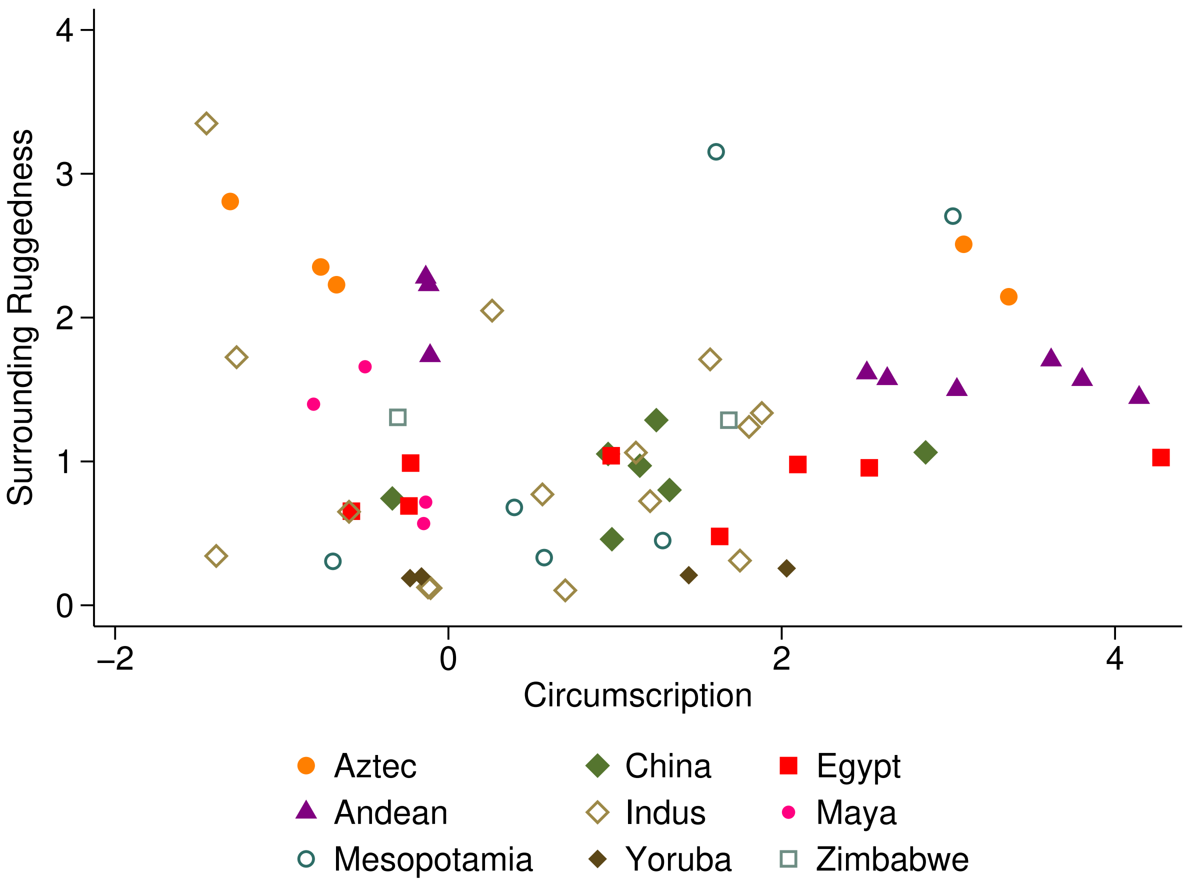


Figure 4.8: The extent of ruggedness and circumscription for all state sites.