



Munich Personal RePEc Archive

Taxes, National Identity, and Nation Building: Evidence from France

Johnson, Noel

11 April 2015

Online at <https://mpra.ub.uni-muenchen.de/63598/>
MPRA Paper No. 63598, posted 12 Apr 2015 06:26 UTC

Taxes, National Identity, and Nation Building: Evidence from France*

Noel D. Johnson[†]
George Mason University

This Version: April 11, 2015

Abstract

What is the relationship between state capacity, national identity, and economic development? This paper argues that increases in state capacity can lower the collective action costs associated with political and economic exchange by encouraging the formation of a common identity. This hypothesis is tested by exploiting the fact that the French Monarchy was more successful in substituting its fiscal and legal institutions for those of the medieval seigneurial regime within an area of the country known as the Cinq Grosses Fermes (CGF). Highly disaggregated data on regional self-identification from the 1789 Cahiers de Doléances confirm that regions just inside the CGF were more likely than regions just outside the CGF to identify themselves with national, as opposed to local, institutions. We also show that regions inside the CGF that affiliated with national identity were more economically developed during the first half of the nineteenth century and more likely to contribute towards local public goods.

Key words: Culture; Institutions; State Capacity; Economic Development

JEL classification: D03; N43; O43

*I am grateful to Saumitra Jha for bringing the data on the Cahiers de Doléances to my attention. John Markoff also very graciously offered assistance in locating data. Werner Troesken provided vital support at the beginning of the project. David Le Bris read and commented on an early draft as well as kindly providing data used in Sections 3 and 6. Nico Voigtländer and Mara Squicciarini provided the data on education drawn from Furet and Ozouf (1977). I also benefited greatly from conversations with Sascha Becker, Jean-Laurent Rosenthal, Nico Voigtländer, Saumitra Jha, Mauricio Drelichman, Jenny Guardado, Philip Hoffman, Mark Koyama, Scott Abramson, James Fenske, John V.C. Nye, Remi Jedwab, Ann Carlos, Sumner La Croix, David Mitch, Karla Hoff, Tony Gil, David Levy, Catalina Viscarra, Richard Hornbeck, Dietrich Vollrath, Nancy Qian, Alan Dye, David Weiman, Suresh Naidu, Arthur Blouin, participants in the January 2014 Economic History Mini-Conference sponsored by the Mercatus Center, audience members at the 2013 SSHA Conference in Chicago, participants in the Public Choice Seminar at George Mason, seminar participants at the University of Vermont, participants at the 2014 Economic History Association Meetings, audience members at ISNIE 2014, participant at the Association for the Study of Religion, Economics, and Culture 2015, and participants in the Barnard/Columbia University Economic History Seminar. Jessi Troyan provided able research assistance. Joy Suh at the George Mason GIS department graciously provided support for the spatial analysis. All remaining errors are the fault of the Author. This paper previously circulated under the title, "From State Capacity to Rule of Law in Old Regime France".

[†]njohnsoL@gmu.edu. Center for Study of Public Choice, Carow Hall, George Mason University, VA 22030.

It is only government by a single man that in the long run irons out diversities and makes each member of a nation indifferent to his neighbor's lot.

Alexis de Tocqueville, *The Old Regime and the French Revolution*, Book II, ch. 8

1 INTRODUCTION: CAPACITY, CONSTRAINT, AND COOPERATION

The relationship between political development and economic development is one of the most active areas of research in social science today. There are two main approaches to this question. One focuses on state capacity – the ability of government to extract revenue (fiscal capacity) and enforce rules (legal capacity).¹ The other approach focuses on the establishment of rule of law, which is commonly associated with credible constraints on government and the uniform application of rules across all segments of society.² At the risk of over-simplifying, the proponents of state capacity claim that any rules consistently enforced on a subject population (excluding the government itself), are conducive to growth. The rule of law approach, on the other hand, argues that ‘good’ rules in the sense that they are inclusive (Acemoglu and Robinson (2006)) and lead to credibly enforced contracts between any group in society are what matters most.

This paper provides evidence that increases in state capacity can facilitate economic development through the creation of shared norms that facilitate the development of growth-enhancing formal and informal institutions. We will show that regions with higher state capacity at the end of the eighteenth century in France – before the constitutional reforms of the Revolution were introduced – were more likely to identify with national, or general, interests as opposed to local, or particularistic, interests. We will also show that in these same areas economic development was higher during the first half of the nineteenth century. Finally, we will provide evidence that within high state capacity regions this economic development was preceded by the formation of shared norms in favor of the establishment of uniform institutions such as common weights and measures and laws applied uniformly to all groups.

Following the work of Besley and Persson (2011), economists tend to treat state capacity and rule of law as being simultaneously determined. This is consistent with observations of present day countries in which rule of law and state capacity invariably ‘cluster’ together. It is very difficult to identify a country with strong rule of law, but no ability to collect taxes or enforce legal codes. Our argument, however, suggests that it is important to unbundle the concepts of state capacity and rule of law since there is evidence that state capacity may lay the ground-work for the formal establishment of rule of law through generating shared

¹The literature is vast. See Weber (1968); Tilly (1990); Besley and Persson (2011); Dincecco (2009); Johnson and Koyama (2013, 2014*b,a*) and the citations therein.

²This literature is also vast. See North and Weingast (1989); Rodrik et al. (2004); Acemoglu et al. (2005) and the citations therein.

norms of cooperation.

The fundamental identification problem this paper addresses is that there may be unobservable variables correlated with both state capacity and national identity. For example, linguistic fractionalization or economic development. To minimize these potential sources of bias, two rich historical episodes will be exploited. The first is the creation of the internal free trade zone known as the Cinq Grosses Fermes (CGF) in 1664 by Louis XIV's finance minister, Jean-Baptiste Colbert. Within the area of the CGF, the absolute monarchy achieved more success in suppressing the local privileges that were inherited from the feudal regime than it did outside the CGF boundaries (Heckscher, 1955, 103-106). More importantly, the monarchy was much more successful in imposing its centralized fiscal institutions on the populations within the CGF. This forced individuals within the CGF to interact with, and eventually identify with, a centralized French state. Instead of (or in addition to) paying local landlords taxes and dealing with seigneurial justice, merchants, nobles, and clergy within the CGF were increasingly subjected to royal, or 'French', tax bureaucracies and justice systems.

In order to identify the effect of increased state capacity on social identity, we use the highly disaggregated information contained within the Cahiers de Doléances sent to the Estates General in 1789. Confronted with an intractable fiscal situation in 1788, Louis XVI agreed to the calling of the French representative assembly known as the Estates General for the first time since 1614. In anticipation of the debates which would emerge from the meeting, every baillage (town or village) in France was asked to compile a list of grievances. Representatives of each Estate (first = clergy, second = nobility, and third = everyone else) in each town were given the opportunity to submit one of these documents.³ These baillage-level Cahiers were sent to the capital of the electoral district in which they were located. Then, each of these approximately 200 electoral districts summarized the baillage Cahiers for each estate into a 'General' Cahier. It is these General Cahiers that we use to measure regional identification with the king or the French state. This task is made possible because of the work of Beatrice Hyslop who, in the 1930's, undertook the massive endeavor to read and analyze all the General Cahiers with an eye toward the question of what types of 'nationalism' existed across France on the eve of the Revolution (Hyslop, 1934, 1936). Importantly for the current study, Hyslop was unconcerned with the particular research question of this paper - 'Whether greater national identification was associated with being located in the Cinq Grosses Fermes?' Indeed, she hardly mentions fiscal boundaries in her work and draws no definite links between a region's membership in the CGF and any of the forty-nine variables she codes for each of the Cahiers.

These two historical sources - the geographic and fiscal reality of the CGF region along with the data on national identity provided by the Cahiers de Doléances - allow us to show a robust association between state capacity and national identity by comparing the reported national identity of similar groups 'close' to both sides of the CGF border. This focus on the local treatment effect of the CGF border results in an estimate of the effect of state capacity on national identity that is less subject to bias due to unobservables than simply

³In general, the third estate represented urban interests rather than peasant, agricultural, interests.

comparing groups inside the CGF to those outside (the average treatment effect). This is because CGF and non-CGF regions close to the border are more likely to have similar characteristics than regions more distant from the border. We will confirm this is the case by comparing the values of observables on both sides of the border in Section 4. In particular, we are confident that the association we show between fiscal capacity and national identity does not stem from geographic factors or differences in initial political institutions – such as whether the region used civil or customary law – which we explicitly control for. We also show that migration is unlikely to explain this correlation using data on urban populations from Bairoch (1988). Finally, we show that linguistic diversity – as proxied by the literacy of regional populations from Furet and Ozouf (1977) – also does not explain the association between fiscal capacity and the development of national identity. We conclude that the development of state capacity in Early-Modern France played a robust role in fostering norms in favor of national, as opposed to local, institutions before the Revolution.

In addition to the literatures on state capacity and rule of law mentioned above, this paper touches on several other topics of concern to social scientists. It follows in the tradition of a recent microeconomics literature which uses within country variation to measure the effect of institutions on economic development (e.g. Banerjee and Iyer (2005), Iyer (2010), and Dell (2010)).⁴ In particular, we build on work by those such as Michalopoulos and Papaioannou (2013) and Dincecco (2009) who argue that historically strong states helped generate economic development.⁵ Unlike these authors, however, we are able to show an explicit link between state capacity and nascent support for national institutions – thereby providing a plausible mechanism through which early state building led to economic development. We also provide empirical evidence which contributes to the debate on whether formal state institutions crowd-in or crowd-out informal norms of cooperation. For example, Bowles and Polania-Reyes (2012) and Lowes et al. (2015) both find crowding-out. Tabellini (2010) and Becker et al. (2014), by contrast, find crowding-in. The evidence we present on the stated preferences for uniformly shared institutions in the Cahiers as well as our findings concerning local public goods spending is consistent with the crowding-in argument.

This paper also contributes to the historical literature on the origins of the French Revolution, and more specifically, French Revolutionary nationalism, at the end of the Old Regime. Many of these studies focus on questions related to the present work such as linguistic or religious affiliation and Revolutionary sentiment (de Certeau et al., 2002; Bell, 1995).⁶ Historians such as Hyslop (1934) and Shapiro et al. (1998) did tremendous work compiling and studying the contents of the Cahiers de Doléances, often with an analytical eye towards the origins of French Revolutionary nationalistic sentiment. To the extent that they relate national identification with what we term state capacity here, however, they focus on average treatment effects in their formal analysis. For example, Markoff (1998) looks at the correlation between a region being a *pays d'état*

⁴Recent papers focusing specifically on institutions associated with political and economic borders include Pinkovskiy (2013), Schulze and Wolf (2009), and Ploeckl (2013).

⁵A closely related argument is that of Bockstette et al. (2002) who argue that older states are associated with greater economic development.

⁶More generally, it is difficult to over-state the influence of Anderson (1991) on the study by national sentiment within the humanities and its resultant emphasis on print culture.

(less monarchical control) versus a *pays d'élection* (more monarchical control) and the degree of agreement between First and Second Estates on various issues brought up in the *Cahiers*. This is a valuable analysis, however, the correlations he presents fall short of presenting a robust relationship between state capacity and national identity, which is a prime focus of the present study.⁷ Our analysis is also closely related to studies of internal variation in *ancien regime* institutions and subsequent economic performance. For example, Le Bris (2015) investigates whether a region using customary versus roman law influences early nineteenth century development. Daudin (2010) uses data on product prices to show that at the end of the eighteenth century individuals within the *Cinq Grosses Fermes* tended not to trade with those outside of it. Our results suggest, however, that the political and economic implications of the CGF tariff region went well beyond the dead-weight losses from foregone trades. Social preferences were molded differently inside the CGF than outside it – and to such an extent that economic outcomes were better in the nineteenth century.

More generally, our work supports those who argue that preferences matter for economic development. For example Hirschman (1977), Nunn (2012), and the empirical work of Hoff and Pandey (2014) on caste identity versus national identity in India. Furthermore, given the unique role played by France as an early participant in the Industrial Revolution, our argument complements the work of those who claim preference changes mattered in the transition to modern economic growth in Europe. Specifically, McCloskey (2010) claims there was a ‘bourgeois revolution’ in which individuals adopted attitudes towards markets that were conducive to growth. Similarly, Mokyr (2005) emphasizes the Baconian Revolution as a turning point in ability of European society to generate and implement scientific ideas. Our unique contribution in this paper is to identify an important historical mechanism – increases in state capacity – through which preferences were formed and directed towards broader goals, such as the creation of common institutions governing market exchange.

2 STATE CAPACITY, PREFERENCES, AND COOPERATION

We make two arguments in this paper: (1) Increases in state capacity in Early-Modern France were closely associated with increased preferences for national, as opposed to local, political and economic institutions. (2) Areas which affiliated with a national identity at the end of the eighteenth century achieved higher economic development in the first half of the nineteenth century. In this section we will briefly survey the literatures addressing two questions: What is state capacity? How might increases in state capacity generate preferences which favor formal and informal institutions that are national rather than local?

In ‘Politics as a Vocation’ Max Weber defines states as political communities that ‘...claim a monopoly on the legitimate use of physical force within a given territory’ (Weber, 1946, 1918). More recently, researchers have operationalized Weber’s definition as consisting of two, measurable, characteristics: (1) The ability to extract revenue, or fiscal capacity, and (2) the ability to enforce rules, or legal capacity. Fiscal and legal

⁷To reiterate, however, the collected articles in Shapiro et al. (1998) and in Markoff (2010) are in broad agreement with the present, general thesis, that stronger states are conducive to the creation of individuals with less particularistic interests.

capacity tend to be positively correlated with each other or, to use the terminology of Besley and Persson (2011), they ‘cluster’. It is very rare to observe a polity with either high taxes and an inability to enforce rules or low taxes and an effective legal environment. As such, and given the difficulties of measuring legal capacity, most historical studies focus on just fiscal capacity in order to trace the development of the state over time (e.g. Dincecco (2009)). In Europe, these studies show that around the end of the sixteenth century, state capacity was dramatically increasing in places like England, France, the Low Countries, and Prussia.⁸

But what effect did these increases in capacity have? Aside from imposing some minimal rules that probably facilitated trade and eventually allowed for greater investment in education (see, e.g., Epstein (2000), Dincecco (2009), and Johnson and Koyama (2014*b*)), is it possible that individuals living in high state capacity regions actually had their preferences molded so that they were more likely to identify themselves with the newly powerful governments which were taxing them? Historians and sociologists have been aware of this possibility for some time. Most famously, one of the main themes in Tocqueville (1998) is that the French Revolution would have never been possible if the institutions of the absolute monarchy had not been so successful in undermining the feudal regime during the eighteenth century.⁹

In economics, one of the more influential arguments concerning the importance of preferences for economic development is Hirschman (1977). For him, the increasing importance of markets, or ‘doux commerce’, during the seventeenth and eighteenth centuries had a profound effect on what individuals valued. The shift was from treating money-making as shameful and personal honor as paramount, towards a more ‘modern’ outlook which valued individualism and market institutions. More recently, McCloskey (2010) has pursued the argument that it was a change in beliefs (as manifested in the rhetoric of individuals) which made possible the Great Divergence in incomes of the late eighteenth century and signaled the onset of modern economic growth.

For Hirschman (1977) the driving force behind preference changes was trade. For McCloskey (2010) it was a combination of the Reformation, the printing press, and the revolutions of the late early-modern period. Recently, however, Tocqueville’s view that strong states may play a significant role in forming individual preferences has received support from work in behavioral economics. In particular, experimental work by Fehr et al. (2008) and Hoff and Pandey (2014) suggests that the values people affiliate with can be strongly affected by the institutions which surround them.¹⁰ Additionally, historical work by Becker et al. (2014) and Lowes et al. (2015) explicitly investigates how the strength of formal political institutions can have a long-run

⁸For the overall picture after 1600, see Dincecco (2009), the essays in Bonney (1995), and Karaman and Pamuk (2013). For a classic treatment of England, see Brewer (1988). For a more recent attempt at endogenizing the explanation for investment in capacity in England and France, see Johnson and Koyama (2014*a*).

⁹The role of state power in affecting individual beliefs is a common theme of historians, especially since the work of Foucault (1977) which influenced many subsequent studies (for example, Sabean (1987) or Rebel (1983)). In sociology, Swidler (1986) argues that the ‘ascriptive’ identity of individuals can be influenced by external institutions and can shift relatively quickly during periods of instability. This theme is built upon by Ross and Nisbett (1991) who discuss the work in social psychology supporting the idea that the social identity of individuals is strongly influenced by their environment.

¹⁰For a recent survey of this work in behavioral economics, see Fehr and Hoff (2011).

impact on informal norms of cooperative behavior.

There are also theoretical studies which support the idea that a high capacity state can mold the preferences of individuals to favor norms of cooperation. In particular, Acemoglu and Jackson (2014) argue that in societies with forward looking individuals it is possible for a ‘prominent agent’ to reverse a history of non-cooperative play through it’s actions. While most of the motivating examples provided by Acemoglu and Jackson (2014) concern individuals (e.g. Mandela in South Africa or Gandhi in India) one could also interpret their prominent agent as a high capacity state. Another line of research by Gintis (2014, 2010) develops a theoretical framework to explain the emergence of cooperation based on the concept of correlated equilibrium Aumann (1987). In this setting a ‘choreographer’ signals to agents how to play in a coordination game. Importantly, the choreographer can be interpreted as either an informal institution (e.g. norms concerning generalized trust) or a formal institution (e.g. regulations concerning weights and measures, taxation, or education). In Appendix A we present an outline of this theory based on Gintis (2014) to show explicitly how cooperation in a coordination game can be increased by the presence of a state acting as a choreographer of a correlated equilibrium.

It is, therefore, possible that the changes in preferences highlighted by Hirschman (1977), McCloskey (2010) and others that laid the foundations for modern economic growth may have been largely a by-product of the increases in state capacity that occurred between 1500 and 1800 in Europe. In particular, these changes were characterized by a shift in social preferences such that individuals affiliated less with particularistic and local institutions. Instead, they started thinking of themselves as citizens in a shared national culture. One consequence of this was that they also became more concerned with dismantling local institutions in favor of the new national institutions. This manifested itself as a desire to standardize economic and legal institutions and to eliminate the vestiges of the feudal regime such as serfdom (we will investigate these claims specifically in Section 6.1).

The evidence for the importance of preferences in generating positive political and economic outcomes is extensive. There is also evidence that changing preferences played an important role in what is probably the most important event in modern economic history, the Great Divergence. What is less clear is what caused this change in preferences. In what follows below, we will present evidence from Old Regime France that suggests Tocqueville was right when he asserted that it was state capacity (the strong absolute monarchy) that ‘ironed out the diversities’ of the French people and made possible the political and economic changes of the Revolution and beyond.

3 HISTORICAL BACKGROUND AND DATA

3.1 *The Cinq Grosses Fermes and French Fiscal Capacity Before and After 1664*

Before Louis XIV took power in 1661, the French fiscal system was characterized by a wildly confusing amalgam of both direct and indirect taxes. The indirect taxes were mainly the octrois, péages, and various adjunct taxes to the aides, traites, and gabelles. The octrois were primarily taxes on collected at city gates and included the huge fiscal resource of the entrées de paris. The aides were mainly excises on wine and other spirits. These and the droits de marque on iron, oil, soap, paper, cloth, and leather, were often collected as border tariffs in the regions in which they were imposed. Also, movement of salt or tobacco was often taxed on border of the gabelles and tabac regions. The main direct taxes were the taille, the capitation (permanent from 1704), and the first and second vingtièmes (from 1749 and 1760 respectively) (see Heckscher (1955, 78-93) and Bosher (1964, 1-5)). These latter taxes, at least in theory, were supposed to be surtaxes of 5% each on the income of all Frenchmen. In practice, however, they were resisted and manipulated with varying levels of success by provincial and class interests at all levels.¹¹

Onto this scene Louis XIV's finance minister, Jean-Baptiste Colbert, arrived with ambitious ideas for reform. Colbert was mercantilist and wanted primarily to adjust border taxes so as to favor the export of French finished goods and the import of raw materials from abroad. This implied high export taxes at the border on raw materials and low import taxes. Similarly, it implied low export taxes on finished goods, but high import taxes on these items. This plan for a customs union with unified export and import taxes on the border was described by Colbert in a document of 18 September, 1664. He succeeded in bringing about half the country into this unified customs union and placed the management of taxes in it under a group of tax farmers known as the Cinq Grosses Fermes.¹² The end result were two separate tax regions: The Cinq Grosses Fermes in the center and The Provinces Reputed Foreign in the East, South, and West (See Bosher (1964, 5-9) and Heckscher (1955, 96-102)).¹³

The proposed unified customs union was not accepted by all the provinces, but was instead centered on the

¹¹A telling anecdote is that when the second vingtième was introduced there was an attempt at the same time to introduce a general land survey by the monarchy's Finance minister Henri Léonard Jean Baptiste Bertin. This was fiercely resisted by the parlements (legislators) across the country since they didn't want the government to know how much they had (and could therefore tax).

¹²For more background on tax farming see Johnson (2006), Balla and Johnson (2009), and the citations therein. The original Cinq Grosses Fermes were the (1) traite foraine, le rêve et le haut passage de Champagne et de Normandie, (2) la traite foraine de Normandie, (3) le traite domaniale de Champagne, Picardie, Normandie, et Bourgogne, (4) la douane de Lyon, and (5) les droits d'entrée sur l'épicerie, drogueries et grosses denrées. These were united under the ownership of René Brunet in 1589. The process was pushed through during the high period of the wars of religion by a cash strapped Henri iii (r. 1574-1589) (Roux, 1916, 70-73). The reasons for unification were, in all likelihood, as more due to a desire to increase the amount of lending the farms could support than to gain administrative efficiencies (Johnson and Koyama, 2014a).

¹³The Provinces Reputed Foreign were Angoumois, Artois, Auvergne, lower Navarre, Béarn, Brittany, Cambrésis, Foix, Dauphiné, Flanders, Forez, Franch-Comté, Gascony, Guyenne, Hainaut, Ile-de-Rhé, Ile d'Oléron, Languedoc, Limousin, Lyonnais (in parts, see below), Marche, Provence, Roussillon, Rouergue, Saintonge, and Vivarais.

traditional lands of the Bourbon monarchs (these will be described in detail in section 3.2 below).¹⁴ As such, the CGF border was not necessarily random and this is a threat to our identification strategy. One piece of evidence suggesting that this is not an important threat to our results is that Hyslop finds no evidence for the nationalist sentiment she observes in the General Cahiers of 1789 in the Cahiers compiled for the Estates General of 1614 – about fifty years before the creation of the CGF border (Hyslop, 1934, 59-63). Nonetheless, we will deal with the potential endogeneity of the border in several ways. First, we will include a host of control variables covering geographic, political, and cultural factors. Second, we will focus on the local treatment effect and will show that as we compare regions closer to the CGF boundary, differences on observables disappear. This suggests, that regions close to each other but on opposite sides of the CGF boundary were very similar. Third, in section 3 we will attempt to test for differences in linguistic background and for potential selection due to migration.

The creation of the CGF unleashed a process which plausibly had a profound impact on the desire of French citizens for reform as well as on their likelihood to identify themselves in relation to either the institutions of the monarchy or those of a more local character. In particular, Boshier (1964, 13-15) says of the Company of General Farms, ‘The methods of collection had been reformed [by Colbert] without reforming the system of duties.’ As this quote suggests, as a result of the creation of the CGF, two processes unfolded: (1) There was suppression of local, or seigneurial, control of fiscal instruments, especially within the area of the CGF (Heckscher, 1955, 103-106).¹⁵ (2) There was a dramatic increase in the amount of exactions, especially within the area of the CGF (we will provide extensive evidence for this in section 4 below).

To see how these two factors conspired to affect opinion about economic reform and national affiliation, consider that the creation of the internal tariff barriers raised the cost of commerce significantly – and that the more barriers one had to cross, the greater this distortion. Thus, for example, in 1702 Jean Anisson de Hauteroche, deputy of Commerce for Lyon, in writing up a reform project for internal customs emphasized that they raised the price of French goods by between 6 and 15 percent (Boshier, 1964, 32).

Furthermore, these distortions had different effects on citizens depending on whether they lived inside or outside the CGF. Inside the CGF the tax system was particularly baroque and oppressive. This meant that individuals received a high return for learning how to work within the system. Thus, members of the nobility often focused on confirming their noble heritage (as ratified through the institutions of the monarchy) so as to retain their exemptions and privileges (Conchon, 2002). Likewise, members of the third estate, who were often represented by bourgeois merchants, had to learn to work within the tax and legal system to survive. This also meant individuals inside the CGF were more focused on a discussion of reform of monarchical

¹⁴An important exception to this observation is Burgundy. It was in the CGF but also had a strong history of independence as the Kingdom of Burgundy. Indeed, the civil conflict known as the *Fronde* (1648-53) was led by princes from Burgundy who were attempting to overthrow the Bourbons.

¹⁵Thus, for example, in 1724 the royal commission on péages started investigating who had legitimate claim and to confiscate those that didn't. The commission claimed to have reduced the number of tolls from 5,688 to 2,054 during the reign of Louis XV. By 1789 there were supposed to be only 1,600 (Boshier, 1964, 2).

institutions. Outside the CGF, by contrast, trade and life in general was more oriented away from the center of the country and more likely to be with foreigners. Thus, the eastern region of Metz, for example, was opposed to tax reform at the end of the eighteenth century (‘the single tariff project’) because they were content with focusing on foreign trade Boshier (1964, 135).

3.2 *The Construction of the Cinq Grosses Fermes Boundary*

To determine which Hyslop cities are included within the boundary of the Cinq Grosses Fermes we use the following procedure. As a first cut, we use the entry on ‘Cinq Grosses Fermes’ in le Rond d’Alembert and Diderot (1784) as the source for the delineation of the CGF border. The relevant entry is reproduced in Appendix B, Figure 8. The entry lists all the provinces included within the border of the CGF as described in Colbert’s ordinance of 1664. For our baseline estimates, we also include Lyon since, as explained in Appendix B, Figure 9, Lyon was granted special status with ‘privileged communications’ with the CGF.¹⁶ After establishing the provinces contained in the CGF we then used the shapefiles of French provinces provided by *Euratlas* (2012) to reproduce a first-cut at the boundary.

The CGF boundary did not actually perfectly correspond to the borders of the provinces delineated in the 1664 ordinance. Thus, in the second step, we use two sources to check on membership status of each Hyslop city individually. The first source we use is Saugrain (1720) which was a population census commissioned at the end of the seventeenth century.¹⁷ With the Saugrain (1720) census we confirm whether each Hyslop city is in a CGF province as defined by the 1664 edict. We next consult Expilly (1768) which is an eighteenth century geographic dictionary. It contains entries for most of the Hyslop cities and explains their legal and economic status. Where there is not an explicit entry for the city, we use the relevant entry on the province or the généralité (another fiscal unit) in which the city was located (as recorded in Saugrain (1720)), which in all cases is sufficient to explain whether the city is in or out of the CGF. After confirming the status of the Hyslop cities, we adjust the initial boundary of the CGF border accordingly using GIS software. The resulting border is illustrated in Figure 2 below along with the Hyslop cities.

3.3 *Measuring Fiscal Capacity at the End of the Old Regime*

In order to establish that fiscal capacity was greater in the region of the Cinq Grosses Fermes than in the rest of the country we rely on data on taxes per capita recorded by Jacques Necker, the finance minister of Louis XVI, in 1784. In his *De l’administration des finances de la France*, Necker recorded for almost every tax district the total taxes per capita collected by the monarchy.¹⁸ These include both direct and indirect taxes.¹⁹ The tax district boundaries in 1700 are illustrated in Figure 1. Since our analysis of the Cahiers will be done at the city rather than the province level, we use GIS software to extract the values of taxes p.c.

¹⁶We also repeat the analysis with Lyon excluded in in Section 4, Table 3.

¹⁷The exact origins and methods used to construct the Saugrain census are uncertain. However, the institutional information it contains is much more precise than the actual population estimates (Hussenet, 1996).

¹⁸These data are reproduced in Nordman and Ozouf-Marignier (1989).

¹⁹The tax districts, of which there were about 31 in 1784, were originally created in 1542 by Francis I (Wolfe, 1972).

for each city in each province. To account for the resulting correlation in fiscal capacity across cities, we will cluster standard errors at the province level in all regressions. We call the resulting fiscal capacity measure *Log Taxes per capita*.

We assume the level of taxes per capita at the district level in 1784 is a good proxy for increases in fiscal capacity due to the creation of the CGF in 1664. This is a safe assumption since fiscal capacity was relatively low throughout France before the accession of Louis XIV in 1661. In Appendix C we describe in detail these fiscal changes during the second half of the seventeenth century.

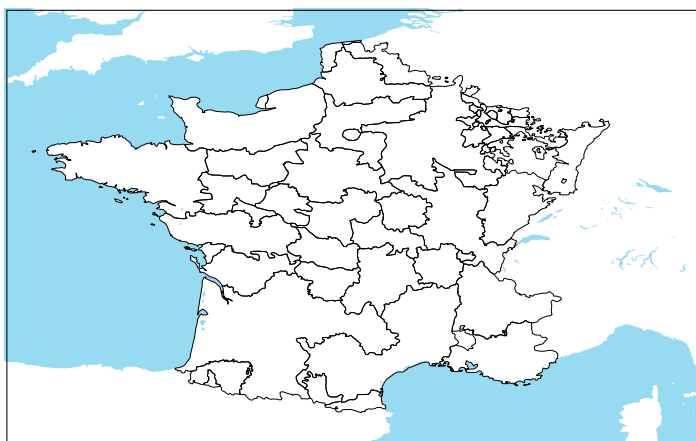


Figure 1: The borders of the provinces in 1700.

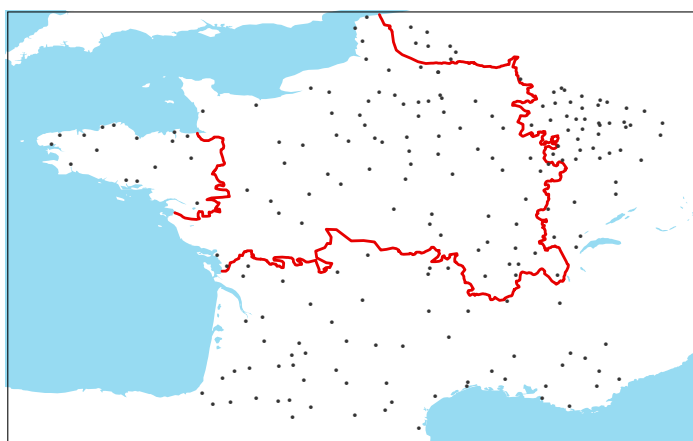


Figure 2: Cities of origin for the General Cahiers and the CGF boundary.

3.4 *The Cahiers de Doléances and Hyslop's Coding*

The locations of the cities in which the General Cahiers were compiled and then sent to Paris are shown in Figure 2. As can be seen from the map, their distribution is not uniform. Particularly in the south-central and western regions of the country the distance between Cahiers observations is further than in the east.

The main outcome variable we use is called *National Identity* and comes from the coding of the General Cahiers by Hyslop (1934) for various degrees of what she terms ‘nationalism’.²⁰ To Hyslop, nationalism was a ‘common consciousness’ that manifested itself on four different margins as shared ‘French’ notions of: (1) geography, (2) race, (3) traditions and, (4) formal institutions (a ‘desire for uniformity’) (Hyslop, 1934, ch 3).

‘Geography’ is interpreted by Hyslop, and coded accordingly, as a recognition of defined national boundaries for an entity called ‘France’. ‘Race’ consists of references in the Cahiers to supposed shared French ‘virtues’, or, more commonly, notions of citizenship. ‘Traditions’ include references to a shared French ‘Constitution’, a common religion, a common language, or the desirability of a national plan for educational instruction. ‘Formal institutions’ (or what Hyslop refers to as a ‘desire for uniformity’) include references to legal and ad-

²⁰Hyslop also published an extensive guide to her analyses as Hyslop (1936).

ministrative uniformity or economic uniformity. This last category, is of particular relevance for the influence of the CGF on national sentiment and, as such, we will investigate the relationship between membership in the CGF and the desire for uniformity in greater detail in Section 6.1.

Hyslop counts the instances in which each of these four characteristics (and their sub-components) of national sentiment appear in each of the General Cahiers. She then constructs a single index out of her sub-indexes in which she records for each Estate whether ‘national patriotism is strongest (to the king or Nation)’, ‘loyalties were mixed’, ‘loyalties to localities, class, or both outweigh national patriotism’, or, ‘no sentiment towards nation or locality is shown’. We code these entries as 3, 2, 1, and ‘missing’ respectively. As such, the main outcome variable, *Identity* runs from 3 to 1. We then take the average of this number over the Cahiers for the nobility and the third estates as the main outcome variable to measure *National Identity*.²¹

4 ANALYSIS

4.1 Data Descriptions and Balance of Observables

The identification strategy is to investigate the discontinuity in national sentiment close to the border of the CGF. This regression discontinuity approach should provide an unbiased estimate of the effect of state capacity on national identification assuming that regions close to the border are relatively similar and if the membership in the CGF results in a ‘treatment’ of living under state institutions with higher fiscal capacity.

As a first cut at seeing if there is some relationship between national sentiment and membership in the CGF, Figure 3 illustrates the spatial distribution of the value of the variable *National Identity*. In the Figure, a heat map of the value of *National Identity* is created based on the inverse weight of it’s value for each of the surrounding 12 cities for each grid on the map. Visual inspection of the map provides support for the hypothesis that regions within the CGF were more likely to identify themselves with France or the king, in 1789. With the notable exception of the Gironde (Bordeaux) and La Rochelle regions in the mid-west and, to a lesser extent, the area around Dijon in the mid-east, the discontinuity at the border in national sentiment appears large and robust.

²¹We exclude the cahiers of the clergy and the relatively small number of ‘unified’ cahiers in which the grievances of all three estates are lumped together. The rationale for excluding the clergy is that their identities were more likely affected by religious policy than tax policy. Including the clergy results in slightly smaller coefficients in the baseline regressions, but our main results are largely supported. These regressions are available upon request.

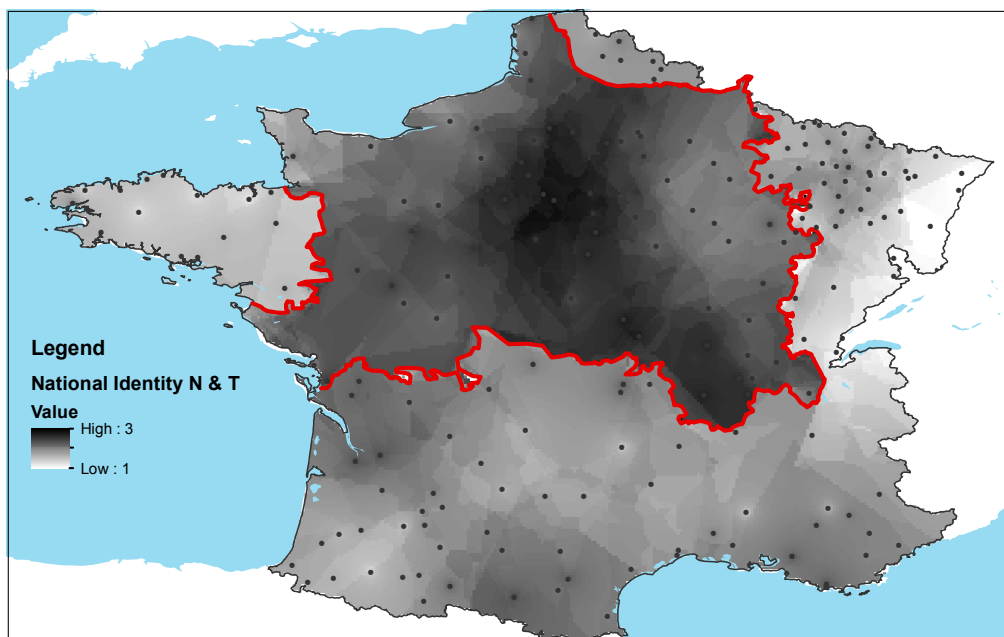


Figure 3: *National Identity* in the 1789 General Cahiers. Darker color represents greater identification in the Cahiers by the Nobility and Third Estate with either the ‘King’ or ‘France’ according to Hyslop (1934). Each grid point in map assigned a value based on the inverse-weighted distance of surrounding 12 cities that sent in Cahiers (exponent of distance used is the default of 2). The Cinq Grosses Fermes region is delineated in red and treated as a barrier.

In order for the identification strategy to be convincing, the characteristics (other than national affiliation) of nearby regions on either side of the CGF border should be similar. Since we can’t test whether potentially important unobservable characteristics differ, in Table 1 we instead investigate whether our identifying assumption holds for the observables which we will also be including as control variables. In columns (1), (2), and (3) we show the average value of each control variable outside the CGF, inside the CGF, and the difference between the two. In brackets under the difference we report the p-value of a two-sided t-test of significance. In columns (4) - (6) we report the statistics on the variables for only the cities within 150 kilometers of the CGF boundary. In columns (7)-(9) we report the statistics for only the cities within 75 kilometers of the boundary. If our identification strategy is appropriate, differences in variables for cities inside and outside the CGF should decrease as we restrict our sample to the area closer to the boundary.

The control variables we investigate are the following (see Appendix F for more details on the control variables): *Urban* measures urban density around the city and is constructed using data on all cities in France with populations greater than 5,000 in 1800 provided in Bosker et al. (2013). *Wheat Suitability* data come from the FAO and are described in Fischer et al. (2002). These data are constructed by combining characteristics of wheat (optimal growing temperature, soil type, etc.) with highly disaggregated climatic and geographic data covering variables such as precipitation, cloud cover, ground-frost frequency, soil types and slope characteristics. The data have a spatial resolution of 0.5 degree x 0.5 degree (or about 60 x 60 kilometers at 45 degrees latitude, typical for France). We extract the wheat suitability for each of our cities using

geospatial software. *Ruggedness* is from Sappington et al. (2007) and measures the variation in the altitude of the terrain surrounding each city. *DRivers* and *DSeas* are simply measures of the distance of each city from major rivers and seas contained in the *Euratlas* (2012) shapefiles. *Communes* is from the Bosker et al. (2013) dataset and measure the average number of cities in the province the Hyslop city is located in that had some form of self-governance. *Bishoprics* and *Archbishoprics* are similarly constructed measures made using the Bosker et al. (2013) data and indicate how close the Hyslop city is to major centers of the Catholic religion. *Universities* is also constructed using the Bosker et al. (2013) data and measures how many cities, on average, had a major university in the province of the Hyslop city. Finally, the two main dependent variables we focus on, *Log Taxes per capita* and *National Identity*, are constructed as described above in sections 3.3 and 3.4.

Column (3) of Table 1 indicates that, using the full sample, four variables - *Urban*, *DSeas*, *Bishoprics*, and *Communes* - are all shown to be significantly different outside the CGF than inside. Furthermore, *Universities* is significant at the 0.104 level and *DRivers* is also very close. When we restrict the sample to cities within 150 kilometers of the CGF boundary, only *Urban* and *Communes* retain their significance. Most of the other differences in the variables become much less significant. When the sample is restricted to cities only 75 kilometers from the boundary, then *Communes* continues to be significant and *DSeas* becomes barely significant at the 10 % level. All the other variables are indistinguishable from zero. This suggests that, at the border of the CGF, at least on observables, regions were fairly similar. Given our relatively small sample of cities (only 208 in the full sample), it is not clear whether 150 km or 75 km is the more appropriate distance to investigate for the local treatment effect. For example, at 75 km there are only 29 cities in the CGF group. Because of this, we will report estimates of the local treatment effect of the CGF boundary for all three samples (full, 150 km, and 75 km) for most regressions. Only when we focus on individual border segments will we restrict ourselves to using the full data set due to sample size considerations.

In the last two rows of Table 1 we show that the two main outcome variables - *Log Taxes per capita* and *National Identity* - are significantly different between cities inside and outside the CGF regardless of how close we approach the border. Furthermore, these are large differences. In the full sample, a standard deviation in *Log Taxes per capita* is 0.39 whereas the CGF cities pay, on average 0.57 more in taxes per person. Similarly, a standard deviation in *National Identity* is about 0.61 whereas the value of the variable is about 0.51 higher in cities inside the CGF boundary. Furthermore, these differences in the outcome variables are stable as we restrict our attention to cities closer to the border. This suggests that the local average treatment effects we estimate below are also relevant more globally.

	Full Sample			<150 km of CGF Boundary			<75km of CGF Boundary		
	Inside CGF (1)	Outside CGF (2)	Difference (3)	Inside CGF (4)	Outside CGF (5)	Difference (6)	Inside CGF (7)	Outside CGF (8)	Difference (9)
Urban	35.24	21.34	-13.90*** [0.000]	25.86	21.25	-4.62** [0.0391]	21.88	21.09	-0.79 [0.626]
Wheat Suitability	3.56	3.65	0.09 [0.609]	3.65	3.52	-0.13 [0.553]	3.62	3.57	-0.06 [0.842]
Ruggedness	0.72	0.69	-0.03 [0.234]	0.73	0.69	-0.04 [0.213]	0.72	0.67	-0.06 [0.234]
DRivers	57555.46	73887.09	16331.63 [0.131]	63189.89	60201.12	-2988.776 [0.804]	67979.8	59884.77	-8095.029 [0.628]
DSeas	198996.9	160707.4	-38289.57** [0.022]	184022.6	202359.4	18336.79 [0.391]	144701.3	194888.5	50187.22* [0.076]
Communes	0.63	0.48	-0.15*** [0.006]	0.65	0.36	-0.29*** [0.000]	0.69	0.40	-0.29*** [0.002]
Bishoprics	0.48	0.31	-0.18*** [0.000]	0.39	0.25	-0.15*** [0.006]	0.33	0.24	-0.09 [0.158]
Archbishoprics	0.09	0.11	0.02 [0.407]	0.10	0.08	-0.02 [0.566]	0.08	0.10	0.02 [0.725]
Universities	0.08	0.12	0.04 [0.104]	0.06	0.07	0.01 [0.783]	0.06	0.07	0.02 [0.674]
Log Taxes per capita	3.33 (n=73)	2.75 (n=130)	-0.574*** [0.000]	3.18 (n=51)	2.73 (n=71)	-0.46*** [0.000]	3.16 (n=29)	2.75 (n=46)	-0.411*** [0.000]
National Identity (2nd & 3rd Estates)	2.44 (n=69)	1.93 (n=110)	-0.51*** [0.000]	2.39 (n=48)	1.87 (n=60)	-0.52*** [0.000]	2.41 (n=27)	1.91 (n=39)	-0.50*** [0.000]
Observations	73	135	208	51	71	122	29	46	75

Table 1: The Balance of Observables Inside and Outside the Cinq Grosses Fermes. Columns (1), (2), and (3) give the raw averages and differences, along with p-values (in brackets) of a ttest of no difference in the variables. *, **, and *** represent 1%, 5%, and 10% significance respectively. Columns (4), (5), and (6) show the same statistics using only cities within 150 kilometers of the CGF boundary. Columns (7), (8), and (9) do the same analysis for cities within 75 kilometers of the CGF boundary.

4.2 Baseline Estimates

While Table 1 is reassuring for our identification strategy and suggests support for our main hypothesis, we would like to estimate the local treatment effect of the border while controlling for potentially confounding variables we can observe. To do this, we use OLS to estimate the following regression equation:

$$y_i = \alpha + \beta \text{CGF}_i + f(\text{geographic location}_i) + \theta \text{Urban}_i + \mathbf{G}'_i \cdot \mathbf{\Lambda} + \varepsilon_i \quad (1)$$

Where y_i is either the value of *Log Taxes per capita* from Necker or the Hyslop measure of *National Identity* for General Cahier city i . CGF is a dummy variable equal to one if the city is located in the Cinq Grosses Fermes. $f(\text{geographic location}_i)$ is either a set of spatial trends on both sides of the CGF boundary or a third order polynomial in the latitude and longitude of the cities.²² We will investigate several different functional forms for the spatial trends, however, in most regressions we will report estimates from specifications that include either linear or cubic trends. Urban_i is our measure of urban density. \mathbf{G}'_i is a vector of geographic (*Wheat Suitability*, *Ruggedness*, *DRivers*, and *DSeas*), religious (*Bishoprics* and *Archbishoprics*), educational (*Universities*), and political (*Communes*) controls. ε_i is an iid error term.

Table 2 reports the results from estimating Equation 2 for our two outcome variables using the full sample as well as using only cities 150 km and 75 km from the CGF boundary. We report the value of β , which is the coefficient of interest. We are testing two hypotheses in Table 2. First, were cities within the CGF taxed at a higher rate than those outside (i.e. was fiscal capacity higher in the CGF)? Second, did cities inside the CGF have a greater affiliation with national as opposed to local identity?

Regressions (1)-(3) provide strong support for the first hypothesis. Cities located within the CGF were taxed significantly higher than those outside. According to the estimate using the full sample and linear trends (Panel A), a city in the CGF paid approximately 27% more in taxes per capita than a city outside. When we restrict our attention to cities within 75 km of the boundary, this estimate increases to 39%. The estimates are somewhat less robust to using a quartic trend, though the sizes of the coefficients are still quite large (24% at 75 km) even when we narrow the sample. Under estimation using the cubic polynomial in latitude and longitude, the coefficients are similar in magnitude to those in Panels A and B, though they are also more precisely estimated.

²²Following Dell (2010), the specific polynomial we use is $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$, where x denotes longitude and y denotes latitude.

	Dependent Variable					
	Log Taxes Per Capita			National Identity (2nd & 3rd Estates)		
	Full Sample	<150 km	<75 km	Full Sample	<150 km	<75 km
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Linear Polynomial in Distance to CGF Border						
CGF Dummy	0.266** (0.116)	0.324** (0.148)	0.388** (0.161)	0.669*** (0.154)	0.679*** (0.236)	0.707** (0.253)
R-sq	0.671	0.557	0.473	0.222	0.237	0.204
Panel B: Cubic Polynomial in Distance to CGF Border						
CGF Dummy	0.448** (0.169)	0.211 (0.152)	0.238 (0.251)	0.624** (0.290)	1.319** (0.560)	0.566 (0.878)
R-sq	0.677	0.582	0.457	0.237	0.227	0.250
Panel C: Cubic Polynomial in Latitude and Longitude						
CGF Dummy	0.378*** (0.0827)	0.414*** (0.0775)	0.382*** (0.0800)	0.532*** (0.119)	0.571*** (0.119)	0.606*** (0.134)
R-sq	0.722	0.715	0.759	0.275	0.246	0.274
Spatial Trends	Y	Y	Y	Y	Y	Y
Pop Density Control	Y	Y	Y	Y	Y	Y
Geography Controls	Y	Y	Y	Y	Y	Y
Rel., Educ., & Political Controls	Y	Y	Y	Y	Y	Y
Observations	203	122	75	179	108	66

Table 2: The Effect of the Cinq Grosses Fermes Border on *National Identity*: Baseline Estimates. Estimates set the entire land border of the CGF as 0 distance. Dependent variable is either *Log Taxes per capita* or *National Identity*. Robust standard errors clustered on province reported in parentheses. *, **, and *** represent 1%, 5%, and 10% significance respectively.

Our second hypothesis, that greater state capacity causes greater national affiliation, implies that the β coefficient should be positive and significant. This is precisely what is found in the estimates in columns (4)-(6). Assuming linear trends (Panel A) and using the full sample, a city inside the CGF boundary has about a one standard deviation higher affiliation with *National Identity* than a city just outside the border. Furthermore, the effect size remains relatively stable as we restrict the sample to cities closer to the border. Under the cubic spatial trend specification in Panel B, these results are supported for the full sample and restricting the sample to cities within 150 km of the boundary. It is only under the most restrictive specification, using just the 66 cities within 75 km of the border, that the estimate becomes imprecise. Though even then, its economic significance is still large ($\beta = 0.57$). In Panel C, using the cubic polynomials in latitude and longitude, the coefficients are large, precisely estimated, and stable as the border is approached. The estimate of 0.61 in Column (6), Panel C suggests individuals just inside the CGF border have about a one standard deviation greater affiliation with *National Identity* compared to those just outside the CGF.

Overall, the results in Table 2 support our claim from section 3.1 that fiscal capacity was much greater within the region of the CGF. They also support our primary hypothesis that citizens living in regions with higher fiscal capacity were more likely to identify with national as opposed to local institutions. We illustrate these results visually in Figure 4.²³ In the Figure distance to the CGF border is on the x-axis (negative is inside the CGF) and the value of national identification is on the y-axis. The values of *Identity* averaged over province are graphed along with the quartic fit of these against distance from the border after partialling out the control variables.²⁴ The vertical distance between the fitted line at ‘0’ distance from the CGF border is the estimated treatment effect.

The Figure does raise a couple potential concerns that we will address in the sections below. First, as we mentioned above in reference to Figure 3, the Gironde and La Rochelle regions in the west are outliers. The cities in the province of La Rochelle (Rochefort, St. Jean, La Rochelle, and Saintes) are actually split by the CGF boundary. Also, the relatively prosperous city of Bordeaux and it’s surrounding regions exhibits much higher national identification than other cities in the South. In Section 4 below we will address this potential source of bias in our estimates by investigating the treatment effect of CGF membership on cities near the individual border segments of the West, South, and East separately.

Another potential source of concern is that Lyon appears to be an influential observation. This could be inappropriate given our decision to include it as a CGF city due to its special status. Because of this concern, we will show in section 3 below that our results are robust to the exclusion of Lyon from the sample.

²³The figure is generated using a local polynomial regression with a bandwidth of 200,000 meters and allowing for fourth order polynomials.

²⁴The average distance of the individual cities from the border within the province is used as the ‘distance’ of the province from the border.

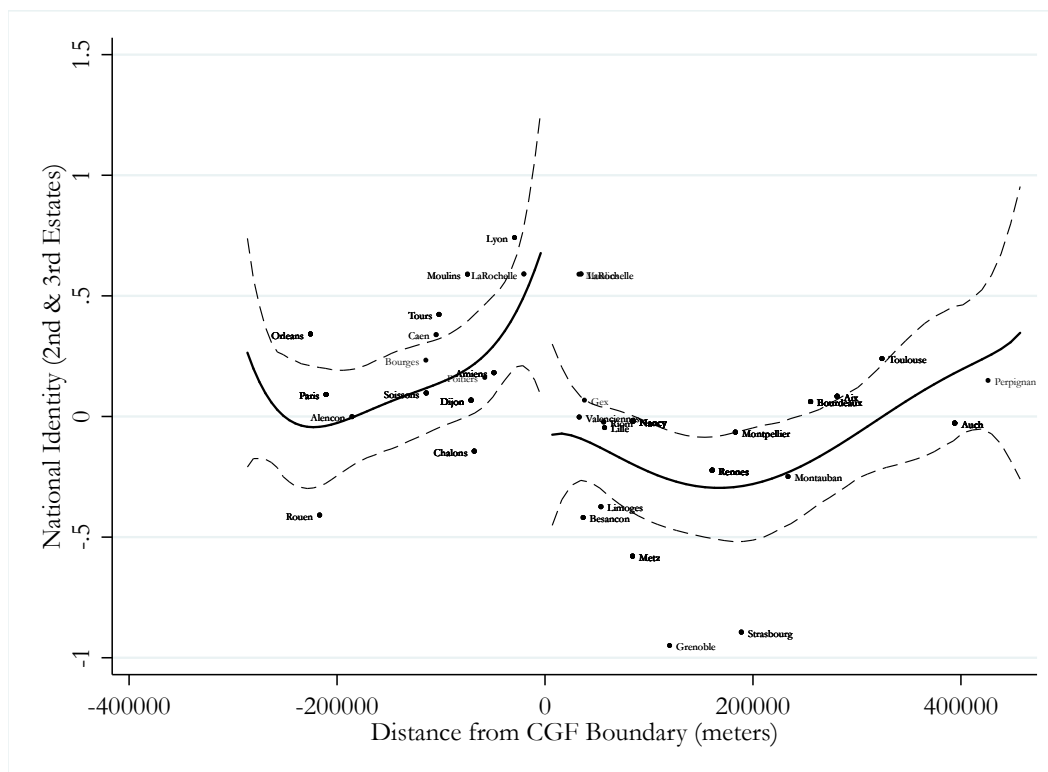


Figure 4: The discontinuity in *National Identity* at the CGF border. Non-parametric estimate of effect of distance from CGF border on *Identity* estimated for both sides border after partialling out all control variables. 95% confidence intervals shown by dashed lines.

5 ROBUSTNESS

5.1 Specification Tests and Alternative Political Boundaries

In Table 3 below we demonstrate the robustness of our main results. In Panels A and B we test the specification of the spatial trends using a quadratic polynomial (Panel A) and a quartic polynomial (Panel B). The results for both *Log Taxes per capita* and *National Identity* using a quadratic polynomial are similar to using the linear trends. All the estimates are both economically and statistically significant and are robust to restricting the sample down to 75 kilometers from the CGF boundary. When we use quartic polynomials in distance to the border in Panel B, the results look very similar to our baseline estimates using a cubic trend. Both fiscal capacity and *National Identity* are much higher for cities inside the CGF region than outside up to the restriction of using only observations within 150 kilometers of the border. However, when we restrict the sample to 75 kilometers from the border, then the estimates on both *Log Taxes per capita* and *National Identity* become indistinguishable from zero – though the effect sizes are still quite large.

In Panels C and D we investigate the robustness of our main results to the exclusion of Lyon - the second most important city in France in the eighteenth century. The estimates using both outcome variables under both the linear and the cubic trends are similar in size and significance to the baseline results in Table 2.

In Panels E and F we control for whether a city was in a region using civil law as opposed to customary law as well as controlling for whether the region was a Pays D'état or a Pays D'élection. Court cases in Early-Modern France that did not deal with 'high crimes' (e.g. witchcraft, lèse majesté, patricide, etc. . .) were adjudicated under one of two general forms of law. In the North, customary law was used and in the south roman law was dominant. Figure 22 in Appendix G shows the regions using these two different legal regimes. The main differences between these two legal regimes are analogous to those between common law and civil law (See Le Bris (2015); La Porta et al. (1998)).²⁵ It's thus possible that state development, economic development, or willingness to use the court system could have been endogenous to the judicial regime in place.

The other variable controlled for in Panels E and F is whether a city was located in a Pays D'État or Pays D'Election. The Pays D'État were more independent of the absolute monarchy. They had their own high courts (parlements) and possessed greater control over taxation (Marion, 1999). In the Pays D'Élection, by contrast, the monarchy had more control over the courts and fiscal matters. The distinction between Pays D'État and Pays D'Election, furthermore, preceded the establishment of the CGF boundary – often by hundreds of years. As such, it is important to control for this in order to distinguish the effect of the increases in state capacity taking place within the CGF from those that already existed. Figure 21 in Appendix G shows the map of the Pays D'États and Pays D'Elections regions. They match quite closely with the CGF boundary except for the region south of the Gironde into Gascony and, importantly, Burgundy. This latter difference is especially interesting since Burgundian was one of the main sources of opposition to the Bourbon monarchy – for example many of the leaders of the civil war known as the Fronde (1648-1653) were from there. Also Burgundy, or as it was known for centuries, the 'Kingdom of Burgundy' had its own distinct set of cultural and political institutions (Davies, 2012).

As can be seen from Panels E and F in Table 3, controlling for a region's legal origin and Pays D'État versus Pays D'Election has little appreciable effect on our estimates. Both the size and significance of the coefficients showing the effect of the CGF border on *National Identity* are similar to the baseline specifications.

Lastly, in Panel G we show simple OLS estimates without any spatial trends so as to give an idea of the average treatment effect of CGF membership. As expected, these all exhibit the correct sign and are highly significant. The sizes of the coefficients in specifications (1) and (4), furthermore, are comparable to those in Table 2. This suggests our estimates of the local treatment effect of the CGF boundary on fiscal capacity and social identity can be generalized to regions receiving the treatment, but further from the boundary.

²⁵The data on legal regime and parlements come from Le Bris (2015).

	Dependent Variable					
	Log Taxes Per Capita			National Identity (2nd & 3rd Estates)		
	Full Sample	<150 km	<75 km	Full Sample	<150 km	<75 km
(1)	(2)	(3)	(4)	(5)	(6)	
Panel A: Quadratic Polynomial in Distance to CGF Border						
CGF Dummy	0.326** (0.158)	0.522*** (0.183)	0.405* (0.194)	0.707*** (0.202)	0.879* (0.430)	1.633*** (0.514)
R-sq	0.670	0.562	0.470	0.244	0.225	0.239
Panel B: Quartic Polynomial in Distance to CGF Border						
CGF Dummy	0.329* (0.183)	0.331* (0.189)	-0.359 (0.458)	0.852* (0.475)	1.190* (0.646)	1.708 (1.242)
R-sq	0.675	0.576	0.467	0.231	0.217	0.298
Panel C: Linear Polynomial in Distance to CGF Border, Drop Lyon						
CGF Dummy	0.264** (0.118)	0.312* (0.153)	0.346** (0.158)	0.657*** (0.156)	0.641** (0.232)	0.687** (0.282)
R-sq	0.669	0.550	0.465	0.213	0.222	0.170
Panel D: Cubic Polynomial in Distance to CGF Border, Drop Lyon						
CGF Dummy	0.463** (0.191)	0.122 (0.133)	0.0603 (0.213)	0.619* (0.320)	1.226* (0.625)	0.620 (0.907)
R-sq	0.675	0.578	0.457	0.228	0.210	0.216
Panel E: Linear Polynomial in Distance to CGF Border, Control for Civil Law and Pays D'Etats						
CGF Dummy	0.263** (0.121)	0.314* (0.164)	0.399** (0.182)	0.651*** (0.155)	0.620** (0.241)	0.640** (0.261)
R-sq	0.670	0.544	0.457	0.226	0.220	0.201
Panel F: Cubic Polynomial in Distance to CGF Border, Control for Civil Law and Pays D'Etats						
CGF Dummy	0.452** (0.185)	0.203 (0.167)	0.282 (0.260)	0.616** (0.275)	1.155** (0.530)	0.212 (0.871)
R-sq	0.676	0.572	0.441	0.228	0.202	0.271
Panel G: OLS						
CGF Dummy	0.426*** (0.0820)	0.418*** (0.0925)	0.392*** (0.114)	0.391*** (0.102)	0.556*** (0.0947)	0.588*** (0.149)
R-sq	0.651	0.558	0.489	0.209	0.200	0.191
Spatial Trends	Y	Y	Y	Y	Y	Y
Pop Density Control	Y	Y	Y	Y	Y	Y
Geography Controls	Y	Y	Y	Y	Y	Y
Rel., Educ., & Political Controls	Y	Y	Y	Y	Y	Y
Observations	203	122	75	179	108	66

Table 3: Robustness and Specification Tests. Robust standard errors clustered on province reported in parentheses. *, **, and *** represent 1%, 5%, and 10% significance respectively.

5.2 Analysis of Separate Boundary Segments

In Table 4 we investigate how robust our estimates are by focusing on separate segments of the CGF boundary. Natural choices for the segments are in the West on the Brittany border, in the East, and in the South. These choices are illustrated in Figures 5, 6, and 7.



Figure 5: Western CGF Boundary.

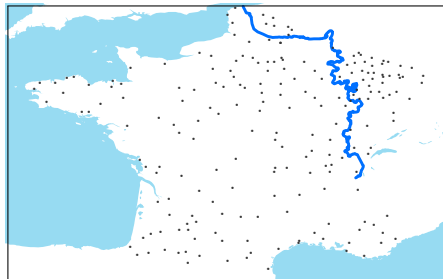


Figure 6: Eastern CGF Boundary.



Figure 7: Southern CGF Boundary.

In Table 4 we report the results of running our baseline specification on samples of cities that are restricted to be near either the Western, Eastern, Southern CGF boundary segments. Because the sample sizes become small very quickly when we focus on individual borders, we only report results using the full sample of cities.²⁶

Using linear trends, the impact of the CGF boundary on the Western border is economically significant and very large. The β coefficient on fiscal capacity in column (1) suggests that taxes per capita are between one-half and two-thirds lower just over the Western CGF boundary. The coefficient of 0.85 estimated in column (2) in Panel A suggests a local treatment effect of CGF membership of 1.4 times a standard deviation in the *National Identity* variable ($\sigma = 0.61$). When we use the cubic trends in Panel B, the results on the Western border remain robust. The estimate in column (2) in Panel B suggests a treatment effect of being inside the CGF boundary on the Western border of three times a standard deviation in *National Identity*.

The results using the Eastern border are also highly robust. Taxes are about 50% lower just over the Eastern boundary according to the estimates in column (3). *National Identity* is also much higher just inside the CGF. The estimates with linear trends suggest a being in the CGF raises national affiliation by a little over one-half a standard deviation. Using the cubic trends, this estimate triples and retains its statistical significance.

The Southern border yields the least robust results. Under both linear and cubic trends, there is no effect of being inside the CGF on fiscal capacity. Specification (6) yields mixed results on *National Identity*. The estimate in Panel A suggests an economically significant effect, however, under the cubic specification in Panel B, the coefficient is basically reduced to zero. This suggests a non-robust effect in the South - a fact that is consistent with the outlier observations in La Rochelle and around Bordeaux we observe in Figure 4.

²⁶When using the full sample, we exclude cities outside the CGF that are not near the relevant border. So for example, when investigating the Western border we exclude cities in the South and East that are not in the CGF.

Dependent Variable	West CGF Border		East CGF Border		South CGF Border	
	Log Tax pc (1)	Identity (2)	Log Tax pc (3)	Identity (4)	Log Tax pc (5)	Identity (6)
Panel A: Linear Polynomial in Distance to CGF Border						
CGF Dummy	0.660*** (0.120)	0.845*** (0.252)	0.480*** (0.146)	0.376* (0.207)	0.0557 (0.101)	0.446** (0.214)
R-sq	0.671	0.264	0.675	0.346	0.660	0.148
Panel B: Cubic Polynomial in Distance to CGF Border						
CGF Dummy	0.484* (0.229)	1.823*** (0.426)	0.441** (0.199)	0.905* (0.514)	0.00208 (0.160)	0.114 (0.371)
R-sq	0.673	0.281	0.667	0.344	0.661	0.168
Spatial Trends	Y	Y	Y	Y	Y	Y
Pop Density Control	Y	Y	Y	Y	Y	Y
Geography Controls	Y	Y	Y	Y	Y	Y
Rel., Educ., & Political Controls	Y	Y	Y	Y	Y	Y
Observations	89	80	130	112	130	125

Table 4: The Effect of the Cinq Grosses Fermes Border on *Log Taxes per capita* and *National Identity*: Estimates Using Separate Border Segments. Robust standard errors clustered on province reported in parentheses. *, **, and *** represent 1%, 5%, and 10% significance respectively.

Overall, the results in Table 4 suggest that support for our hypothesis that higher fiscal capacity is strongly associated with greater affiliation with national identity is coming primarily from cities close to the Eastern CGF boundary and the boundary between Brittany and the rest of France in the West. The results for the Eastern boundary are particularly supportive of our hypothesis since, as mention in above, Burgundy is the furthest East region inside the CGF, and yet it is also did not have historically strong roots to the French monarchy.

5.3 The CGF Boundary and Linguistic Differences

Another concern raised in Section 3.1 was that our main results may be spurious because of unobserved differences in cultural or linguistic composition between the CGF and non-CGF regions. It would be less surprising if individuals in the CGF regions affiliated themselves more with a national as opposed to a local identity if the very reason they joined in the first place was because they all spoke the same language. Though, we should note that even if national identity formed around a common ethno-linguistic group (which, to some extent, it almost certainly did), there is still the question of why this happened when it did and what institutions initially facilitated the formation of the common national identity.

There are at least two possible interpretations of of the associations we show in Section 4. The first is

that increases in state capacity caused increases in national identity. The second is that some unobserved characteristic, such as genetic or linguistic heritage, caused state capacity to increase and that this then became the institutional vehicle through which national identity was expressed. Both are interesting stories. One way to distinguish between them is to investigate the correlation between language use and the CGF boundary. In Early-Modern France, the major dialects in use were Langue d'oil, Langue d'oc, and Breton (see de Certeau et al. (2002) for a discussion). Data on which regions spoke these dialects are extremely scarce. There were a few unscientific studies during the late eighteenth century (see, e.g. de Certeau et al. (2002)), but due to the inherent measurement problems with these, we instead use the data from Furet and Ozouf (1977) on literacy rates around 1686 and 1786.

The Furet and Ozouf (1977) data exist because a retired school master named Louis Maggiolo was able to persuade the French government in 1877 to undertake a systematic survey of the country's literacy over time. About 16,000 school teachers across all the regions of France agreed to visit local archives and scrutinize marriage records to see what proportion of inhabitants signed their names versus making their marks. According to Schofield (1968), the ability to sign one's name is equivalent to being between a reading level and a writing level in terms of literacy. This suggests that, especially for the eighteenth century data, signing one's marriage certificate implies a level of education at which you would be conversant in the language of the central government (langue d'oil). The geographic distribution of male and female literacy rates at the end of the seventeenth and eighteenth centuries are shown in Figures 16, 17, 18, and 19 in Appendix F.

Table 5 shows the results of running specifications based on Equation 1 on this measure of French language use. The main finding from the table is that there was no relationship between the use of the patois and CGF membership in either the seventeenth or the eighteenth centuries. Overall, the results in Table 5 provide support for the similarity of regions inside the CGF border and outside it as one approaches the border. This also suggests that there is validity to the story that the increases in state capacity that took place within the CGF border played a causal role in generating the national identity that is present in the Cahiers.

	Dependent Variable					
	Male Literacy in 1686			Male Literacy in 1786		
	Full Sample	<150 km	<75 km	Full Sample	<150 km	<75 km
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Linear Polynomial in Distance to CGF Border						
CGF Dummy	0.0410 (0.0632)	-0.0541 (0.0838)	0.0128 (0.0854)	-0.0764 (0.0984)	-0.210 (0.134)	-0.0991 (0.129)
R-sq	0.291	0.211	0.249	0.194	0.150	0.247
Panel B: Cubic Polynomial in Distance to CGF Border						
CGF Dummy	-0.123 (0.117)	-0.120 (0.154)	0.00496 (0.163)	-0.294 (0.182)	-0.206 (0.260)	-0.0438 (0.271)
R-sq	0.339	0.191	0.198	0.265	0.121	0.241
Spatial Trends	Y	Y	Y	Y	Y	Y
Pop Density Control	Y	Y	Y	Y	Y	Y
Geography Controls	Y	Y	Y	Y	Y	Y
Rel., Educ., & Political Controls	Y	Y	Y	Y	Y	Y
Observations	206	121	75	206	121	75

Table 5: The difference in language use inside and outside the CGF in 1686 and 1786. Robust standard errors clustered on province reported in parentheses. *, **, and *** represent 1%, 5%, and 10% significance respectively. *Source:* Furet and Ozouf (1977)

5.4 The CGF Boundary and Migration

Another possible source of bias in our estimates of the relationship between the CGF border and national identity is migration. It's possible that individuals who valued strong national institutions moved into the CGF and those who did not moved out. There are no data on migration flows *per se* for Early-Modern France. However, aside from the Huguenot diaspora in the aftermath of the Revocation of the Edict of Nantes (1685) there are no known major population movements between 1661 and 1789. Furthermore, we can look at the correlations between city sizes in 1700, 1750, and 1800 using the Bairoch et al. (1988) data to verify this intuition.²⁷ The Pearson correlation coefficient for log city size in 1700 and 1800 is 0.92 and is highly statistically significant. Between 1750 and 1800 the coefficient is 0.97. These statistics suggest no major population flows.

In addition to these simple correlations, we can also use the data from Bairoch et al. (1988) to test whether city growth was systematically different on either side of the CGF border during the eighteenth century. In Table 6 we present estimates of the baseline specification using city growth during the seventeenth and

²⁷The Bairoch cities are different than the Cahiers cities. The data set contains population estimates based on primary and secondary sources for all cities in France with populations greater than 5,000. The map of the distribution of the Bairoch cities is presented in Appendix F, Figure 20.

eighteenth centuries as the dependent variables. In columns (1), (2), and (3) we investigate the whether urban growth was bigger on one side of the CGF border during the eighteenth century. None of the coefficients are statistically significant, suggesting there was no difference in the growth of urban density on either side of the CGF border during the eighteenth century. Of course, to interpret this as evidence of no migration requires the assumption that migration would have been biased in one direction or the other. If there were balanced sorting across the border, then one would expect no difference in urban growth rates.

In columns (4)-(6) we investigate whether urban growth was greater inside the CGF than outside it during the seventeenth century. We do this to exclude the possibility that regions enjoying faster or slower economic development may have been chosen to join the CGF in the first place. The regressions show no robustly discernible difference in growth rates. Only the cubic polynomial regression in Panel B, column 6 is statistically significant. However, this is based on only 22 observations and the regressions in columns (5) and (6) suggest it is not robust.

	Dependent Variable					
	City Growth 1700-1800			City Growth 1600-1700		
	Full Sample	<150 km	<75 km	Full Sample	<150 km	<75 km
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Linear Polynomial in Distance to CGF Border						
CGF Dummy	-0.126 (0.132)	-0.0992 (0.186)	-0.254 (0.227)	-0.0488 (0.224)	0.0302 (0.278)	0.0192 (0.302)
R-sq	-0.00255	0.0679	0.0807	0.225	0.369	0.523
Panel B: Cubic Polynomial in Distance to CGF Border						
CGF Dummy	-0.0483 (0.249)	-0.127 (0.377)	-0.198 (0.514)	0.442 (0.391)	0.395 (0.665)	1.480* (0.699)
R-sq	0.00112	0.117	0.0467	0.254	0.296	0.698
Spatial Trends	Y	Y	Y	Y	Y	Y
Geography Controls	Y	Y	Y	Y	Y	Y
Observations	169	87	57	61	32	22

Table 6: Growth in urban population inside and outside the CGF during the seventeenth and eighteenth centuries. Standard errors reported in parentheses. *, **, and *** represent 1%, 5%, and 10% significance respectively. City data from Bairoch (1988). Geography controls from Nunn and Qian (2011) and include soil suitability for potatoes, soil suitability for other Old World crops, and terrain ruggedness.

6 NATIONAL IDENTITY, COOPERATION, AND ECONOMIC OUTCOMES

Our argument is that increased state capacity generated preferences in favor of uniform, national, institutions regulating economic and political activity. Furthermore, we claim that these preferences generated higher

economic development in the nineteenth century. In this section we will provide evidence for these claims. First, we will look at the disaggregated data on the Cahiers in order to see whether individuals inside the CGF had preferences for economic and political reforms that were more conducive to coordinating action on the national rather than the local scale. Second, we will look at whether the nobility and the third estate (mainly urban bourgeois in the Cahiers) were more likely to have the similar preferences inside the CGF than outside or it. Third, we will look at measures of economic development in the nineteenth century. The primary measure we will focus on is direct tax receipts collected between 1817 and 1821. This is a good measure of the income of regional populations since, after the Revolution, the French government collected these in a uniform way (same rates etc. . .) across the entire country. Additionally, we will look at measures of use of the formal contracting activity, as proxied by notarial transaction. We will also look at fertility rates in 1821 inside and outside the CGF in order to evaluate the extent to which each region was transitioning from a Malthusian to a modern economic growth regime Galor (2011). Lastly, we will look at amount of local public goods spending inside and outside the CGF between 1878 and 1902 – the earliest these data are available. This will provide us with additional evidence that regions inside the CGF had norms which inclined them more towards cooperation than those regions outside.

6.1 Preferences for National Institutions Inside the CGF

In this section we look at the disaggregated data from the Cahiers in order to test whether regions inside the CGF were more likely to have preferences for formal institutions that encouraged a uniform treatment of individuals on political and economic margins. We will look at separate correlations for different institutions for each of the first, second, and third estates. The way the Cahiers are coded is that when an institution is mentioned, then this gives that city/estate a one. If the institution is never mentioned, then the Cahier is coded as a zero on that topic. The *National Identity* variable that we've been using so far is an average over many topics that were coded up. Once the disaggregated data are used, unfortunately, sample sizes generally become very small. For this reason, we will restrict our attention to investigating simple correlations in this sub-section.

Columns (1), (2), and (3) of Table 7 investigates stated preferences for a national education system, uniform administrative and legal codes, and uniform economic regulations respectively. All of these institutions match closely with what we have been calling 'national institutions' so far in the paper. All three estates expressed strong support for them. For example, the statistic in column (3) of Panel A suggests that there was 20% more support for regulations supporting greater economic uniformity inside the CGF than outside. The p-value for the t-test of difference in means is given in brackets and the number of positive ('yes') observations across the whole noble cahiers sample is reported as 30. The only national institution which does not have statistically significant greater support inside the CGF is national education, however, that statistic is still positive (6%).

Columns (4) - (5) report on institutions that explicitly undermined the feudal regime. Column (4) addresses

whether or not the Cahier expressed support for giving the third estate enough power to outvote the first and second estates in the upcoming Constituent Assembly. This was a major issue since in all previous Estates General, the traditional forces of the feudal regime – the nobility and clergy – were always able to outvote the third estate despite its larger numbers since votes were tallied by estate rather than by head. The statistics in column (4) are all positive, but relatively small. Notably, the nobility were 6% more likely to support vote by head inside the CGF than outside. Column (5) reports the number of Cahiers indicating that they wanted legal uniformity in the sense that all estates would be treated equally under the law. There is significantly more support for this inside the CGF than outside.

In columns (7) and (8) we look at which Cahiers asked for feudal dues or serfdom to be abolished. Across the board there is more support for these institutional changes inside the CGF.

Overall, the numbers in Table 7 support the argument that regions inside the CGF held preferences for institutions that would coordinate political and economic activity at the national, as opposed to at the local, level.

	Topic Mentioned in Cahiers							
	Support for National Education	Uniformity of Admin and Legal System	Greater Economic Uniformity	Approve Vote by Head in Estates General	Same Law for All Estates	Suppress vs Maintain Guilds	Abolish Feudal Dues	Abolish Serfdom
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Cahiers of Nobility							
Mean Diff (CGF - Not CGF)	0.123**	0.101*	0.201***	0.060*	0.108***	0.191	0.067**	0.034*
	[0.015]	[0.055]	[0.000]	[0.083]	[0.001]	[0.355]	[0.042]	[0.092]
number "yes" obs =	31	37	30	14	11	Suppress=13 Maintain=3	11	5
	Panel B: Cahiers of Third Estate							
Mean Diff (CGF - Not CGF)	0.057	0.111*	0.300***	0.007	0.236***	0.142	0.095	0.095**
	[0.294]	[0.057]	[0.000]	[0.464]	[0.000]	[0.194]	[0.162]	[0.051]
number "yes" obs =	35	54	111	1	68	Suppress=53 Maintain=40	147	31
	Panel C: Cahiers of Clergy							
Mean Diff (CGF - Not CGF)	0.189***	0.100**	0.094**	0.045	0.079**	0.200	0.094*	0.082***
	[0.001]	[0.008]	[0.008]	[0.399]	[0.036]	[0.297]	[0.065]	[0.001]
number "yes" obs =	41	17	16	32	15	Suppress=10 Maintain=1	33	7

Table 7: The effect of CGF membership on preferences for standardization and reform. T-tests of difference in mean number of positive responses reported between the CGF cahiers and the non-CGF cahiers. P-values in brackets. Number “yes” Obs = gives the number of districts advocating the reform topic in the General Cahiers. For every test, there are 208 total General Cahiers for each of the Nobility, Third Estate, and the Clergy. There are 135 cahiers outside the CGF and 73 inside. *, **, and *** represent 1%, 5%, and 10% significance respectively.

6.2 Agreement Between Second and Third Estates

An important part of the argument of this paper is that higher state capacity makes it more likely that individuals cooperate because they come to share a common sense of identity. In Table 8 we test this by looking at whether being located within the CGF made the second and third estates more likely to hold the same value for *National Identity*. Specifically, we construct a variable called *SD Identity* which measures the standard deviation of the identities of the two estates in each city. We then run our baseline specifications in Table 8 using this as the dependent variable.

	Dependent Variable = Standard Deviation of 2nd and 3rd Estate Identity					
	Linear Polynomial in Distance to CGF Border			Cubic Polynomial in Distance to CGF Border		
	Full Sample	<150 km	<75 km	Full Sample	<150 km	<75 km
	(1)	(2)	(3)	(4)	(5)	(6)
CGF Dummy	-0.315*	-0.514*	-0.774**	-0.560*	-0.868*	-0.908
	(0.164)	(0.257)	(0.272)	(0.306)	(0.448)	(0.864)
R-sq	0.107	0.121	0.260	0.131	0.147	0.201
Spatial Trends	Y	Y	Y	Y	Y	Y
Pop Density Control	Y	Y	Y	Y	Y	Y
Geography Controls	Y	Y	Y	Y	Y	Y
Rel., Educ., & Political Controls	Y	Y	Y	Y	Y	Y
Observations	108	66	43	108	66	43

Table 8: The effect of CGF membership on variance of *National Identity* between second and third estates. Robust standard errors clustered on province reported in parentheses. *, **, and *** represent 1%, 5%, and 10% significance respectively.

The coefficient estimates in Table 8 strongly support the hypothesis that the second and third estates were more likely to hold common social preferences inside the CGF than outside. Given that a standard deviation in *SD Identity* is 0.37, the linear specifications reported in columns (1) - (3) suggest being inside the CGF increases agreement between the two estates by between one and two standard deviations. Under the cubic specification, these estimates increase to about a two to two and half standard deviation decrease in the outcome variable. Overall, these results suggests that not only did national affiliation increase inside the CGF during the eighteenth century, but that this process also generated a *shared* sense of identity between the nobility and urban merchants inside the CGF region.

6.3 National Identity and Economic Outcomes in the Nineteenth Century

We now look at the level of economic development inside the CGF compared to outside during the nineteenth century. Since the data we have on nineteenth century outcomes are at the Department rather than the city level, we abandon the regression discontinuity approach used in previous sections and instead run simple OLS

regressions, with a large number of control variables, to identify our effects. The main specification we will run is:

$$y_{ij} = \alpha + \beta \text{cgfXidentity}_i + \theta \text{Urban}_i + \mathbf{G}'_j \cdot \mathbf{\Gamma} + \mathbf{P}'_j \cdot \mathbf{\Delta} + \varepsilon_{ij} \quad (2)$$

where y_{ij} is equal to one of our six department-level measures of nineteenth century outcomes for city i in department j . The six measures are *Tax1817-21*, *Doors&Wind1822*, *Contracts1817-21*, *Fertility1821*, *LPGVillage*, and *LPGState*. *Tax1817-21* measures the average amount paid between 1817 and 1821 to the four direct wealth taxes known as the *contribution foncière*, *contribution personnelle mobilière*, *patente* and the *portes et fenêtres*.²⁸ These taxes, which eventually became known as the *quatre vieilles*, were collected in the same way (e.g. same rates and bases) across all the territories of France for decades. As such, they can be used as a gauge of economic development in the same way that, for example, Becker and Woessmann (2009) use tax receipts for nineteenth century Prussia. *Doors&Wind1822* comes from the census of 1822 and measures the number of visible doors and windows from the outside of all residential buildings adjusted for population of the department. As such, this variable is a rough measure of the wealth of a region. *Contracts1817-21* is the average amount of tax paid on notary contracts in the department between 1817 and 1821. This tax, known as the *Droit d'enregistrement* was assessed at a uniform rate (by type of transaction) on the value of all notary contracts. These sorts of contracts were used for mortgages, sales of high value items, or any transaction involving a real asset.²⁹ As such, this measure is particularly useful for us since it proxies for the degree to which individuals are using national-level formal sector institutions to intermediate transactions. *Fertility1821* measures completed fertility in 1821 and comes from Bonneuil (2007). This year is chosen as it is the earliest observation we can get that is unaffected by the upheavals of the Revolutionary or Napoleonic periods. Furthermore, we focus on early fertility since we want to identify the origins of the transition from the Malthusian to the Modern growth regimes and, over the subsequent decades, all Departments in France are converging to relatively low fertility rates.

We also investigate two dependent variables that measure spending on public goods. The first concerns spending by Communes, which are the equivalent of county, or district, level spending in the United States.³⁰ These were the lowest level administrative divisions in nineteenth century France for which data is generally available. *LPGVillage* measures the per capita Departmental average of spending on education by all of its communes between 1855 and 1902. These data are based on the series collected by Diebolt and Trabelsi

²⁸These data were collected by David LeBris from the *Annuaire Statistiques* and used in Le Bris (2015). As described in Le Bris (2015): ‘The *contribution foncière* is a tax on the value of real estate (land and building) paid by the owner. The *contribution personnelle mobilière* is a tax on the value of houses, but paid by the occupant and not the owner. The *patente* is a tax on business (trade and industry) mainly based on the value of the building occupied. The *portes et fenêtres* is a tax on the number of doors and windows visible from the street, again paid by the occupant and not the owner.’

²⁹See Le Bris (2015) for more discussion.

³⁰The three main administrative divisions in nineteenth century France, from the lowest level of hierarchy to the highest, were the Commune, the Department, and the State.

(2009) and extended by Franck and Johnson (2014).³¹ $LPGState$ represents education spending by state level governments in the Department. We do not expect this last variable to be related to norms of cooperation since national spending on public goods relied less on Department level norms of cooperation than on uniform allocation rules set at the national level.

On the right hand side of Equation 2 we have $cgfXidentity_i$, which is a dummy variable equal to one times the value of *National Identity* for city i if that city is inside the CGF boundary. If city i is outside the CGF boundary, then $cgfXidentity_i$ is equal to zero. In effect, this variable is a weighted measure of the state-capacity treatment for city i . It is also our variable of interest.

As controls we include $Urban_i$ which is equal to the late eighteenth century urban density of city i and is identical the variable used in the baseline regressions above.³² G_j is a vector of department and city-level geographic controls including: wheat suitability, linear distance of the city from the sea or nearest major river, ruggedness of the terrain surrounding the city, and the average altitude of the department in which the city is located.³³ P_j is a vector of city or department-level political and cultural controls including: bishoprics, archbishoprics, university, and commune, which are identical to the controls used in Section 2. Additionally, P contains a measure of the protestant population of the department in which the city is located in 1866 as well as two dummy variables indicating whether the department used to be part of a Pays D'État or Pays D'Élection and whether the department used roman or customary law during the ancien régime.³⁴ ε_{ij} is an error term. To take into account the fact that all the dependent variables are measured at the Department level even though we are also using city-level data when we estimate Equation 2 we will always cluster errors at the Department level.

Table 9 reports the results of estimating Equation 2 on $Tax1817-21$, $Doors\&Wind1822$, $Contracts1817-21$, $Fertility1821$, $LPGVillage$, and $LPGState$. Columns (1) and (2) show the estimates using $Tax1817-21$ as the dependent variable. In column (1), when no controls are included, the association between cgf membership weighted by national identity is both economically and statistically very strong. The coefficient of 2.06 is interpreted as a one standard deviation increase in $cgfXidentity$ is associated with 0.62 of standard deviation increase in $Tax1817-21$. When all the controls are included, the coefficient size is cut in half but remains economically and statistically convincing. Regression (2) in Table 9 is our strongest evidence for an effect of nationality identity on subsequent economic development in France. Non-parametric estimates of all the regressions that include controls in Table 9 are shown in Figure 11 in Appendix E. As can be seen from the non-parametric figure the relationship between $Tax1817-21$ and $cgfXidentity$ increases monotonically which is exactly what we would predict.

³¹We use population from the Annuaire Statistique in 1886 to convert these numbers to per capita terms.

³²We have also run regressions using later measures of urban density and the results are unchanged. These regressions are available on request.

³³With the exception of the altitude measure, which is from Le Bris (2015), these controls are identical to those used in the baseline regressions.

³⁴These last three, department level, control are from Le Bris (2015).

In columns (3) and (4) we report the results of estimating Equation 2 on *Doors&Wind1822*. The results here are again robust. When all controls are included the coefficient is 0.046 which is interpreted as a one standard deviation in *cgfXidentity* is associated with about a fifth of a standard deviation increase in *Industry1860*. This offers additional support for the hypothesis that the support for strong national institutions that was associated with increases in fiscal capacity during the ancien regime, also were associated with greater economic development.

In columns (5) and (6) we look at the *Contracts1822* which measures the value of transactions intermediated by notaries. The results here, again, robustly show that higher association with national identity was also associated with greater use of formal financial instruments. Furthermore, the conditional non-parametric relationship shown between *cgfXidentity* and *Contracts1822* in panel (c) of Figure E suggests a gradual increase in the use of notaries as reported identity from the cahiers increases.

Columns (5) and (6) investigate the relationship between fertility rates in 1821 and *cgfXidentity*. We look at this variable for two reasons. First, the demographic transition plays a central role in unified growth models which attempt to explain the transition from the Malthusian economy to modern economic growth (see, e.g. Galor and Weil (2000)). Furthermore, the French demographic transition is particularly interesting since it seems to have been both very early and very influential across Europe (Spolaore and Wacziarg, 2014). This set of regressions can offer insight into whether or not the influence of the French state on social identity may have been correlated with the shift in fertility behavior that was associated with the French demographic transition. The results in columns (5) and (6) are suggestive, but somewhat weaker than for the purely economic outcomes. Both coefficients are negative and, with no controls, the coefficient in (5) is also precisely estimated. This suggests that there is some association between affiliation with national identity and fertility restrictions. When all controls are included, however, the coefficient on *cgfXidentity* is statistically insignificant and the size of the coefficient is cut in half. This is perhaps consistent with the work of Spolaore and Wacziarg (2014) which suggests that the underlying processes driving economic innovation were different than those driving fertility adjustment.

We next investigate the effect of *National Identity* on local public goods provisionment. It is well known in the economics and political science literatures that in developing economies, fractionalization is negatively related to a localities overcoming the costs of collectively providing LPG's such as education, welfare, and other infrastructure (Alesina et al., 1999; Miguel, 2004). We predict that local public goods provisionment, measured as *LPGVillage* should be positively correlated with *cgfXidentity*. This is because those who affiliate more strongly with national identity will be more willing and able to overcome the coordination problems necessary to contribute to public goods than those who have a more local outlook (see, e.g., Alesina et al. (1999) or Miguel (2004)). By similar logic, we predict that spending by the national government in a Department should be uncorrelated with *National Identity* since national spending was less contingent on local preferences for public goods (and thereby should have been unaffected by heterogeneity etc. . .).

	Dependent Variable											
	Tax1817-21		Doors&Wind 1817-21		Contracts 1817-21		Fertility1821		LPG Village 1855-1902		LPG State 1855-1902	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
cgfXidentity	2.063*** (0.351) [0.62]	1.140*** (0.242) [0.35]	0.108*** (0.0303) [0.46]	0.0462** (0.0220) [0.20]	0.506*** (0.121) [0.46]	0.252** (0.112) [0.23]	-0.0225** (0.00946) [0.25]	-0.0138 (0.0108) [0.15]	0.126*** (0.0297) [0.43]	0.0705** (0.0274) [0.24]	-0.0122 (0.0261) [0.06]	-0.00328 (0.0276) [0.02]
R-sq	0.391	0.588	0.200	0.499	0.205	0.484	0.0616	0.227	0.190	0.511	-0.00279	0.467
Pop Density Control	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y
Geography Controls	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y
Rel., Educ., & Political Controls	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y
Observations	176	176	176	176	176	176	166	166	168	167	168	167

Table 9: National identity and nineteenth century economic development. Standard errors clustered on department reported in parentheses. *, **, and *** represent 1%, 5%, and 10% significance respectively. Numbers in brackets represent the proportion of a standard deviation of the dependent variable a one standard deviation increase in cgfXidentity induces. *Source:* See text.

Columns (9) and (10) of Table 9 show the association between national identification inside the CGF and education spending at the village level around the middle of the nineteenth century (earlier data on local public goods are, unfortunately, not available). Overall, these results support our hypothesis that *National Identity* should be correlated with the ability of a region to invest in local public goods such as state provided education. With all the controls included, the coefficient on *cgfXidentity* suggests a one standard deviation increase in identity is associated with a quarter of a standard deviation increase in local public goods spending. This result provides support for the crowding-in hypothesis concerning the relationship between strong state institutions and informal norms of cooperative behavior (Bowles and Polania-Reyes, 2012; Tabellini, 2010; Becker et al., 2014). We should note, however, that we are unable to identify whether the mechanism through which local public goods spending is higher is through formal (e.g. better bureaucracy) or informal (e.g. trust behavior) institutions. Encouragingly, when we investigate the relationship between *cgfXidentity* and spending by the state on local public goods (*LPGState*) we find no relationship. This suggests that the results for Village level spending are not being generated by some unobserved factor such as rent seeking from the central government.

Overall, the results of our regressions looking at outcomes for the nineteenth century provide robust support for the correlation between preferences for national institutions and economic development. Furthermore, given that we restrict our attention to national identity within the CGF region, and combined with our other results, there appears to be a strong association between state capacity, national identity, and economic development.

7 CONCLUSIONS

In *The Old Regime and the French Revolution*, Alexis de Tocqueville describes how, in creating a strong fiscal and legal state, the Absolute Monarchy sowed the seeds of its own destruction and paved the way towards modernity. The research described in this paper provides vital empirical support for this claim by showing that, even within France, variation in the strength of the institutions of the Monarchy had a profound effect on the social identity of its subjects and their subsequent support for national economic and political institutions. Furthermore, these institutions appear to explain a great deal of variation in the regional distribution of economic development during the early nineteenth century.

APPENDICES FOR ONLINE PUBLICATION ONLY

A CORRELATED EQUILIBRIUM AND THE STATE AS CHOREOGRAPHER

The purpose of this section is to summarize the model of coordinating behavior described in Gintis (2014) and to relate it to an explanation of how state capacity can increase coordination between individuals.

Consider the following simple n -player game described by Gintis (2014): Each player chooses an integer k from 1 to 10. Nature also chooses an integer k . Each player wins 1 if they all choose the same integer. They receive 0 if not. Each player can also query nature for a noisy signal equal to the true value of k with probability $p > 0.10$. The simple best response in this game is for each player is to query nature and play the value of the signal. The expected payoff is then p^n .

There also exists a correlated equilibrium of this game in which the players recognize a social rule which requires the youngest player to reveal her choice (based on the signal she receives from nature). This is a correlated equilibrium because, assuming players n_{-i} follow the rule, it is a best response for player i to follow it as well. Payoffs are now p . The social rule (to follow the youngest player) in this example can be interpreted as the *choreographer* of a correlated equilibrium.

According to Gintis (2014, 32), ‘A correlated equilibrium of an epistemic game G is a Nash equilibrium of a game G^+ , in which G is augmented by an initial move by a new player, whom we call the choreographer, who observes a random variable γ on a probability space (Γ, p) , and issues a ‘directive’ $f_i(\gamma) \in S_i$ to each player i as to which pure strategy to choose. Following the choreographer’s directive is a best response for each player [assuming all other players follow choreographer].’ This definition may seem unrealistic since the choreographer implies some external agency, however, there is a well known theorem of Aumann (1987) that shows Bayesian rational agents in an epistemic game G with common subjective priors play a correlated equilibrium of G . So there need not be an external agent. In terms of our simple game, so long as the players know the probability distribution of the types of the other players and their payoffs (and this is common knowledge) and if they are rational, then they will play as if there is an explicit choreographer. To take a relevant example from the paper, priors might consist of beliefs about the probability that each player assigns to which system of weights and measures to adopt.

How does being subject to the same set of fiscal and legal institutions generate common priors? There is no agreement on the philosophical grounding of probabilistic statements (von Mises, 1981; de Laplace, 1996; Gillies, 2000; Keynes, 2004). However, if we assume *symmetric reasoners* subject to *mutually accessible* events, then common knowledge can be generated (Gintis, 2014, 140).

Let an *indicator* be an event N that specifies a *social event* E to all individuals in a group. E.g., the indicator

(N) that we are all having the monarchy's unit of account, the *Livre Tournois*, imposed on us specifies the social event (E) that the *Livre Tournois* is the appropriate unit of account to use in exchange. We can then define *Symmetric Reasoning* as follows: Individual i is a symmetric reasoner with respect to j for an indicator N of event E if, whenever i knows N , and i knows that j knows N , then i knows that j knows E (See (Gintis, 2014, 141) and (Vanderschraaf and Sillari, 2007)). Additionally, If an indicator (N) is *mutually accessible* to a given population, then Gintis (2014) proves that Bayesian rational symmetric reasoners will adopt E as common knowledge.

We follow Gintis (2014) in interpreting the development of *symmetric reasoning* as stemming primarily from sociobiological processes such as gene-culture coevolution (Cavalli-Sforza and Feldman, 1973; Boyd and Richerson, 1985). However, we interpret the role of increases in state capacity as increasing the *mutual accessibility* of indicators (N). This has two components:

1. Increasing the *number* of mutually accessible indicators.
2. Increasing the *size* of the population subject to the mutually accessible indicators.

Within the area of the CGF more mutually accessible indicators were present since the monarchy was imposing common fiscal and legal institutions there at the same time as it was suppressing local, feudal, institutions. Furthermore, the size of the population subject to these accessible indicators was dramatically increased by the creation of the CGF. We conclude, that individuals in the CGF were more likely to reach a cooperative equilibrium in which national level institutions were taken as the choreographer guiding them how to behave. Individuals outside the CGF, where state capacity was both weaker, and the populations generally smaller, did not benefit to the same extent from these choreographing institutions.

B INCLUSION CRITERIA FOR CITIES IN THE CGF

Les provinces des *cinq grosses fermes*, c'est-à-dire celles qui, suivant l'article 3 du titre premier de l'ordonnance de 1787, sont comprises dans l'étendue du tarif de 1664, à l'entrée & à la sortie desquelles se perçoivent les droits qu'il impose, sont la Normandie, la Picardie, la Champagne, la Bourgogne, dont il faut aujourd'hui distraire le petit pays de Gex, le Bourbonnois, le Berry, le Poitou, l'Aunis, l'Anjou, le Maine, & toutes celles qui sont renfermées dans le cercle intérieur que forment ces provinces, comme l'Orléanois, le Nivernois, la Tourraine, l'île de France & autres.

Le Beaujolois a été ajouté aux provinces des *cinq grosses fermes*, par arrêt du conseil du 10 avril 1717.

Figure 8: The Provinces Contained within the Cinq Gross Fermes. Extract from the 'Cinq Grosses Fermes' entry in le Rond d'Alembert and Diderot (1784).

La Provence en laissa établir de tous côtés; le Languedoc, dont le gouvernement comprenoit le Quercy, le Rouergue & le Vivarais, fut également séparé des provinces voisines par des bureaux; mais le Lyonnais conserva des communications privilégiées avec les *cinq grosses fermes*, & avec le Languedoc & la Provence.

Figure 9: Explanation of Special Lyonnais Relationship with Cinq Gross Fermes. Extract from the 'Cinq Grosses Fermes' entry in le Rond d'Alembert and Diderot (1784).

C INCREASES IN FRENCH FISCAL CAPACITY DURING THE SEVENTEENTH CENTURY

The sixteenth and seventeenth century French fiscal system was a patchwork of different taxes and loans. Receipts were categorized in two ways: extraordinary and ordinary revenues.³⁵ Extraordinary revenues were particularly important during times of economic distress (usually wartime) and came from temporary surtaxes (e.g. the *tailion*) or loans, sometimes forced, from the wealthy (e.g. the *dons gratuits* paid by the clergy). Ordinary revenues, by contrast, derived from taxes that were relatively well established and recurring. After the establishment of the Company of General Farms in 1664 all ordinary taxes were collected under the direct control of the state – that is the ‘capacity’ of the state to tax was well represented by the size of these tax revenues.³⁶ They included the local city taxes known as *octrois* (Franck et al., 2014) and indirect taxes of all sorts (e.g. taxes collected through the tax farms).



Figure 10: Total tax revenues and *tailles* revenues, 1600–1695. Vertical dashed lines indicate the dates France entered the Thirty Years War (1635) and the beginning of Louis XIV’s personal rule (1661)

The increase in the fiscal capacity of the centralized French state during the seventeenth century is shown below in Figure 10.³⁷ The dashed line shows the *per capita* silver value of all revenues collected. These include both ordinary tax revenues in addition to loans and temporary financial expedients. The solid line shows the revenues coming directly from the primary direct tax collected by the crown known as the *taille*. As is clear from Figure 10, while revenues were increasing throughout the century, there is a marked jump in the

³⁵See Guéry (1978) for an overview of ancien regime financial records.

³⁶For a detailed discussion of the establishment of fiscal capacity in France around this time, see Johnson and Koyama (2014a)

³⁷Data are from Richard and Bonney (2011) and are converted to silver equivalents Wailly (1857). Population data from 1700 are from Dupâquier (1988). ‘Total Revenues’ are equal to the sum of extraordinary revenues and ordinary revenues. ‘Tailles Revenues’ are a subset of ordinary revenues, the other major component coming from the tax farms Johnson (2006).

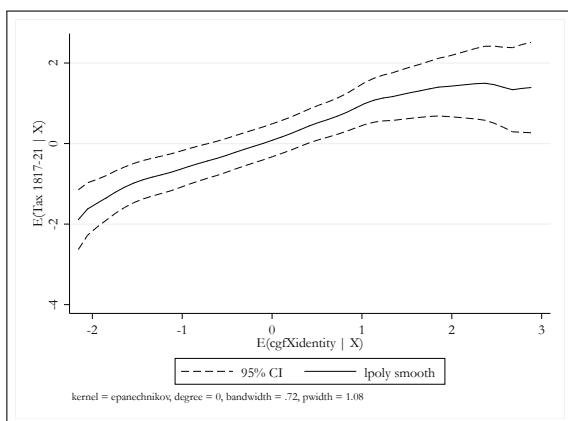
regime's capacity to collect taxes through the *tailles* starting around 1635 when France entered the Thirty Years War (1618-1648). Moreover, whereas increases in revenues between 1600 and 1635 tended to come from so-called 'extraordinary' sources, such as loans (voluntary and forced) and temporary taxes, after 1635 there was a dramatic rise in revenues coming from 'ordinary' sources that were managed directly by the crown. The move away from extraordinary taxes along with the massive increase in tax revenues collected through ordinary sources represents the increase in fiscal capacity we discuss in the paper.

D DESCRIPTIVE STATISTICS FOR NINETEENTH CENTURY VARIABLES

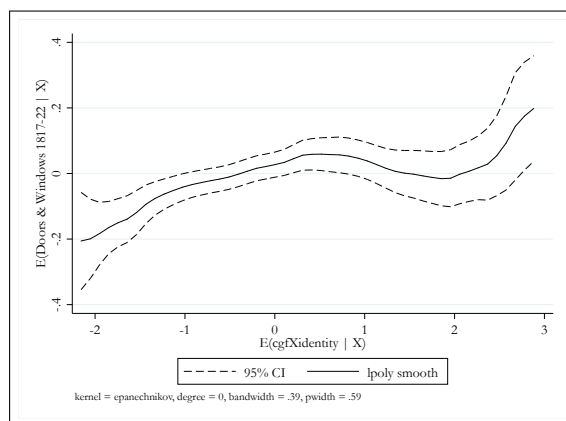
Variable	Obs	Mean	Std. Dev.	Min	Max
Variable of Interest and 19th c Outcomes					
cgfXidentity	176	0.95	1.23	0.00	3.00
Direct Tax 1817-21	176	11.97	4.06	5.51	22.13
Doors & Wind 1817-21	176	1.15	0.29	0.60	1.94
Contracts 1817-21	176	4.66	1.36	1.85	8.28
Fertility 1822	166	0.38	0.11	0.25	0.82
LPGVillage 1855-1902	167	-0.02	0.36	-0.96	0.50
LPGState 1855-1902	167	0.74	0.27	-0.02	1.58
Control Variables					
Wheat Suitability	176	3.61	1.22	1.00	7.00
Dseas	176	177244.40	115271.90	0.00	417348.30
Drivers	176	64832.67	72663.16	49.09	360389.10
Ruggedness	176	0.70	0.18	0.12	0.97
Altitude	176	292.41	295.64	34.00	1671.00
Bishoprics	176	0.39	0.30	0.00	1.00
Archbishoprics	176	0.11	0.18	0.00	1.00
Universities	176	0.11	0.15	0.00	1.00
Communes	176	0.55	0.37	0.00	1.00
Protestant	176	0.02	0.04	0.00	0.31
Civil Law	176	0.55	0.87	0.00	3.00
Pays D'Etat	176	0.34	0.48	0.00	1.00
Urban	176	26.91	17.79	7.64	103.50

Table 10: Descriptive statistics for nineteenth century regressions in Section 6.3. *Source:* See text.

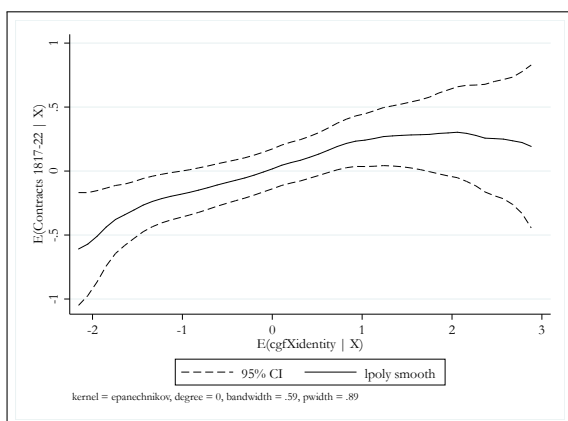
E NON-PARAMETRIC FIGURES FOR 19TH CENTURY OUTCOME REGRESSIONS



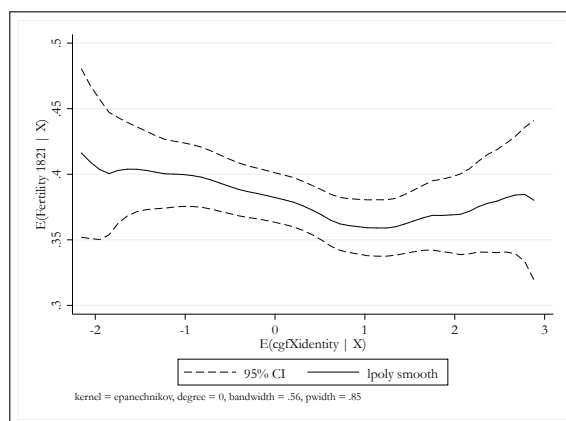
(a) Direct Tax Receipts 1817-1821 and National Identity



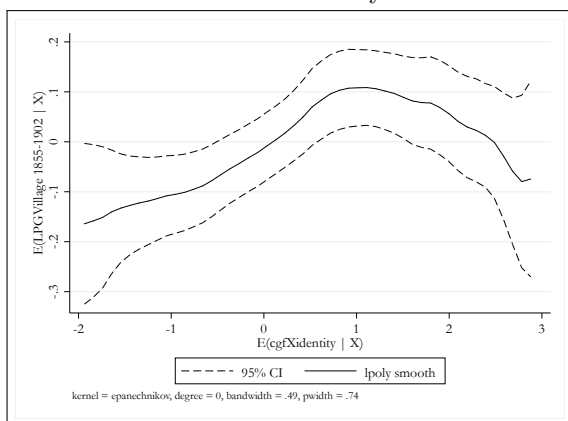
(b) Tax on Doors & Windows 1817-1821 and National Identity



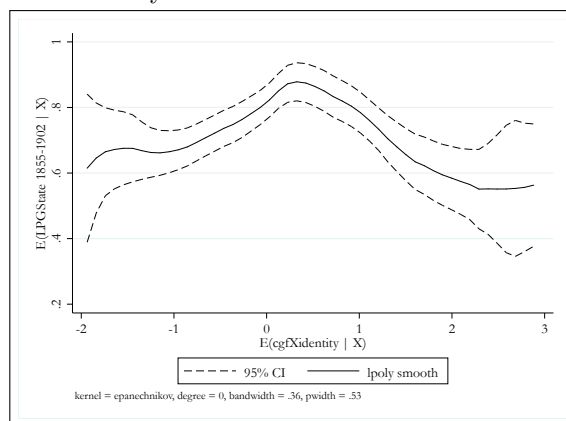
(c) Tax on Notary Contracts 1817-1821 and National Identity



(d) Fertility in 1822 and National Identity



(e) Village LPG Spending 1855-1902 and National Identity



(f) State LPG Spending 1855-1902 and National Identity

Figure 11: Non-parametric relationship between nineteenth century outcomes and cgFXidentity . All figures condition on control variables used in regressions in Section 6.3. Dashed lines show 95% confidence intervals.

F SUPPORTING FIGURES

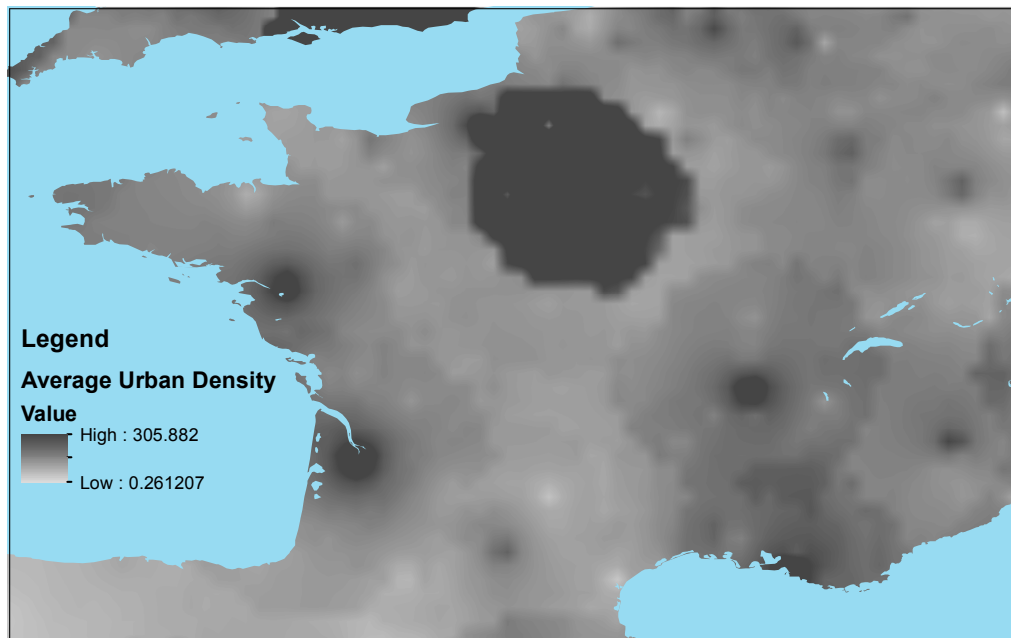


Figure 12: Urban Density. Constructed by assigning a value for urban density to each grid cell using the inverse weighted average (power of 1.5) of the surrounding Bosker et al. (2013) cities with populations of at least 5,000 in 1800.

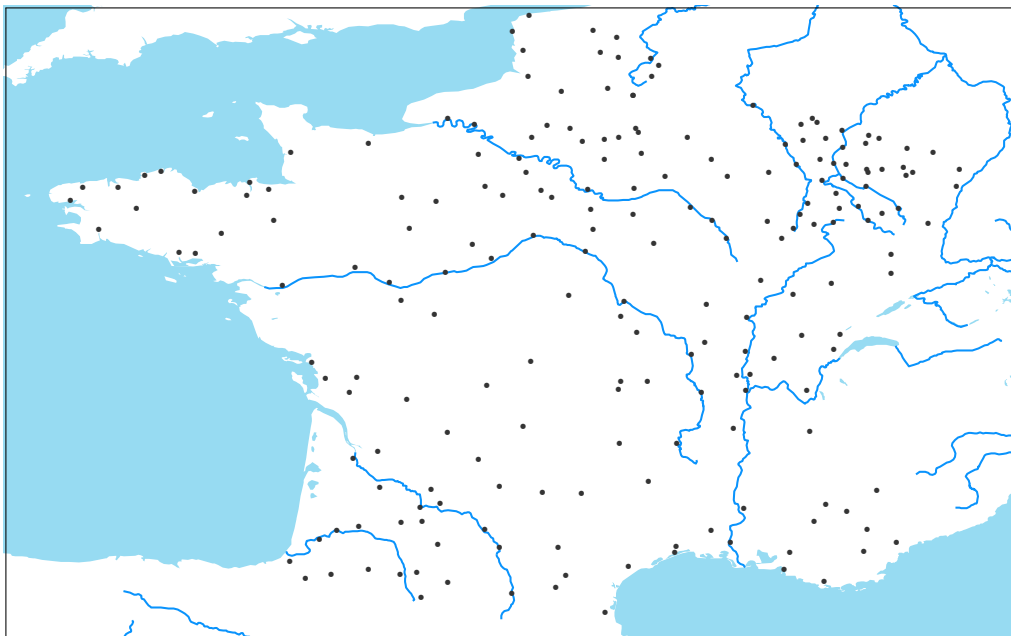


Figure 13: Major Rivers and Seas and the Cahiers Cities. *Source: Euratlas (2012).*

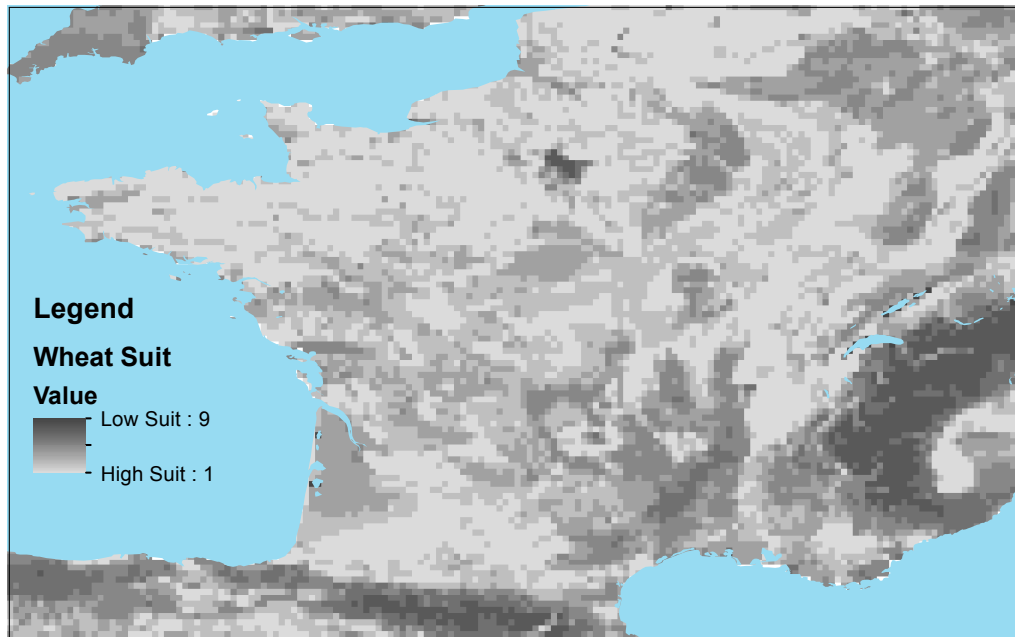


Figure 14: Wheat Suitability and Cahiers Cities. *Source:* Fischer et al. (2002).

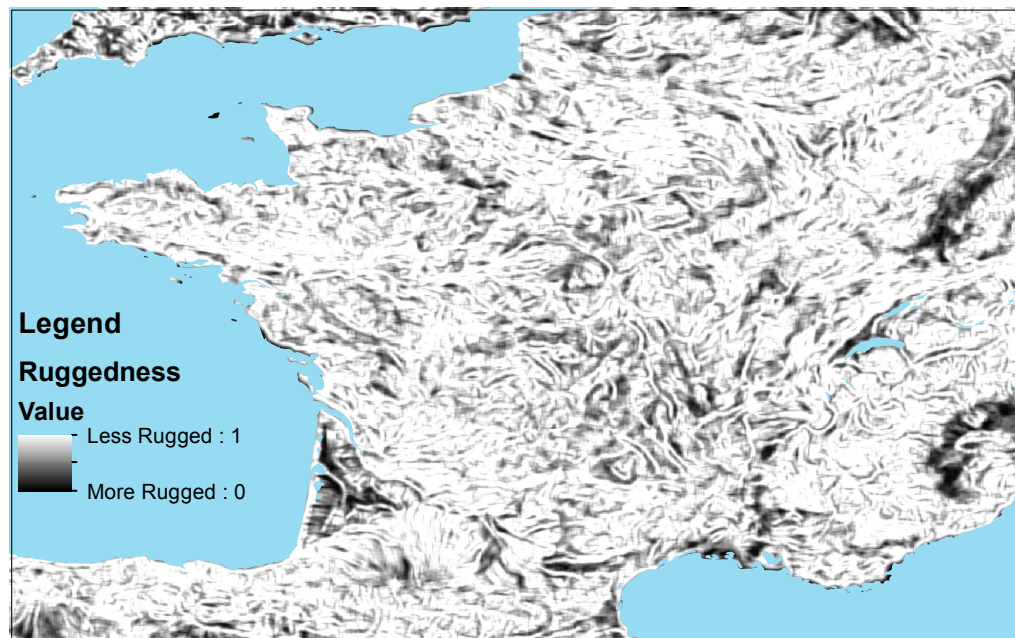


Figure 15: Terrain Ruggedness and Cahiers Cities. *Source:* Sappington et al. (2007).

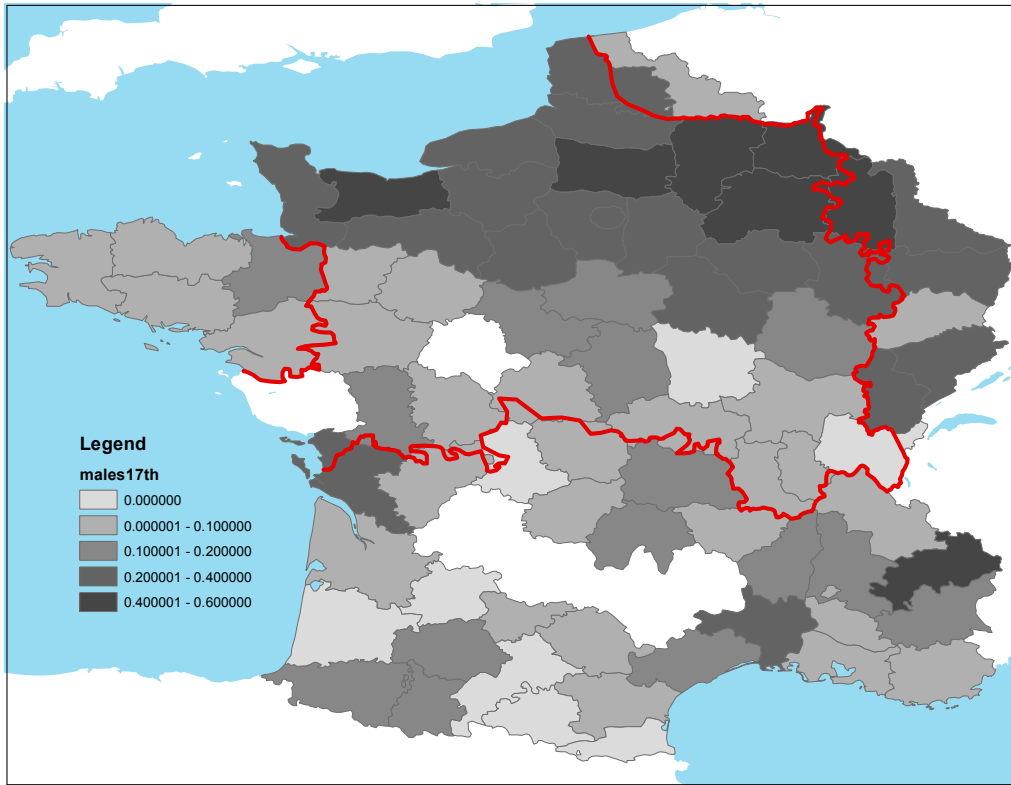


Figure 16: Male literacy rates in 1686. *Source:* Furet and Ozouf (1977).

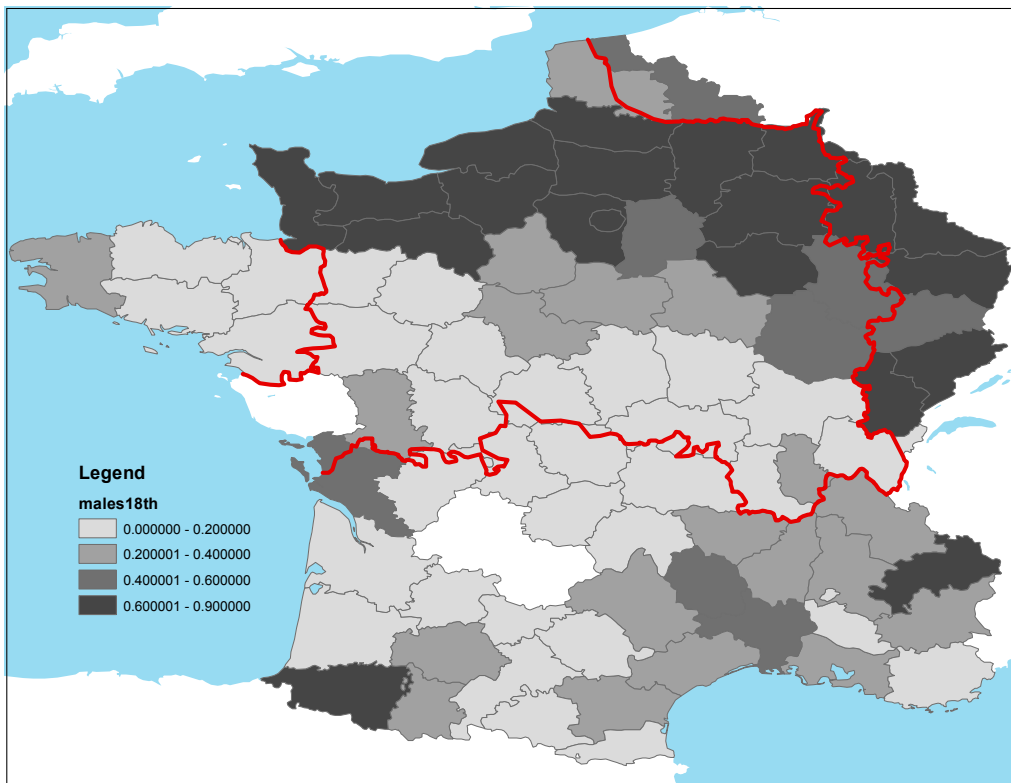


Figure 17: Male literacy rates in 1786. *Source:* Furet and Ozouf (1977).

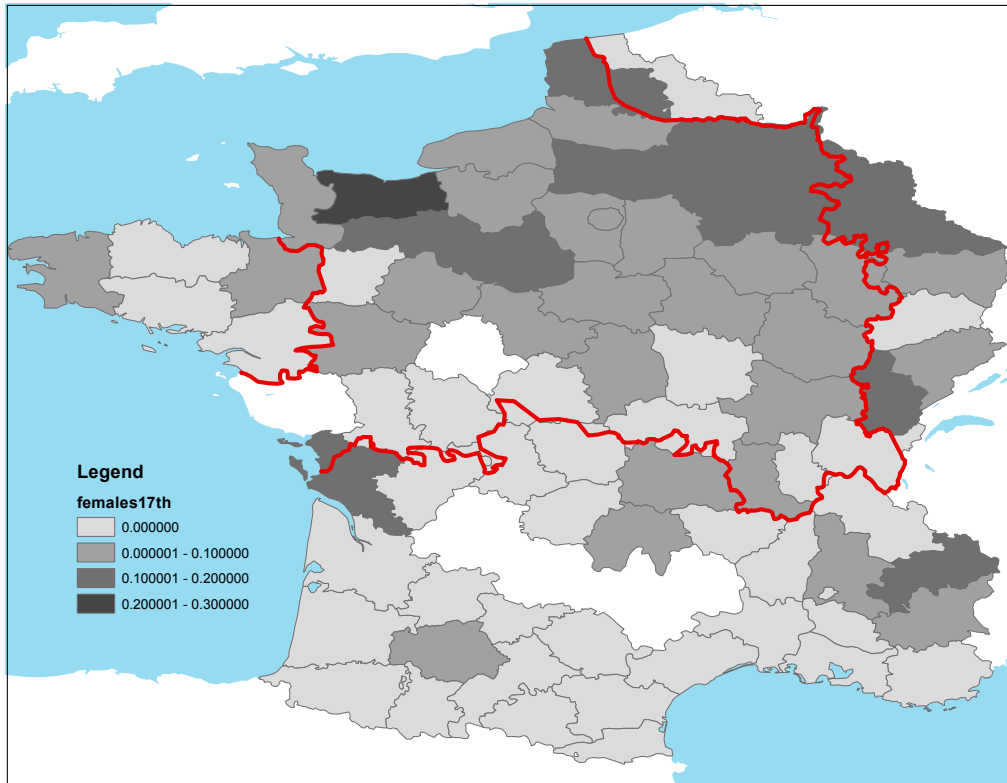


Figure 18: Female literacy rates in 1686. *Source:* Furet and Ozouf (1977).

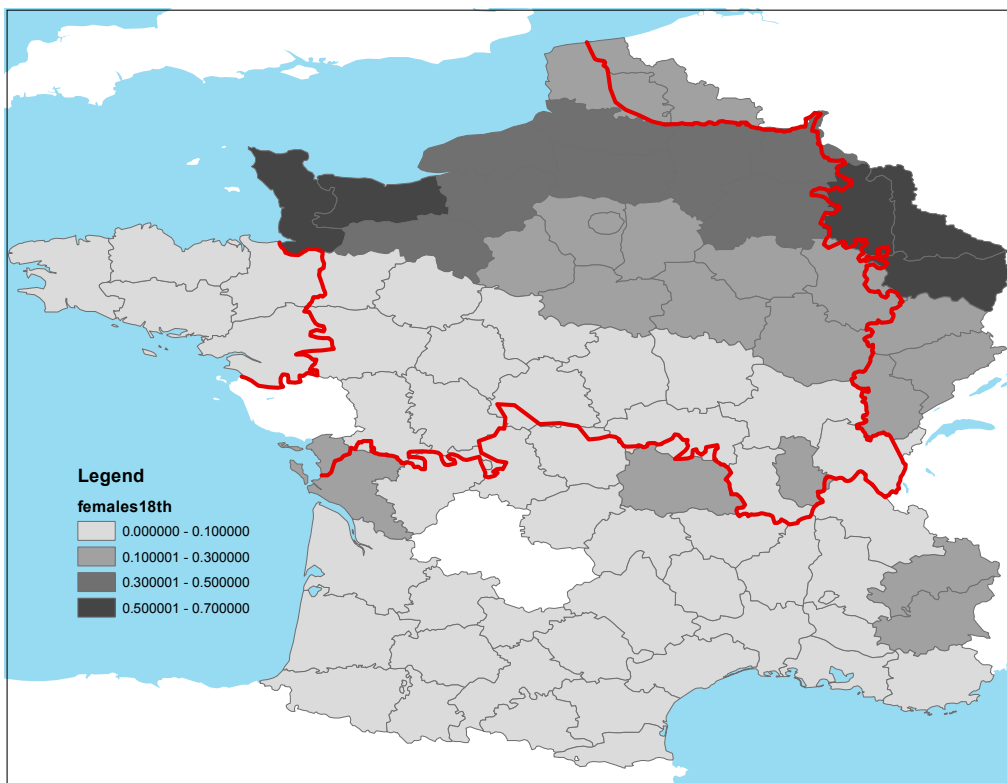


Figure 19: Female literacy rates in 1786. *Source:* Furet and Ozouf (1977).



Figure 20: Bairoch Cities and the CGF boundary. *Source:* Bairoch (1988).

G OTHER JURISDICTIONAL BORDERS

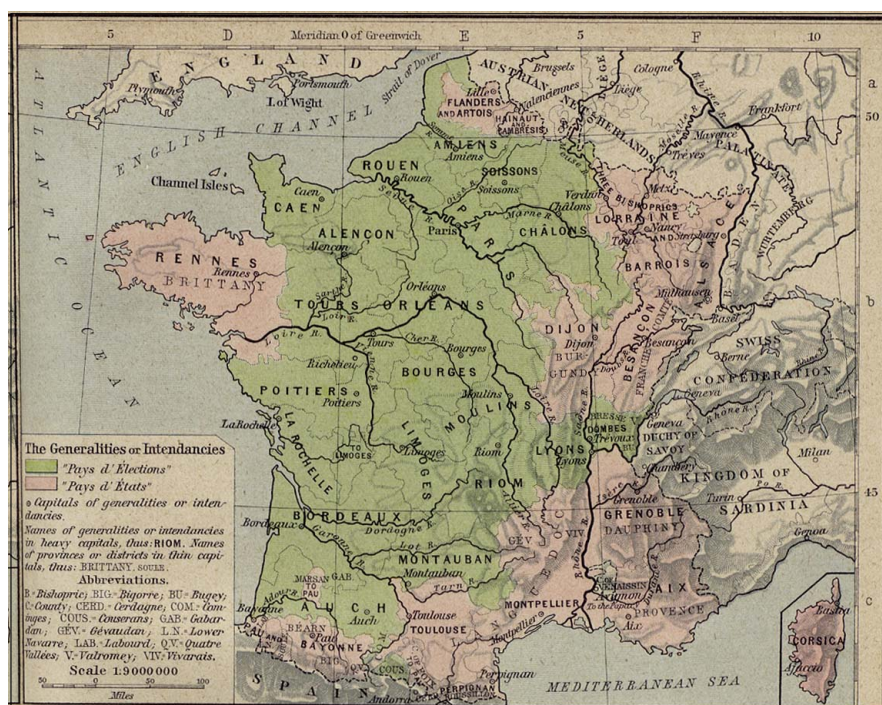


Figure 21: Map showing the Pays d'État and the Pays d'Élection around 1789. Source: Shepherd (1956).



Figure 22: Map showing Law Courts and Legal Regions around 1789. Source: Shepherd (1956).

REFERENCES

- Acemoglu, Daron and James A. Robinson (2006), *The Economic Origins of Dictatorship and Democracy*, Cambridge University Press, Cambridge.
- Acemoglu, Daron and Matthew O Jackson (2014), ‘History, expectations, and leadership in the evolution of social norms’, *The Review of Economic Studies* .
- Acemoglu, Daron, Simon Johnson and James Robinson (2005), ‘The rise of Europe: Atlantic trade, institutional change, and economic growth’, *American Economic Review* **95**(3), 546–579.
- Alesina, Alberto, Reza Baqir and William Easterly (1999), ‘Public goods and ethnic divisions’, *The Quarterly Journal of Economics* **114**(4), 1243–1284.
- Anderson, Benedict (1991), *Imagined Communities: Reflections on the Origin and Spread of Nationalism*, Verso.
- Aumann, Robert (1987), ‘Correlated equilibrium as an expression of bayesian rationality’, *Econometrica* **55**(1), 1–18.
- Bairoch, Paul (1988), *Cities and Economic Development, from the dawn of history to the present*, University of Chicago Press, Chicago. translated by Christopher Braider.
- Bairoch, Paul, Jean Batou and Pierre Chèvre (1988), *La Population Des Villes Européennes*, Droz, Geneva.
- Balla, Eliana and Noel D. Johnson (2009), ‘Fiscal crisis and institutional change in the Ottoman Empire and France’, *The Journal of Economic History* **69**(03), 809–845.
- Banerjee, Abhijit and Lakshmi Iyer (2005), ‘History, institutions, and economic performance: The legacy of colonial land tenure systems in india’, *The American Economic Review* **95**, 4.
- Becker, Sascha O, Katrin Boeckh, Christa Hainz and Ludger Woessmann (2014), ‘The empire is dead, long live the empire! long-run persistence of trust and corruption in the bureaucracy’, *The Economic Journal* .
- Becker, Sascha O. and Ludger Woessmann (2009), ‘Was Weber wrong? a human capital theory of Protestant economic history’, *The Quarterly Journal of Economics* **124**(2), 531–596.
- Bell, David A. (1995), ‘Lingua populi, lingua dei: Language, religion, and the origins of french revolutionary nationalism’, *The American Historical Review* **100**(5), 1403–1437.
- Besley, Timothy and Torsten Persson (2011), *Pillars of Prosperity*, Princeton University Press, Princeton, New Jersey.
- Bockstette, Valerie, Areendam Chanda and Louis Putterman (2002), ‘States and markets: The advantage of an early start’, *Journal of Economic Growth* **7**(4), 347–69.
- Bonneuil, Noël (2007), *Transformation of the French Demographic Landscape, 1806-1906*, Clarendon Press, Oxford, UK.
- Bonney, Richard, ed. (1995), *The Rise of the Fiscal State in Europe c.1200-1815*, Clarendon Press, OUP, Oxford.
- Bosher, J. F. (1964), *The Single Duty Project: A Study of the Movement for a French Customs Union in the Eighteenth Century*, University of London Press, London.
- Bosker, Maarten, Eltjo Buringh and Jan Luiten van Zanden (2013), ‘From Baghdad to London: unravelling urban development in Europe and the Arab world 800-1800’, *Review of Economics and Statistics* **95**(4), 1418–1437.
- Bowles, Samuel and Sandra Polania-Reyes (2012), ‘Economic incentives and social preferences: substitutes or complements?’, *Journal of Economic Literature* pp. 368–425.
- Boyd, Robert and Peter J. Richerson (1985), *Culture and the Evolutionary Process*, University of Chicago Press, Chicago.
- Brewer, John (1988), *The Sinews of Power*, Harvard University Press, Cambridge, M.A.
- Cavalli-Sforza, Luca and Marcus W. Feldman (1973), ‘Models for cultural inheritance: Within group variation’,

- Theoretical Population Biology* **42**(4), 42–55.
- Conchon, Anne (2002), *Le pèage en France au XVIII^e siècle*, CHEFF, Paris.
- Daudin, Guillaume (2010), ‘Domestic trade and market size in late-eighteenth-century France’, *The Journal of Economic History* **70**(03), 716–743.
- Davies, Norman (2012), *Vanished Kingdoms: The Rise and Fall of States and Nations*, Penguin.
- de Certeau, M., D. Julia and J. Revel (2002), *Une politique de la langue: la Révolution française et les patois : l’enquête de Grégoire*, Collection Folio. Histoire, Editions Gallimard.
- de Laplace, Marquis (1996), *A Philosophical Essay on Probabilities*, Dover, New York.
- Dell, Melissa (2010), ‘The persistent effects of peru’s mining mita’, *Econometrica* **78**(6), 1863–1903.
- Diebolt, Claude and Jamel Trabelsi (2009), ‘Human capital and French macroeconomic growth in the long run’, *Economies et Sociétés (Serie 'Histoire Economique Quantitative')* (40), 901–917.
- Dincecco, Mark (2009), ‘Fiscal centralization, limited government, and public revenues in Europe, 1650-1913’, *Journal of Economic History* **69**(1), 48–103.
- Dupâquier, Jacques (1988), *Histoire de la population Française*, Vol. 4 volumes., PUF.
- Epstein, S. R. (2000), *Freedom and Growth, the rise of states and markets in Europe, 1300–1700*, Routledge, London.
- Euratlas* (2012).
URL: http://shop.euratlas.com/maps_gis
- Expilly, J.J. (1768), *Dictionnaire géographique, historique et politique des Gaules et de la France*, number 5 vols., Desaint et Saillant.
- Fehr, Ernst and Karla Hoff (2011), ‘Introduction: Tastes, castes and culture: The influence of society on preferences’, *The Economic Journal* **121**, 396–412.
- Fehr, Ernst, Karla Hoff and Mayuresh Kshetramade (2008), Spite and development, Research Paper 4619, World Bank Policy.
- Fischer, Gunther, Harrij van Nelthuizen, Mahendra Shah and Freddy Nachtergaele (2002), *Global Agro-Ecological Assessment for Agriculture in the 21st Century: Methodology and Results*, Food and Agriculture Organization of the United Nations, Rome.
- Foucault, Michel (1977), *Discipline and Punish: The Birth of the Prison*, Peregrine Books, Vintage Books.
- Franck, Raphaël and Noel D. Johnson (2014), Can public policies lower religiosity? Evidence from school choice in late nineteenth century France.
- Franck, Raphaël, Noel D. Johnson and John V. C. Nye (2014), ‘From internal taxes to national regulation: Evidence from a French wine tax reform at the turn of the twentieth century’, *Explorations in Economic History* **51**, 77–93.
- Furet, François and Jacques Ozouf (1977), *Lire et écrire. L’alphabétisation des Français de Calvin à Jules Ferry*, Minuit, Paris.
- Galor, Oded (2011), *Unified Growth Theory*, Princeton University Press, Princeton, New Jersey.
- Galor, Oded and David N. Weil (2000), ‘Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond’, *American Economic Review* **90**, 806–828.
- Gillies, Donald (2000), *Philosophical Theories of Probability*, Routledge, London.
- Gintis, H. (2014), *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences*, Princeton University Press.
- Gintis, Herbert (2010), ‘Social norms as choreography’, *Politics, Philosophy and Economics* **9**(3), 251–264.
- Guéry, Alain (1978), ‘Les finances de la monarchie française sous l’ancien régime’, *Annales E.S.C.* **33**, 216–39.
- Heckscher, Eli F. (1955), *Mercantilism*, Vol. I, George Allen & Unwin LTD, London. translated by E.F. Soderlund.

- Hirschman, Albert O. (1977), *The Passions and the Interests: Political Arguments For Capitalism Before Its Triumph.*, Princeton University Press, Princeton, New Jersey.
- Hoff, Karla and Priyanka Pandey (2014), ‘Making up people – the effect of identity on preferences and performance in a modernizing society’, *Journal of Development Economics* **106**, 118–131.
- Hussonnet, Jacques (1996), ‘Les dénombremments Saugrain. Genèse et hypothèses’, *Annales de démographie historique* pp. 295–312.
- Hyslop, Beatrice (1934), *French Nationalism in 1789 According to the General Cahiers*, Cambridge University Press.
- Hyslop, Beatrice (1936), *A Guide to the General Cahiers of 1789: With the Texts of Unedited Cahiers*, Columbia University Press.
- Iyer, Lakshmi (2010), ‘Direct versus indirect colonial rule in india: Long-term consequences’, *The Review of Economics and Statistics* **92**(4), 693–713.
- Johnson, Noel D. (2006), ‘Banking on the king: The evolution of the royal revenue farms in Old Regime France’, *Journal of Economic History* **66**(4), 963–91.
- Johnson, Noel D. and Mark Koyama (2013), ‘Legal centralization and the birth of the secular state’, *Journal of Comparative Economics* **41**(4), 959–978.
- Johnson, Noel D. and Mark Koyama (2014a), ‘Tax farming and the origins of state capacity in England and France’, *Explorations in Economic History* **51**, 1–20.
- Johnson, Noel D. and Mark Koyama (2014b), ‘Taxes, lawyers, and the decline of witch trials in France’, *Journal of Law & Economics* **57**(1).
- Karaman, Kivanc and Şevket Pamuk (2013), ‘Different paths to the modern state in Europe: The interaction between warfare, economic structure and political regime’, *American Political Science Review* **107**(3), 603–626.
- Keynes, John Maynard (2004), *A Treatise on Probability*, Dover, New York.
- La Porta, Rafael, Florencio Lopez de Silanes, Andrei Shleifer and Robert W. Vishny (1998), ‘Law and finance’, *Journal of Political Economy* **106**(6), 1113–1155.
- Le Bris, David (2015), Testing legal origins theory within France: Customary laws versus roman code.
- le Rond d’Alembert, Jean and Denis Diderot, eds (1784), *Encyclopédie méthodique: Finances*, chez Panckoucke.
- Lowe, Sara, Nathan Nunn, James A Robinson and Jonathan Weigel (2015), ‘The evolution of culture and institutions: Evidence from the Kuba Kingdom’, *Harvard University, manuscript*.
- Marion, Marcel (1999), *Dictionnaire des Institutions de la France aux XVIIe et XVIIIe Siècles*, Picard, Paris.
- Markoff, John (1998), *Revolutionary Demands: A Content Analysis of the Cahiers de Doléances of 1789*, Stanford University Press, chapter Allies and Opponents: Nobility and Third Estate in the Spring of 1789.
- Markoff, John (2010), *The Abolition of Feudalism: Peasants, Lords, and Legislators in the French Revolution*, Pennsylvania State University Press.
- McCloskey, Deirdre N. (2010), *Bourgeois Dignity: Why Economics Can’t Explain the Modern World*, University of Chicago Press, Chicago.
- Michalopoulos, Stelios and Elias Papaioannou (2013), ‘Pre-colonial ethnic institutions and contemporary African development’, *Econometrica* **81**(1), 113–152.
- Miguel, Edward (2004), ‘Tribe or nation? nation building and public goods in Kenya versus Tanzania’, *World Politics* **56**, 328–362.
- Mokyr, Joel (2005), ‘The intellectual origins of modern growth’, *The Journal of Economic History* **65**(2), 285–351.
- Nordman, D. and M. Ozouf-Marignier, eds (1989), *Atlas de la Révolution française, vol 5: Le territoire*.
- North, Douglass C. and Barry Weingast (1989), ‘Constitutions and commitment: the evolution of institutions govern-

- ing public choice in seventeenth century England', *Journal of Economic History* **49**, 803–32.
- Nunn, Nathan (2012), 'Culture and the historical process', *Economic History of Developing Regions* **27**(sup1), S108–S126.
- Nunn, Nathan and Nancy Qian (2011), 'The potato's contribution to population and urbanization: Evidence from a historical experiment', *The Quarterly Journal of Economics* **126**(2), 593–650.
- Pinkovskiy, Maxim L. (2013), Economic discontinuities at borders: Evidence from satellite data on lights at night.
- Ploeckl, Florian (2013), 'The internal impact of a customs union; Baden and the Zollverein', *Explorations in Economic History* **50**, 387–404.
- Rebel, Hermann (1983), *Peasant Classes: The Bureaucratization of Property and Family Relations Under Early Habsburg Absolutism, 1511-1636*, Princeton University Press.
- Richard and Margaret Bonney (2011).
URL: <http://esfdb.websites.bta.com/Database.aspx>
- Rodrik, Dani, Arvind Subramanian and Francesco Trebbi (2004), 'Institutions rule: The primacy of institutions over geography and integration in economic development', *Journal of Economic Growth* **9**(2), 131–165.
- Ross, Lee and Richard Nisbett (1991), *The Person and the Situation: Perspectives of Social Psychology*, McGraw-Hill, New York.
- Roux, Pierre (1916), *Les fermes d'impôts sous l'ancien régime*, Rousseau et cie, Paris, France.
- Sabeau, D.W. (1987), *Power in the Blood: Popular Culture and Village Discourse in Early Modern Germany*, Cambridge University Press.
- Sappington, J.M., K.M. Longshore and D.B. Thomson (2007), 'Quantifying landscape ruggedness for animal habitat analysis: A case study using bighorn sheep in the mojave desert', *Journal of Wildlife Management* **71**(5), 1419–1426.
- Saugrain, C.M. (1720), *Nouveau dénombrement du royaume, par généralitez, élections, paroisses et feux...*, Chez Saugrain et P. Prault.
- Schofield, Roger (1968), *Literacy in Traditional Societies*, Cambridge University Press, Cambridge, chapter The Measurement of Literacy in Pre-industrial England, pp. 311–325.
- Schulze, Max-Stephan and Nikolaus Wolf (2009), 'On the origins of border effects: insights from the Habsburg Empire', *Journal of Economic Geography* **9**, 117–136.
- Shapiro, Gilbert, Timothy Tackett, P. Dawson and John Markoff (1998), *Revolutionary Demands: A Content Analysis of the Cahiers de Doléances of 1789*, Stanford University Press.
- Shepherd, William R. (1956), *Historical Atlas*, 8th edn, The Colonial Offset Co, Inc.
- Spolaore, Enrico and Romain Wacziarg (2014), Fertility and modernity.
- Swidler, Ann (1986), 'Culture in action: Symbols and strategies', *American Sociological Review* **51**, 273–286.
- Tabellini, Guido (2010), 'Culture and institutions: economic development in the regions of Europe', *Journal of the European Economic Association* **8**(4), 677–716.
- Tilly, Charles (1990), *Coercion, Capital, and European States, AD 990-1990*, Blackwell, Oxford.
- Tocqueville, Alexis de (1998), *The Old Regime and the Revolution*, Vol. 2, University of Chicago Press, Chicago.
- Vanderschraaf, Peter and Giacomo Sillari (2007), *The Stanford Encyclopedia of Philosophy*, Stanford University Press, Stanford, California, chapter Common Knowledge.
- von Mises, Richard (1981), *Probability, Statistics, and Truth*, Dover, New York.
- Wailly, Natalis de (1857), 'Mémoire sur les variations de la livre tournois, depuis le règne de saint louis jusqu'à l'établissement de la monnaie decimate.'', *Mémoires de l'Academie des Inscriptions et Belles-Lettre* **21**(11), 398–401.

Weber, Max (1946, 1918), *From Max Weber*, Oxford University Press, Oxford, chapter Politics as a Vocation.

Weber, Max (1968), *Economy and Society*, Bedminster Press, New York.

Wolfe, Martin (1972), *The Fiscal System of Renaissance France*, Yale University Press, New Haven.