On the Origins of National Identity

Felix Kersting (HU Berlin)
Nikolaus Wolf (HU Berlin)

Discussion Paper No. 217

December 13, 2019

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Dr. Klaus M. Schmidt, University of Munich, 80539 Munich, Germany
+49 (89) 2180 3405 | info@rationality-and-competition.de

On the Origins of National Identity*

Felix Kersting Nikolaus Wolf

December 13, 2019

Abstract

What are the origins of national identity? We extend the model by Alesina et al. (2019) to analyze the incentives of elites to use specific types of identity policies in response to shocks, and the extent to which such policies should be effective. To elicit changes in identity we use data on first names given in German cities between 1800 and 1875. We show that parents in cities treated by nation building policies responded by choosing first names of German origin for their children. To control for family-specific confounding factors, we exploit *within family* variation. We also show that the response can be conditional on cultural distance to the elite. Finally, Germanic first names had remarkable predictive power for behaviour. We find that individuals with Germanic first names made different marriage choices and were more likely to get actively involved and decorated during the German-French War in 1870/71 and the First World War.

JEL Classification:
Keywords:

*Kersting: Humboldt-University Berlin. Email: f.kersting@hu-berlin.de. Wolf: Humboldt-University Berlin, CEPR, and CESifo. Email: nikolaus.wolf@wiwi.hu-berlin.de. We would like to thank Robert Akerlof, Thilo Albers, Davide Cantoni, Neil Cummins, Theocharis Grigoriadis, Klaus Schmidt, Moses Shayo, and Thierry Verdier for helpful comments as well as seminar and conference participants at WEHC Boston, Heidelberg, Oxford, CRC Workshops Ohlstadt, Munich, and Schwanenwerder. We also would like to thank Jesper Zedlitz and his team for providing the loss lists of World War I, Walter Bauernfeind and the city archive of Nuremberg for as well as numerous other city archives for providing data on first names. Christopher Prömel provided excellent research support. Financial support by Deutsche Forschungsgemeinschaft through CRC TRR 190 is gratefully acknowledged

1 Introduction

Why are people willing to die for their country? Why would parents send their children to battle or sign war bonds even when the situation is hopeless? These and similar questions have fascinated social scientists at least since the writings of Durkheim, Veblen and Weber. The modern approach evolves around the concept of "social identity", namely the idea that people value their membership in social groups, and that such valuations are interdependent and changing over time. National identity is a type of social identity that stands out because it seems to evoke extreme emotions and forms of behaviour that only compare to religion. Moreover, national identity relies less than other types of social identity (such as gender or race) on physical categories and more on everyday cultural practise in the form of singing anthems, waving flags, or marching in parades. But when and why did people begin to do this?

Our starting point is the idea, formulated by Alesina et al. (2019), that elites can have an incentive to invest in policies to "homogenize" preferences in the population, e.g. via schooling or state propaganda. But clearly, such policies are not always successful. They can even backfire, see e.g. Fouka (2019). Moreover, the formation of national identities in the early 19^{th} century in many parts of Europe was not mainly about reducing the cultural or geographical distance to some existing elite. It was rather about the creation of an altogether new type of social identity, the collective imagination of a national community (Anderson, 1983). To address these aspects, we suggest a simple modification of Alesina et al. (2019). Here elites can invest in either homogenization or national identity policies. Under quite general assumptions, the incentive to invest in identity policies rather than in homogenization will be larger, the larger the cultural distance between elite and population. If so, some event that would suddenly increase this distance could foster identity policies, like some territorial change. Their success in turn would depend on the level of cultural distance to the elite.

Based on this simple framework we analyze the origin of a German national identity during the Napoleonic wars. There existed a German national movement since the last third of the 18^{th} century, carried by a new middle-class. But it was limited to small, intellectual circles. Only after 1813, when the Prussian elites started to appeal openly to national sentiments, did the movement spread to larger parts of society. Prussia had suffered a devastating defeat against France in 1806. What is more, the French Revolution had fundamentally shaken the legitimacy of the Ancien Regime, like the rule of the Hohenzollern over Prussia. In 1813 the Prussian elites adopted identity policies, not unlike those of Napoleon, to enhance ideas of a German nation, e.g. using speeches, pamphlets and theatre performances. They did so with the strategic aim to fight the French occupation, to regain legitimacy, and in 1815 to integrate the Rhineland and Westphalia into their enlarged territory.

We examine how changing exposure to Prussian identity policies during this time led to the formation of national identities on a larger scale. Our empirical approach is a

difference-in-difference framework at the level of families. We compare first names given to children in cities that became Prussian in 1815 to names in cities that stayed outside Prussia. In 1815 Prussia had gained large territories in the West, against her own intention, but due to a British intervention at the Congress of Vienna (Huning and Wolf, 2019). To elicit changes in identity within this setting, we focus on *within family* variation, which helps us to control for unobservable family characteristics. Using a large new dataset on name choices for families in 12 German cities, we show that exposure to Prussian policies in 1815 had a strong causal effect to increase names of Germanic origin. We can show that the same parents that neither had a national family tradition nor had given their children German names before becoming a Prussian subject, were much more likely to do so afterwards. However, we can also show that this effect depended on cultural distance between treated cities and the Prussian elites. Finally, we provide evidence that national first names had strong predictive power for individual behaviour. Hence, tracking first names is indeed a valid way to elicit the emergence and spread of a national identity. Men and women with national names in all German cities between 1800 and 1875 were more likely to marry a partner that also had a national given name. And men with national first names were more likely to be actively involved in the 1870/71 Franco-German war and the First World War and more likely to receive a medal for bravery.

Our setting in Germany in the early 19^{th} century is attractive for several reasons. First, it has been argued that in difference to France or England, the formation of a German national identity preceded the establishment of a nation state by several decades (Schulze, 1985, p. 58ff.). The Prussian elites appealed to ideas of a German nation for strategic reasons, which later helped to form a German nation state under Prussian leadership. Second, 19^{th} century Germany provides the researcher with a wealth of variety in terms of polities, geography, religious, ethnic and economic differences. This stems from the fact that the political unification of Germany was a slow process, which arguably started with Napoleon but was not finished until 1871. Third, German national identity is notorious for fostering – towards the end of the century – an aggressive nationalism, which contributed to the outbreak of the Great War and the ensuing rise of the NSDAP.

The major challenge for any historical study on social identity is empirical measurement. Our theoretical framework makes predictions about the effect of policies on the identity of individuals and individual behaviour. Identity can be seen as a latent variable, which in a historical context cannot be elicited by opinion polls or lab experiments. Also, there is little evidence to be gained from voting behaviour as the political franchise was very limited and voting infrequent. Hence, we have to find a way to capture identities indirectly, via choices that people made at their time. These choices should be observable at the individual level, they should have an obvious relation to identity and they should be frequent. Following earlier suggestions from historical sociology, notably Gleitze (1962), Lieberson and Bell (1992), Gerhards (1997), and Wolffsohn and Brechenmacher (1999), we will elicit social identities and their changes over time by the type of first names that parents gave their children. We classify names based on the philological five-volume com-

pendium on German first names by Seibicke (1996, 1998, 2000, 2002, 2007) into German national, religious, European, ancient and ruler first names.

Such an approach relies on two main assumptions. First, we need to assume that the choice of given names reflects the value that parents place on membership in a specific social group. Second, the valuation of parents needs to be at least partly transmitted to their children, be it directly by the parents themselves via education or indirectly, via their social environment. Both assumptions find very broad support from the recent literature on first names that also uses first names to capture social identity (Fryer and Levitt, 2004; Head and Mayer, 2008; Cook et al., 2014; Abramitzky et al., 2019; Fouka, 2019). Still, it is difficult to rule out various confounding factors that also might account for the name choices of parents, such as differences in education, social status or individual traditions of parents. This is a problem for modern data, but more so in a historical setting with limited data on control variables. To deal with this, we go beyond the existing literature in several ways. We systematically exclude all parents that had national names themselves and use for the remaining parents only *variation* in the type of name choice made by the same parents over time. Hence, we only need to assume that unobserved parent-level factors remain constant as long as the mother was in child-bearing age. Our evidence on the predictive power of first name type for the behaviour of children further suggests that name choice can indeed capture identity.

Our study is related to theoretical approaches in economics that aim to explain the formation of social identities. A common starting point is that individuals value their perceived distance from specific group prototypes, as suggested by Akerlof and Kranton (2000). Building on this, several authors have suggested mechanisms to account for changes in perceived distance and group status. The most prominent idea is to distinguish between an "elite" - a small group of agents with exogenous preferences (and identity) - and a larger group of agents with endogenous preferences. The elite can use policies and institutions such as pamphlets, state celebrations or schools to shape the preferences of the population (Alesina et al., 2017, 2019). Arguments along these lines rely on the assumption that elites have both, an incentive to form a specific social identity among the population and the capacity to influence identities. We contribute to this by modifying the framework by Alesina et al. (2019), such that elites can chose between policies to foster "national identity" and alternatives. Moreover, we show that the success of such policies will depend on the characteristics of the population. Another approach considers technological and economic change as a driver of identity change. In the literature on evolutionary cultural transmission, social traits simply spread between generations in line with reproductive success, and hence are fully determined by economic factors (Cavalli-Sforza and Feldman, 1981). In contrast, Bisin and Verdier (2001) show that the long-run outcomes depend crucially on whether cultural transmission within and outside the family are substitutes. Gellner (1983) argues that industrialization, accompanied by large-scale migration from villages to cities and the need for standardization devalued old regional identities and contributed to the rise of a new broader "national identity". Related, Shayo

(2009) develops a model, where group status depends among other things on relative income and perceived distances from group prototypes. Here, income growth and improved communication can lead to the spread of a "national identity", first among the middle-class and later among the poor (Shayo, 2009, p.156). In our study we show that exposure to elite policies matters, but also that its success will depend on characteristics of the population.

Next, the previous empirical literature has shown the impact of different dimensions of identity change on individual behavior (Hoff and Pandey, 2006; Benjamin et al., 2010, 2016). Battu and Zenou (2010) and Manning and Roy (2010) investigate determinants of identity formation. However, causal evidence on the determinants of changes of identity formation is scarce. A the notable exception is Fouka (2019), who provides evidence that forced assimilation policies in several US states that targeted the German minority backfired and led to less integration. We add to this literature evidence on the origin and spread of a "national identity"for the 19^{th} century, when arguably for the first time in history national identity became a mass phenomenon. Moreover, we improve identification using variation within families. We also provide an explanation for why identity policies are sometimes successful, and sometimes not.

The rest of our paper is organized as follows: we explain our theoretical framework in section 2. In section 3 we introduce the historical background to our study on the origins of a national identity in Germany. In section 4 we present and discuss our empirical strategy with the main identifying assumptions and the data we use. Section 5 contains our results on the treatment of 1815 for between and within family variation, including various robustness checks. We show that first names have strong predictive power for individual behaviour in section 6, where we first discuss evidence from marriage behaviour and next behaviour during the wars of 1870/71 and 1914/18. Section 7 concludes.

2 Theoretical Framework

When would elites have an incentive to manipulate the identity of their population? To guide our empirical analysis, we suggest a simple modification of the model by Alesina et al. (2019). A ruling elite faces a heterogeneous population, with individuals located at distance d_i from the location of the elite (the "capital"). Individual i's utility is given by

$$u_i = g(1 - d_i \frac{a}{\lambda_1}) + \lambda_2 + y - r \tag{1}$$

The first term $g(1 - d_i \cdot a)$ measures the perceived value of the government to individual i. The parameter g is the maximum utility an individual receives from the government at distance zero, where d_i is individual i's preference distance from the government. Following Alesina et al. (2019), distance can reflect geographical distance but also language, cultural, ideological difference between the individual and the public goods and policies provided by the government. The parameter a measures the cost of this distance. The term y

captures income (exogenously given), and r stands for taxes, which are split as lump sum tax between all individuals in the country. With increasing returns to scale (or scope) in the provision of public goods, the basic trade-off for any individual will be one between the benefits from size and the cost of heterogeneity. In a larger state individuals can benefit from cheaper provision of public goods but might suffer from higher (geographical or ideological) distance from their provision.

In (1) we have extended the model by Alesina et al. (2019) by assuming that the government (or elite) can chose between two types of identity policies, captured by the parameters λ_1 and λ_2. The government can either invest in specific homogenization policies λ_1 as in Alesina et al. (2019), or invest in a common "national" identity λ_2. The first type of policies would be aimed at reducing the distance to the existing elite directly. Examples would be infrastructure investment, but also specific education policies aimed at reducing the perceived distance to the government, such as history classes that glorify the ruling elite. We modify the model to introduce a second type of policies λ_2. These are attempts to create a new type of virtual public good, meant to reduce the relevance of any existing differences between individuals and the ruling elite. Examples would include propaganda against some common enemy, the invention of a common past between elite and all members of society, or the celebration of a national holiday.

Intuitively, the incentive for the elite to invest in either type of policy will depend on the cost of pursuing them (c_1, c_2), the cost of distance a, and the extent of economies of scale in providing public goods (see appendix). It is easy to show that both types of policies are substitutes, but not perfect ones. In particular, the benefit from pursuing "national" identity policies λ_2 instead of homogenization policies λ_1 is increasing in a, the cost of cultural distance between individuals and the elite. As we show in the appendix, an exogenous increase in distance d would have an equivalent effect to an increase in a.

In our context, the defeat of the Prussian army at the battle of Jena-Auerstedt 1806, or the spread of republican and anti-monarchist sentiment could have alienated the Prussian elite from their population. The Prussian elite would thereby have an incentive to invest in "national identity". The gain of new territories in 1815 should have further added to this incentive. Here, the authorities faced the challenge to establish the legitimacy of their rule and win the support of new populations that were geographically and culturally even further apart from the centre of the state.

To summarize, our framework delivers three testable hypotheses:

1. We expect to see that after a crisis of legitimacy or territorial gains elites will invest in the invention of a new "identity" rather than homogenization.

2. These policies should change *individual* identities towards national values.

3. Assuming a common level of optimal investment but a heterogeneous population, individuals will differ in their responsiveness to such policies depending on their distance to the elite.

We want to test these hypotheses in the context of Germany after 1800, where we observe for the first time that ideas of a German nation were spreading beyond small circles of intellectuals. In particular, we use the expansion of Prussian territory in 1815 as a shock to population heterogeneity, which was not anticipated (Huning and Wolf, 2019). We predict that in the new territories the Prussian elites intensified their efforts of nation building, both compared to other Prussian cities and to the efforts of nation building in other German cities outside of Prussia. In particular we predict that the Prussian government would have attempted to create a new narrative of the Prussian monarchy linked to a new "German" identity, instead of simple pro-Prussian propaganda. Moreover, we predict that some individuals will have responded to this and develop a new German identity. Finally, these responses will have been weaker the larger the cultural distance to the Prussian elite. The next section present historical background and first descriptive evidence on this.

3 Historical background

Our empirical study is focused on the years when the German national movement for the first time began to spread beyond small circles. To understand the historical context we will briefly describe the political situation at the time. Next we will discuss why the Prussian state had temporarily a strategic interest to appeal to national sentiment, and what policies were used to shape identities.

In the last third of the 18^{th} century, the Holy Roman Empire with the Emperor in Vienna still existed, but it was little more than an empty shell. The German lands were fragmented into several hundred territories ruled by various competing dynasties, church territories and free city states. The two leading powers were Habsburg and Prussia, both comprising large territories outside the Empire, and both striving for a more or less enlightened type of absolutist state. The dynasties of Habsburg-Lorraine and Hohenzollern had expanded their rule to the East and South, and around 1789 both states encompassed quite heterogeneous populations in terms of languages, religious denominations or levels of urbanization. A the same time, the old feudal order, with a division of power between the ruler and the estates was being replaced by a centralization of power, with an absolutist ruler supported by a modernized state bureaucracy.

This centralization of political power was, according to Schulze (1985, p. 240), accompanied in many societies in Europe by a crisis of loyalty, where old identities were destabilized due to political, economic and cultural change. The dramatic events in America 1776 and France 1789 intensified this loyalty crisis and were initially greeted by many intellectuals on the continent as the beginning of a new era. However, the initial enthusiasm quickly gave way to anti-French feelings. The occupation of large parts of Germany did not only foster modernization, e.g. by eliminating many small states, but also led to large and growing burden in terms of taxes and conscription for the war. Especially

the self-coronation of Napoleon as Emperor in 1804 and the humiliating defeat of Prussia in 1806 sparked a new type of German national sentiment, that was geared against the French occupation. Around the same time in Berlin formed a patriotic circle that included the publisher Georg Andreas Reimer, as well as Ernst Moritz Arndt, Friedrich Jahn, Friedrich Schleiermacher and the Prusssian officer August von Gneisenau (Bartmußet al., 2008). Notably Jahn became prominent as founder of the German gymnastics movement ("Turnbewegung") in 1811, which shortly spread to other German states and became the organizational backbone of the early national movement (Düding, 1984). Ernst Moritz Arndt wrote very popular pamphlets and poems such as "Was ist des deutschen Vaterland?" (1814), which claimed all territory where German was spoken for a new German national state.

But the national movement remained very small, and the authorities in the various German states hesitated to support it. Some states, like Prussia in 1806, had lost not only their independence and military power but also large parts of their territory, while others like Bavaria had been compensated for a loss of independence by territorial gains and upgraded status.[1] Moreover, the aristocracy all over Europe feared the liberal and revolutionary tendency among the national movements, as they could easily turn against their privileges, wealth, and indeed their life. Within this tension, a group of reformers among the Prussian elite realized that the national and anti-French sentiment could be used to the benefit of the Prussian state. A key motive was the mobilization of people for a levee en masse against the occupation. Already in 1797, Gerhard Scharnhorst, then officer in Hanover, observed that the French army had an advantage due to the messianic fighting spirit of the French soldiers. They were much more motivated to fight than the mercenaries or conscripted soldiers of regular troops (Dörner, 1995, p. 112). In 1801 Scharnhorst left Hanover to become an officer in the Prussian army. He worked as a military instructor, founded the Berlin Military Academy and quickly became influential, together with his disciples von Clausewitz and von Gneisenau. After the humiliating peace of Tilsit in 1807 he was promoted to major-general and led the effort of a fundamental reform of the Prussian army. He introduced new recruitment systems to circumvent the limitations imposed on Prussia since 1807 and strengthened the ties between army and population, e.g. by opening the career path of officers to non-noble families.

After Napoleons defeat in Russia in late 1812, Scharnhorst, Gneisenau and other reformers convinced the king of Prussia, Frederick William III to side with Russia (treaty of Kalisch, February 1813) and finally appeal to patriotism and the national movement. As a first step towards this, in February 1813 the king decreed that every man above 20 had to wear in public the national cockade as a sign of their patriotism (Hagemann, 2019, p.165). Next, when Prussia declared war to France in March 1813, the king made a proclamation "To my people", where he appealed for the very first time to national pride of "Germans and Prussians" to fight-off the foreign occupation. This proclamation was

[1] With French support, the Prince-elector Maximilian was crowned King Maximilian I. of Bavaria in January 1806.

published as a pamphlet and as an article in the leading newspaper of Breslau (today Wroclaw, where the king resided in 1813), and sent to all postal offices of Prussia at the time. Also in March 1813 the king established the Iron Cross, a new military decoration to be awarded for bravery without regard to rank nor social status.

The battle of Leipzig in October 1813 brought the military success over France that Prussia had hoped for, but the monarchy still had to face major challenges. At the international peace congress in Vienna 1814/15 Prussia had aimed for an annexation of the Kingdom of Saxony. Instead, due to a British intervention, Prussia gained large territories in the West (the Rhineland and Westfalia), which were disconnected from the Prussian mainland and had a predominantly catholic and often urban population (Huning and Wolf, 2019). This stood in contrast to the more rural and protestant Prussian mainlands and the protestant monarch Frederick William III himself. The main cities in the new territories like Cologne, Muenster, or Aachen had very different traditions as either free cities or catholic Prince-Bishoprics.

Hence, after the new territories had been formally incorporated into the Prussian state in 1815, the authorities faced the challenge to establish the legitimacy of their rule and win the support of their new population. A main instrument for this were royal proclamations and the official homage ceremonies, which took place in 1815. Characteristic for them was the attempt to create a new narrative with the king of Prussia leading the German nation. In the proclamation of 5 April 1815 to the inhabitants of the "Rhinelands unified with the Prussian monarchy" the king Frederick Wilhelm III states that at the congress of Vienna he agreed to the burden of defending the "endangered borderlands" out of respect for the "unified German fatherland" (Kotulla, 2010, p.591). Moreover, he gives the people of the Rhineland back to their "German fatherland and to an old German princely family" (Kotulla, 2010, p.591), and promises to honor their religion. This is in line with our first hypothesis: territorial change can provide elites with an incentive to invest into the invention of a "new identity".

In the same spirit, the Prussian authorities organized homage ceremonies in the newly gained territories, which stood in the tradition of medieval homage to the ruler, yet added quite explicitly a new element of national identity (Schwengelbeck, 2007, p.136ff). The ceremony for the Rhineland took place in Aachen on 15 May 1815, the ceremony for Westphalia in Muenster on 18 October 1815. The decision for Aachen was a reference to the former residence of Charlemagne. As General Gneisenau and Governor Sack stated in their official announcement, the city was chosen because "only the city of Aachen unites age, grandeur and suitable local with the dignity of a coronation city for the most elevated German rulers, where after Charlemagne no less than 35 German Emperors were crowned" (cited after Tschacher, 2010, p.259). The date in turn coincided with Pentecost, which had been abolished under French rule but was now being reintroduced as a religious holiday. The choice of date for the second ceremony in Muenster was a very explicit reference to national identity: it coincided with the second anniversary of the battle of Leipzig, which already then played a pivotal role for the collective memory, mainly due to the publications

by Ernst Moritz Arndt and Friedrich Jahn (Hagemann, 2002, p.481f). The orchestration for both ceremonies was similar: they were extended over several days and included the celebration of a mass in the cathedral, processions, fireworks, official banquets, feeding of the poor and theatre performances. Importantly, in both cases the authorities had invited representatives of all parts of society, including the nobility, clergy, burghers and peasants. For Aachen, there was also a common dress-code, where all civilians were obliged to wear black dress and the national cockade (Tschacher, 2010, p.259). In Aachen, the play "Deutsche Treue" by August Klingemann was performed, which placed the unification of the Rhineland with Prussia in the context of the reconciliation between Frederick the Fair of Austria and Louis IV (the Bavarian) in 1325 after their struggle for the German crown. Afterwards, the celebrations were popularized in a flood of pamphlets and newspaper articles, most prominently a serial written by Ernst Moritz Arndt (Tschacher, 2010, p.267).

The celebration in Muenster a few months later was quite similar, but here the association between the king of Prussia and a new German identity was made even more explicit. The festivities took place over several days, they were elaborate and rather expensive (Lahrkamp, 1976, p.128). Apart from the fact that the date, exactly two years after the battle of Leipzig was a clear reference to the idea of a "German nation", the chosen theatre play "Der Altar im Walde" made it even more obvious. Here, the king of Prussia is called the new Hermann, the legendary chieftain of the Germanic Cherusci tribe who fought the Roman invasion. The topic had been popularised as an allegory to the French occupation since Arndt asked in 1805 for a "new Hermann" to fight Napoleon and Heinrich von Kleist's drama "Die Hermannsschlacht" (1808). Now in Muenster 1815, king Frederick Wilhelm III is depicted as the new Hermann, saviour of the German nation. The author of the play was Wilhelm Aschenberg, a lutheran pastor and writer from the nearby city of Hagen. In 1814 Aschenberg had founded the journal "Hermann" that propagated ideas of a pro-Prussian German nationalism, along the lines that would suit the Prussian government. As in the case of Aachen, the celebration was followed by several articles in newspapers and special publications (e.g. a commemoration book published by Joh. B. Bodde in 1816), to spread the word.

Hence, the Prussian authorities clearly attempted to create a new narrative to show the king of Prussia as the leader of the German nation, particularly in the new provinces in line with the first hypothesis from our theoretical framework. Their intention was purely strategic, with the aim to use the broader idea of a German nation to turn the new population into loyal subjects of the king of Prussia. And indeed, already from 1818, when the territorial order was established did the Prussian authorities try to silence the national movement again. But to what extent did their strategy succeed in 1815? How can we test if individuals that experienced these celebrations in 1815 or had heard and read about these events, actually changed their identity, and behaviour?

4 Empirical Strategy and Data

In order to determine the causal effect of Prussian policies in 1815 on the formation of national identities (our second hypothesis), we use evidence on families in a differences-in-differences framework, and exploit within-family variation.[2]

Our *treated* families live in Aachen (in the Rhineprovince) and Hagen (in Westphalia), two cities that become part of Prussia in 1815. Our *control* group consists of families who live in Mannheim, Heidelberg, Hanover and Frankfurt, i.e. cities that were not part of Prussia. To estimate the causal effect of becoming part of Prussia on "identity" we use first names that parents decided to give their children in different cities and at different points in time. We assume here that the choice of given names reflects the value that parents place on membership in a specific social group. We also assume the valuation of parents needs to be at least partly transmitted to their children, be it directly by the parents themselves via education or indirectly, via their social environment. Both assumptions find very broad support from the recent literature on first names that also uses first names to capture social identity (Fryer and Levitt, 2004; Head and Mayer, 2008; Cook et al., 2014; Abramitzky et al., 2019; Fouka, 2019).

The two cities in our treatment, Aachen and Hagen, provide us with variation regarding distance to the elite (our third hypothesis). In 1815 Aachen was a predominantly catholic city, in difference to the protestant mainlands of Prussia, whereas Hagen was predominantly protestant. The city of Aachen had played an outstanding role in the history of the old Holy Roman Empire, as the place where the Emperor was crowned German king until 1531. The city was an Imperial free city until the French occupation in 1794, and became Prussian in 1815. The town of Hagen had a less glorious past. Importantly, Hagen as a settlement in the county of Mark became part of Brandenburg-Prussia from 1666 onwards. It received town rights in 1746, and started to grow. In 1807 the Mark and with it the city of Hagen passed from Prussia to France before it returned to Prussia in 1815. Hence, we would expect to see that efforts of any type of identity policy by the Prussian elite might meet a more fertile ground in Hagen compared to Aachen, as the two cities differed significantly in terms of their cultural distance to Prussia. Note that territorial change as such will not explain our findings, because during our period of observation this affected both, cities in our treatment group and in the control group.

However, just comparing the frequency of name choices in treated versus non-treated cities could be misleading. There is a host of confounding factors that might also account for the name choices of parents, such as differences in education, social status or individual traditions of parents. This is a problem for modern data, but more so in a historical setting with limited data on control variables. We address these problems by systematically

[2]Previous literature using within–family variation focused on the quantity-quality trade-off for fertility decisions. Multiple births and preferences for balanced gender composition are used in this research as instruments for fertility (Angrist et al., 2010; Black et al., 2005; Cáceres-Delpiano, 2006; Rosenzweig and Zhang, 2009).

excluding all parents that had national names themselves. Moreover, our data allows us to use family-fixed effects: we use for the remaining parents only *variation* in the type of name choice made by the same parents over time. This implies that we restrict our attention on families who get at least one children before and after the treatment. Exploiting within–family variation has several main advantages. First, this strategy allows us to control for all time–invariant family characteristics. Second, we only compare families in similar age groups as we are only interested in those families that get children in the years before and after the specific year. Third, we can rule out that migration into Prussian cities drive our results because we only consider families who already live in the respective city before the treatment year and stay there after the treatment.

We use the following differences-in-differences estimation employing an OLS model[3]:

$$NationalName_{fct} = \alpha_f + \theta_t + \sum_{t=1805}^{1824} \beta_t (TreatedCity_c \cdot Year_t) + \epsilon_{fct} \qquad (2)$$

$NationalName_{fct}$ is a dummy variable for a national first name in family f in city c at time t. Our dependent variable has the value of 1 only when a child is given a national first name, and zero otherwise. α_f are family fixed effects and θ_t year fixed effects. The coefficient of interest is β_t that indicates the effect of living in a treated city in $Year_t$ year. With this specification, we control for time invariant family characteristics as well as more general time trends. The variation we exploit only comes from families who either change from a non-national first name to national first name for their children or vice versa. We allow β to vary over time and thus are able to control for potential treatments. As second specification, we follow Bertrand and Mullainathan (2004) to control for serial auto-correlation and collapse our sample in a pre- and post-treatment period:

$$NationalName_{fct} = \alpha_f + \beta (TreatedCity_c \cdot Post1815) + \gamma Post1815 + \epsilon_{fct}, \qquad (3)$$

where $Post1815$ equals to 1 for the post-treatment period and 0 otherwise. The coefficient of interest in β. We will use this as our main specification and discuss the credibility of our approach in the following section. Our dependent variable equals to 1 if a family has at least one child with national first name in one period. In addition, we use similar and less restrictive specifications where we use city fixed effects instead of family fixed effects. With this less restrictive approach we have much more observations, but a less-clear-cut identification and obviously a lack of control variables.

The identifying assumption in this setting implies that the decisions by parents regarding first names in cities that become part of Prussia and in cities that do not become part of Prussia would follow the same trend in absence of the treatment. We will show that indeed is the case for our treatment analysis. Given the course of events after the defeat of Napoleon in 1813/14, and Prussia's negotiations at the congress of Vienna (Huning and Wolf, 2019), the territorial change that occurred in 1815 was difficult to

[3]Note that we present results using a Logit model as robustness.

anticipate.

In order to measure national identity using first names, we use several new data sets, which we will describe in the following. Moreover, we introduce our way to classify the first names into different social categories.

Birth Registers of German Cities Our main data is based on births registers for 12 German cities based on city archives and several genealogy websites, which we had typed in. Thereby, we construct a sample with almost 1.2 million observations. For these births, we have the following information: first and last name for children and their parents, as well as place and year of birth of the children.

Table B.1 in the Appendix provides an overview about our data set, including the date the city became part of the Kingdom of Prussia. Note, that we have four different groups within our sample: first, cities like Berlin that belonged to the Kingdom of Prussia throughout the 19th century. Second, cities in Westphalia and the Rhine Province that became part of Prussia after the Peace of Vienna in 1815. Third, cities that became part of Prussia after the Prussia-Austria War in 1866. Fourth, cities that belonged to the German Empire after its foundation in 1871. We exploit this variance in terms of national allegiances later in our empirical analysis.

For our main identification strategy based on within-family variation, we construct a common family ID for children with the same parents based on the first name of the father, the first letter of the first name of the mother, and the last names of father and mother.

In the case of Aachen, we have to conduct some adjustments due to the French occupation since 1794. The French bureaucracy probably "francicised" many German names, although we have no direct evidence on this (Kramer, 1993, p.225). In the case of the parents' first names, this is straightforward as it is unlikely that the parents were given French names in late 18^{th} century Aachen before the French occupation. Thus, we use the German translation of the French first names (e.g. Guillaume/Wilhelm). Reassuringly, we get a better matching of the family IDs, as we find many more couples with the same family name and matching first names before and after 1815. For the children, however, this becomes more complicated because we do not know whether parents adjusted to the new rulers by choosing French first names or instead the French administration "francicised" the German names. Such adjustments are potentially important for our results. If we do not use the German versions of the French names for children born between 1794 and 1814, we might underestimate the share of Germanic first names for the pre-treatment period and thus potentially overestimate our treatment effect. To circumvent this problem, we present the results for both, the unadjusted and the most "conservative" option. In the conservative option, we assume that the French administration systematically "francicised" German first names. We lack direct evidence on this, but if anything this introduces a

bias *against* us finding a treatment effect.[4] Note that we do not face a similar problem for families in the city of Hagen.

Our data set on families also allows us to trace the first names of marrying couples over time and in different parts of Germany, conditional on them having children. We will use this below as one piece of evidence (among others) on the predictive power of name types for individual behaviour.

Loss Lists and Honored Soldiers from Wars Furthermore, we collect data from loss lists and lists with honored soldiers. We use this data to show the "predictive power" of national first names. More specifically, we use the digitalized German loss lists for the German-French War in 1871 as well as World War I provided by Verein für Computergenealogie (2014). The loss lists include information about the first and last name, the location, the regiment, the year the soldier was listed on the loss list, and contain around 160.000 entries for the German-French War in 1870/71 and more than 8.5 million entries for World War I. The loss lists indicate not only participation in the war, but active participation, as they refer to soldiers that were either killed, wounded, captured or went missing. However, none of these events needs to reflect voluntary behaviour of soldiers.

In order to capture voluntary engagement during wars, we use lists with honored soldiers. For the German-French War from 1870/71, we digitalize the names of all 1295 soldiers honored with the *Eiserne Kreuz* (Königliche General-Ordens-Kommission, 1878) and categorize them based on four status groups. For World War I, we rely on Ophaus (1936). His book provides information on all non-commissioned officers (*Unteroffiziere*) who were honored with the *preußische goldene Militär-Verdienst-Kreuz* which was the highest honor for soldiers with this rank. Overall, 1773 soldiers got this award during World War I.[5] The basic idea here is that the distinction of soldiers for bravery should be a much better indicator for voluntary behavior than active war participation as reflected on the loss list. We might expect that soldiers with a strong national(ist) identity would be more willing to fight and hence more likely to receive the Iron Cross than others. If Germanic first names would be a good predictor for national identity, we expect to find a positive correlation between names and honors.

Classification To classify the first names in general and national first names in particular, we use a five volume encyclopedia on German first names from Seibicke (1996, 1998, 2000, 2002, 2007). We classify all first names that come up at least 300 times in the 19^{th} century. By doing so, we capture around 95% of all entries in our sample. We differentiate between national, religious, European and ancient first names. To make sure that we only measure these facets of identity and not conflated parts of identity, we furthermore

[4]For instance, we do not observe a break in our data once the French administration left Aachen in January 1814.

[5]Ideally, we would have the first names of the volunteers in World War I. According to some estimates, around 260.000 soldiers volunteered in August 1914. However, these lists were destroyed in World War II.

account for family and ruler tradition. We measure family tradition by counting those who have the same first name as their parents and we account for ruler tradition with first names from local rulers during their tenure as well as Friedrich and Wilhelm as rulers of Prussia for cities that were part of Prussia throughout the 19^{th} century.

We count those first names as Germanic whose origin lies in old high German (*althochdeutsch*) or germanic (*germanisch*) language, according to Seibicke. We cross check our list of national first names with a book from Khull (1909) published around 1900 that aimed to increase the consciousness for Germanic first names in order make sure that the names on our list had a nationalist connotation around 1900. The most popular national first names can be found in Table B.2 in the Appendix. Note, that the distribution of the frequencies for the first names is highly unequal, even more so for men where a four first names (Carl/Karl, Friedrich, Heinrich and Wilhelm) make up about 25% of all males in our sample.

In the Appendix in Figure B.5, we apply these steps to our data and summarize the shares for each city to give an overview of the trends in Germany throughout the 19^{th} century. Religious first names have the highest share among our groups. Figure B.4 shows the development of the shares of national first names from 1810 to 1874 conditional on family tradition and ruler first names. We observe an strong upward trend especially between 1810 and 1830. Afterwards, the share of national first names slowly increases. In addition, we observe also a steady increase in the share of Hermann and Ernst, two prominent nationalist first names not related to rulers, among boys (also in Figure B.4). This descriptive analysis provides first evidence that the first decades of the 19^{th} century were indeed crucial for the spread of national identity in Germany. Moreover, we see that the share of national first names increase which fits to the historiography according to which national identity becomes more salient over the 19^{th} century.

5 The Effect of Nation-Building

As a first result, we show the development of the share of national first names in our treated and control cities in Appendix Figure B.3. The share of national first names slightly increases from around 5% to 8% in our control cities, whereas the share of national first names more than doubles from 12% to around 20% in the treated cities.

Before presenting the results for our regressions, we present evidence in Appendix Figures B.1 and B.2 and in Appendix Table B.3 on the common trend assumption. The coefficients based on equation 2 (with 1814 as reference year) for the pre-treatment period are all statistically insignificant, independent whether we rely on city or family fixed effects. Between 1811 and 1814, we see quite some variation in the point estimates (especially in case we use family fixed effects), however, no upward trend is visible. Overall, this suggests that the common trend assumption holds for our estimation despite losing a lot of observations because this step only includes families that have at least one children

before *and* after 1815.

The results from a difference-in-difference regression using city fixed effects confirm this impression (Table 1 column 1). The coefficients are positive and significant. The result holds after controlling for family fixed effects (column 3). The point estimate suggests that families in treated cities choose a national first name for their child with a 30 percentage point higher probability. The event study graphs (Figures B.1 and B.2) show the yearly effects. Using city effects the effect size remains stable over time (with 1816 as the only non-significant result). The point estimates of our preferred regression exploiting only within-family variation seem to decline over time. Note that the standard errors are quite large, likely due to the small sample size for each year.

In a next step, we allow for heterogeneous effects for our two treated cities, Aachen and Hagen, to account for the difference in terms of "cultural distance" to Prussia. We find strong positive effects for both cities, especially Aachen in column 2 and 4 in Table 1. We discuss the issue of "francicised" first names in Aachen when we discuss several robustness checks.

Table 1: Treatment Analysis

	(1)	(2)	(3)	(4)
	Dep. Var. Dummy National First Name			
DiD Coefficient...				
...overall	0.0869***		0.293**	
	(5.74)		(3.99)	
...Aachen		0.0688***		0.388***
		(23.97)		(13.71)
...Hagen		0.107***		0.212***
		(37.40)		(7.48)
After 1815	0.0350***	0.0350***	0.0939**	0.0939**
	(12.21)	(12.20)	(3.32)	(3.32)
City FE	✓	✓		
Family FE			✓	✓
Mean Dep. Var.	0.1936	0.1936	0.1936	0.1936
Observations	77328	77328	2422	2422

Notes: * p<0.1, ** p<0.05, *** p<0.01. We only include parents without national first given name. Clustered standard errors at the city level. Treated cities: Aachen and Hagen. Control cities: Frankfurt (Main), Hannover, Heidelberg, and Mannheim. Results based on equation 3.

Further Robustness As further robustness checks, we use a Logit estimation in Appendix Table B.4 to account for a dummy variable as dependent variable. The results remain highly significant and positive.

Next, we further restrict our analysis conditional on family tradition, i.e. we exclude families that choose a middle name from the parents as first name of the child in Appendix Table B.5, Panel 1. The coefficient of interest slightly declines, but remains

highly significant.

We also test whether the effect may only be due to a specific Prussian "ruler" effect. Therefore, we exclude all children named Wilhelm after 1815 in Panel 2 in Appendix Table B.5. Our results only change slightly.

The identity policies in 1815 in the new provinces were location-specific as discussed in section 2 above, and should have affected the centre less than the treated cities. Therefore, we use in Panel 3 of appendix Table B.5 Berlin as control group. As before, we find a strong treatment effect for families in Hagen and Aachen.

However, the strong result for Aachen crucially depends on how we adjust for "francicised" names before 1815. In Appendix Table B.6 we show that with a very conservative adjustment, assuming that *all* first names were "francicised" and use the German translation thereof, the data suggests no positive treatment effects for Aachen and also no positive overall effect. The effect for Hagen remains positive and highly significant.

Nation-building, 1866 Finally, consider the territorial expansion of Prussia in 1866, when Frankfurt and Hanover became part of the Prussian state. In difference to 1815, the enlargement of Prussia did not have the connotation of German unification, but rather of Prussian power politics geared against the idea of a German national identity, especially due to the German-Austrian War (Schulze, 1985, p.235f). Here, the incorporation of the new population was not accompanied by a similar national propaganda as in 1815, so we would expect to find no effect for the treated cities in 1866. Indeed, as shown in Table B.7, we do not find similar positive treatment effects for 1866 when Frankfurt and Hanover became part of Prussia. However, we see that the national idea was generally spreading after 1866, albeit not related to Prussia.

To summarize, the empirical evidence supports our three hypotheses. First, we find that after the challenges of 1806/07 the Prussian authorities clearly attempted to create a new narrative to show the king of Prussia as the leader of the German nation. In particular, they attempted to invent a new German identity in the new provinces in 1815 in line with our first hypothesis. Next, we see that their success to affect individual identities was mixed. We find an overall positive and significant treatment effect as documented in column 1 and 3 in Table 1. Moreover, we always find very strong evidence for a positive treatment effect for families in the city of Hagen. In this case, the effect remains strong even after controlling for family fixed effects (which reduces the number of observations in our sample from over 75000 to below 2500). This supports our second hypothesis, namely that identity policies can indeed change individual identities. Our results for Aachen are however less clear, as they depend on our assumption about the administrative "francification" of first names. Hence, our evidence on the third hypothesis about the relevance of cultural distance is ambiguous.

6 The Behavioral Effects of First Names

So far, we provided evidence on the origins of national identity after the collapse of Napoleon. However, we cannot infer from these results whether or to what extent national first names were also associated with nationalist behavior. The argument here is not that national first names would change individual behaviour. Rather, we suggest that national first names can reflect a social environment, which put a high value on national identity and thereby helped to shape the identity and behaviour of a child. In order to provide evidence for this idea, we look at the predictive power of national first names in two arguably important and distinct aspects of life: marriage and war.

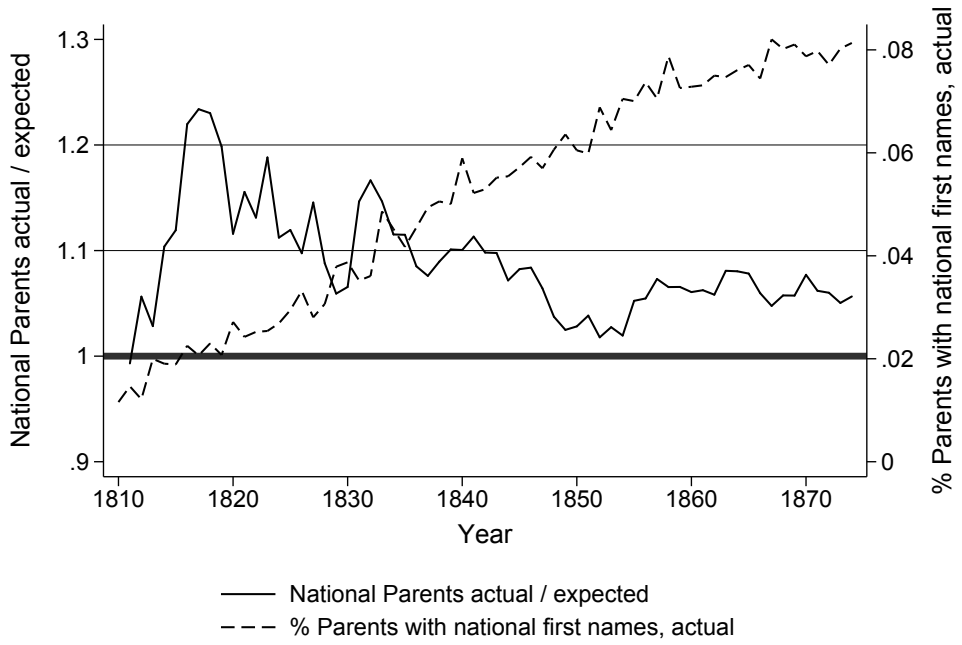
First Names and Marriages

Our data allows us to analyze whether parents with a national first name are more likely to marry a person with a national first name. The basic idea is to test, whether two people of different sex, living in the same city in the same year had a higher than random chance to marry if both partners carried a national first name. If indeed names capture the social environment in which children grow up, we should see that men and women sharing a name type are more likely to marry compared to random matching. To this end, we compute the actual and expected share of parents who both bear a national first name and test this for statistical differences. We calculate the expected share of national marriages, based on the share of men and women in a city at a point in time assuming random matching. In Figure 1 we plot the ratio of actual over expected "national" couples, where both bear a national first name (left axis). We show averages over all cities over time, based on calculations for each city and year separately. We see substantial differences in the first decades of the 19^{th} century, which are declining but not disappearing thereafter. The early differences are based on few observations, because national names were still rare, and more so couples where both partners had a national first name. A possible interpretation for the observed decline that follows is that the overall share of national names in the population is increasing (right axis), such that the signalling function of the national name-type is getting weaker. Overall, we interpret this finding as evidence for assortative mating based on national identity.

First Names and Behavior in World War I

Next, we compare the first names of honored soldiers with all soldiers who appear on loss lists of the First World War. By analyzing honored soldiers, we aim to capture strong engagement during the German-French War in 1870/71 and World War I. Comparing the loss lists and the honored soldiers in Table 2 shows overall substantial and statistically significant differences. More specifically, we find differences between the share of national first names for the soldiers on the loss lists and those who were awarded during the German-French War in 1870/71, also if we condition the analysis on different hierarchy levels (panel

Figure 1: National marriages, 1810-1874



Note: Balanced panel for 1810-1874. Includes: Frankfurt (Main), Hannover, Nuernberg, Hagen, Berlin, Aachen, Mannheim and Heidelberg.

1). Note that the differences are very sizeable. The difference in terms of the share of national first names between *Offiziere* on Loss Lists and Award Lists is more than 30 percentage points. We find similarly large differences for other hierarchy levels. Analyzing more specific national first names, we also find differences. Note that the relative difference is more pronounced for Ernst and Hermann – well-known nationalist first names – than for Friedrich and Wilhelm – names with strongly overlapping connotations, because these were also the most prominent ruler first names of that time. For World War I, we have sufficient data at the city level to compare first the frequency of names in the population, next to the loss list and finally to honored soldiers (panel 3). Moreover, we can condition on last names to improve the matching between our samples. We find similar results compared to 1870/71: 44% of all soldiers on the loss lists have a national first names, comparing to 51% of the awarded non-commissioned officers. Again, we can rule out that the results are mainly driven by ruler first names like Friedrich and Wilhelm (panel 2).

We conclude form this that first names can indeed capture aspects of identity that are relevant for individual behaviour, as documented for marriages and war participation. Again, we do not claim that the names as such shape behaviour. Rather they reflect circumstances or social environments that shape behaviour later on.

Table 2: National First Names, War Participation, and War Decoration

	Panel 1: German-French War 1870/71		
	Share national first names...		
	...Loss Lists	...Honored Soldiers	t-Test
National First Names	0.309	0.553	17.739***
... *General*	0.190	0.604	14.106***
... *Offiziere*	0.200	0.532	20.503***
... *Unteroffiziere*	0.333	0.579	5.864***
... *Mannschaft*	0.314	0.625	2.489**
Friedrich & Wilhelm	0.109	0.140	3.243***
Ernst & Hermann	0.029	0.057	4.297***
	Panel 2: Honored Soliers and World War I		
	Share national first names...		
	...Loss Lists	...Honored Soldiers	t-Test
National First Names	0.441	0.512	5.9109***
Friedrich & Wilhelm	0.112	0.111	0.1239
Ernst & Hermann	0.055	0.064	1.5796*
	Panel 3: Participation World War I		
	Share national first names...		
City	...Loss Lists	...Honored Soldiers	t-Test
No restrictions, national first names			
Hannover	0.65	0.64	39924
Mannheim	0.36	0.48***	18806
Nürnberg	0.31	0.37***	21555
No restrictions, national first names excluding ruler names			
Hannover	0.23	0.27***	39924
Mannheim	0.12	0.18***	18806
Nürnberg	0.10	0.17***	21555
Conditioning on last name, national first names			
Hannover	0.65	0.65	29992
Mannheim	0.36	0.46***	8175
Nürnberg	0.31	0.37***	20985
Conditioning on last name, national first names excluding ruler names			
Hannover	0.23	0.28***	29992
Mannheim	0.12	0.15***	8175
Nürnberg	0.10	0.18***	20985

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

7 Conclusion

In this paper we asked about the origins of national identity. We extended a model by Alesina et al. (2019) to analyze how exposure to national identity policies had a causal effect on individual identity and consequences for behaviour. The model generated testable hypotheses about the incentives of elites to use either homogenization or to invent a new "national" identity in response to shocks, and the extent to which such policies should be effective. To elicit changes in identity we used data on first names given in German cities between 1800 and 1875. We have shown that the Prussian state indeed attempted to create a "national" identity in the newly gained territories. We also showed that parents in cities treated by nation building policies of the Prussian state in 1815 responded by choosing first names of German origin for their children. To control for family-specific confounding factors, we exploited *within family* variation. We also showed that Germanic first names had remarkable predictive power for behaviour: individuals with Germanic first names made different marriage choices and were more likely to get actively involved and decorated during the German-French War in 1870/71 and the First World War. We think that our approach can be easily applied to many other contexts. Using name choices as indicators for identity change, and exploiting variation within families to control for unobservable characteristics rests on data, which is available in many contemporary and historical settings. Our findings based on the nation-building policies of Prussia around 1815 suggest that choices of first names react strongly to policy changes, and have remarkable predictive power for individual behaviour.

References

Abramitzky, R., L. Boustan, and K. Eriksson (2019). Do immigrants assimilate more slowly today than in the past? *American Economic Review: Insights Forthcoming*.

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *Quarterly Journal of Economics 115*(3), 715–53.

Alesina, A., P. Giuliano, and B. Reich (2019). Nation-building and education. *NBER Working Paper No 18839*.

Alesina, A., B. Reich, and A. Riboni (2017). Nation-building, nationalism and wars. *NBER Working Paper No 23435*.

Anderson, B. (1983). *Imagined Communities. Reflections on the Origin and Spread of Nationalism*. London: Verso Books.

Angrist, J., V. Lavy, and A. Schlosser (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics 28*(4), 773–824.

Bartmuß, H.-J., E. Kunze, and J. Ulfkotte (2008). *"Turnvater Jahn" und sein patriotisches Umfeld. Briefe und Dokumente 1806–1812.* Köln: Böhlau Verlag.

Battu, H. and Y. Zenou (2010). Oppositional identities and employment for ethnic minorities: Evidence from england. *The Economic Journal 120*(542), 52–71.

Benjamin, D. J., J. J. Choi, and G. Fisher (2016). Religious identity and economic behavior. *Review of Economics and Statistics 98*(4), 617–637.

Benjamin, D. J., J. J. Choi, and A. J. Strickland (2010). Social identity and preferences. *American Economic Review 100*(4), 1913–1928.

Bertrand, M. and S. Mullainathan (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review 94*(4), 991–1013.

Bisin, A. and T. Verdier (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory 97*(2), 298–319.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005, may). The more the merrier? the effect of family size and birth order on children's education. *The Quarterly Journal of Economics 120*(2), 669–700.

Cáceres-Delpiano, J. (2006). The impacts of family size on investment in child quality. *Journal of Human Resources XLI*(4), 738–754.

Cavalli-Sforza, L. L. and M. W. Feldman (1981). *Cultural Transmission and Evolution*. Princeton: Princeton University Press.

Cook, L. D., T. D. Logan, and J. M. Parman (2014). Distinctively black names in the American past. *Explorations in Economic History 53*, 64–82.

Dörner, A. (1995). *Politischer Mythos und symbolische Politik: Sinnstiftung durch symbolische Formen am Beispiel des Hermannsmythos*. Opladen: Westdeutscher Verlag.

Düding, D. (1984). *Organisierter gesellschaftlicher Nationalismus in Deutschland*. München: Oldenbourg Wissenschaftsverlag.

Fouka, V. (2019). Backlash: The unintended effects of language prohibition in US schools after World War I. *Review of Economic Studies Forthcoming*.

Fryer, R. G. and S. D. Levitt (2004). The causes and consequences of distinctively black names. *Quarterly Journal of Economics 119*(3), 767–805.

Gellner, E. (1983). *Nations and Nationalism*. Ithaca [NY]: Cornell University Press.

Gerhards, Jürgen; Hackenbroch, R. (1997). Kulturelle Modernisierung und die Entwicklung der Semantik bei Vornamen. *Kölner Zeitschrift für Soziologie und Sozialpsychologie 49*, 410–439.

Gleitze, B. (1962). *Beitrag zur Geschichte der kulturellen Integration Deutschlands. Untersucht an der ost- und westdeutschen Vornamensentwicklung 1750-1850*. Würzburg: Holzner-Verlag.

Hagemann, K. (2002). *"Mannlicher Muth und Teutsche Ehre". Nation, Militär und Geschlecht zur Zeit der Antinapoleonischen Kriege Preußens*. Paderbron: Schöningh.

Hagemann, K. (2019). *Umkämpftes Gedächtnis: Die Antinapoleonischen Kriege in der deutschen Erinnerung*. Paderborn: Schöningh.

Head, K. and T. Mayer (2008). Detection of local interactions from the spatial pattern of names in France. *Journal of Regional Science 48*(1), 67–95.

Hoff, K. and P. Pandey (2006). Discrimination, social identity, and durable inequalities. *American Economic Review 96*(2), 206–211.

Huning, T. and N. Wolf (2019). How Britain unified Germany: Endogenous trade costs and the formation of a customs union. *CEPR Discussion Paper No. 13634*.

Khull, F. (1909). *Deutsches Namesnbüchlein. Ein Hausbuch zur Mehrung des Verständnisses unserer heimischen Vornamen und zur Förderung deutscher Namesngebung*. Berlin: Verlag des Allgemeinen Deutschen Sprachvereins.

Königliche General-Ordens-Kommission (1878). *Königlich preussische Ordensliste 1877*. Berlin.

Kotulla, M. (2010). *Deutsches Verfassungsrecht 1806-1918. Eine Dokumentensammlung nebst EinfÃ¼hrungen, 3. Band*. Heidelberg: Springer.

Kramer, J. (1993). Französische personen- und ortsnamen im rheinland 1794–1814. In W. Dahmen (Ed.), *Das Französische in den deutschsprachigen Ländern*, pp. 222–236. Tübingen: Narr.

Lahrkamp, M. (1976). *Quellen und Forschungen zur Geschichte der Stadt Münster. Neue Folge 7/8. Münster in Napoleonischer Zeit 1800-1815*. Münster: Aschendorff.

Lieberson, S. and E. O. Bell (1992). Children's first names: An empirical study of social taste. *American Journal of Sociology 98*(3), 511–554.

Manning, A. and S. Roy (2010). Culture clash or culture club? national identity in Britain. *The Economic Journal 120*(542), F72–F100.

Ophaus, F. J. (1936). *Das preußische goldene Militär-Verdienst-Kreuz*. Berlin: Siegismund.

Rosenzweig, M. R. and J. Zhang (2009). Do population control policies induce more human capital investment? Twins, birth weight and China's "one-child" policy. *Review of Economic Studies 76*(3), 1149–1174.

Schulze, H. (1985). *Der Weg zum Nationalstaat: Die deutsche Nationalbewegung vom 18. Jahrhundert bis zur Reichsgründung*. München: Deutscher Taschenbuch Verlag.

Schwengelbeck, M. (2007). *Die Politik des Zeremoniells. Huldigungsfeiern im langen 19. Jahrhundert*. Frankfurt am Main: Campus.

Seibicke, W. (1996). *Historisches Deutsches Vornamenbuch, Band 1, A - E*. Berlin: Walter de Gruyter.

Seibicke, W. (1998). *Historisches Deutsches Vornamenbuch, Band 2, F - K*. Berlin: Walter de Gruyter.

Seibicke, W. (2000). *Historisches Deutsches Vornamenbuch, Band 3, L - Sa*. Berlin: De Gruyter.

Seibicke, W. (2002). *Historisches Deutsches Vornamenbuch, Band 4: Sc-Z*. Berlin: De Gruyter.

Seibicke, W. (2007). *Historisches Deutsches Vornamenbuch: Band 5: Nachträge und Korrekturen*. Berlin: Walter de Gruyter.

Shayo, M. (2009). A model of social identity with an application to political economy: Nation, class, and redistribution. *American Political Science Review 103*(2), 147–174.

Tschacher, W. (2010). *Königtum als lokale Praxis*. Stuttgart: Franz Steiner Verlag.

Verein für Computergenealogie (2014). Verlustlisten Erster Weltkrieg.

Wolffsohn, M. and T. Brechenmacher (1999). *Die Deutschen und ihre Vornamen: 200 Jahre Politik und öffentliche Meinung*. München: Diana.

Appendix

A Theoretical Framework

When would elites have an incentive to invest in national identity? To guide our empirical analysis, we suggest a simple modification of the model by Alesina et al. (2019). A ruling elite faces a heterogeneous population, with individuals located at distance d_i from the location of the elite (the "capital"). Individual i's utility is given by

$$u_i = g(1 - d_i \frac{a}{\lambda_1}) + \lambda_2 + y - r \tag{A.1}$$

The first term $g(1 - d_i \cdot a)$ measures the perceived value of the government to individual i. The parameter g is the maximum utility an individual receives from the government at distance zero, where d_i is individual i's preference distance from the government. Following Alesina et al. (2019), distance can be language, cultural, ideological difference between the individual and the public goods and policies provided by the government. The parameter a measures the cost of this distance. The term y captures income (exogenously given), and r stands for taxes, which are split as lump sum tax between all individuals in the country.

We extend the model by Alesina et al. (2019) by assuming that the government (or elite) can chose between two types of identity policies, captured by the parameters λ_1 and λ_2. The government can either invest in specific homogenization policies λ_1 as in Alesina et al. (2019), or in a new common "national" identity λ_2. The first type of policies would be aimed at reducing the distance to the existing elite directly. Examples would be infrastructure investment, but also specific education policies that attempt to reduce the perceived distance to the government, such as history classes that glorify the ruling elite. We modify the model and introduce a second type of policies λ_2. These are attempts to create a new type of virtual common good, meant to reduce the relevance of any existing differences between individuals and the ruling elite. Examples would include propaganda against some common enemy, or the invention of a common past between elite and all members of society.

For a country with mass 1, the government budget constraint is given by

$$r = k + c_1 \cdot \lambda_1 + c_2 \cdot \lambda_2 \tag{A.2}$$

where k stands for the cost to provide the public good g, c_1 and c_2 are the cost to implement policies λ_1 and λ_2, respectively. As in Alesina et al. (2019) we assume that the provision of public goods features increasing returns. With this, we can rewrite the utility as

$$u_i = g(1 - d_i \frac{a}{\lambda_1}) + \lambda_2 + y - (k + c_1 \cdot \lambda_1 + c_2 \cdot \lambda_2)$$

Consider some mass s (with $0 < s < 1$) of individuals that are all located (geographically or ideologically) at distance is from the capital. Assume that these individuals are able to separate from the capital and to form their own state. Hence individuals in s have an outside option with utility

$$u_i^o = g + y - r^o \tag{A.3}$$

The government in a separate state would face a budget constraint of $r^o = \frac{k}{s}$, with $s < 1$.

With this, an elite that is trying to prevent the separation of region s will chose policies λ_1 and/or λ_2 such that individuals in s prefer not to split:

$$g(1 - d_{is}\frac{a}{\lambda_1}) + \lambda_2 + y - (k + c_1 \cdot \lambda_1 + c_2 \cdot \lambda_2) \geq g + y - \frac{k}{s} \tag{A.4}$$

What is the optimal investment in homogenization policies (λ_1), or national identity policies (λ_2)? After some simple manipulation we find that the solution for λ_1 has the form

$$\lambda_1^* = \frac{z_2 - \sqrt{((z_2)^2 - 4z_1 z_3)}}{2z_1} \tag{A.5}$$

with $z_1 = c_1$, $z_2 = [(1 - c_2)\lambda_2 + k(\frac{1-s}{s})]$, and $z_3 = g \cdot d_{is} \cdot a$.

The optimal level of investment in national identity (λ_2) is given by

$$\lambda_2^* = \frac{1}{1 - c_2}[k(\frac{s-1}{s}) + g \cdot d_{is}\frac{a}{\lambda_1^*} + c_1 \cdot \lambda_1^*] \tag{A.6}$$

This expression will be positive as long as $c_2 < 1$ and $k(\frac{s-1}{s}) < g \cdot d_{is}\frac{a}{\lambda_1^*} + c_1 \cdot \lambda_1^*$, with $s < 1$. It is simple to show that the incentives to pursue either homogenization λ_1 or national identity policies λ_2 are increasing with the disutility from distance to the elite (a), with distances d_{is} and with size s of the separating region.

Moreover, the elite faces a trade-off between the two types of policies:

$$\frac{\partial \lambda_2}{\partial \lambda_1} = \frac{1}{(1 - c_2)}[c_1 - \frac{g \cdot d_{is} \cdot a}{\lambda_1^2}] < 0, \text{ with } c_1 < \frac{g \cdot d_{is} \cdot a}{\lambda_1^2} \tag{A.7}$$

In particular, we see from equation A.6 above that the incentive for the elite to invest into a common "national" identity (λ_2) increases with c_1, the cost to pursue homogenization policies λ_1, and also with the disutility from distance to the elite (a). Finally, the trade-off between investment in identity policies vs. homogenization policies (in equation A.7) is increasing with the disutility from distance to the elite a. Equivalently, it is increasing with distance d_{is} of the separating region from the center. Put differently, for a higher distance d_{is} or disutility from distance (a), a given increase in national identity investment is equivalent to a larger change in homogenization investment.

This has several interesting implications. For example, if the perceived distance from the elite would increase (e.g. due to some failure of the elite, like a lost battle, corruption or anti-monarchist sentiment against a ruling king), this will increase the incentive of the elite to pursue policies of national identity λ_2, *more* than the incentive to pursue the alternative, homogenization policies λ_1. In our context, both the defeat of the Prussian army at the battle of Jena-Auerstedt 1806 and the spread of republican and anti-monarchist sentiment in the wake of the French Revolution would have shifted (a) against the Prussian elite. Related, the gain of a new, more distant territory, such as the Rhineland and Westphalia in 1815 can trigger policies of national identity λ_2 to prevent such a region from separating.

Finally, we see that such national policies would not always succeed to affect individual utility and hence behavior. Rather their success should vary with distance from the elite (e.g. in terms of religious, or language differences). If we assume two types of individuals in region s with different cultural distance from the elite but the same outside option, exposure to national identity policies by the elite could change the behavior of one type sufficiently close to the elite, but not the behaviour of the other type, further away.

B Tables and Figures

Table B.1: Data Set

City	**Province**	**Prussia since**	**Births**	**Years**
Berlin	Brandenburg	1701	204.596	1804 - 1877
Stettin	Pomerania	1720	78.253	1815 - 1877
Danzig	West Prussia	1772	38.948	1824 - 1874
Aachen	Rhine Province	1815	95.553	1800 - 1874
Crefeld	Rhine Province	1815	92.788	1816 - 1875
Hagen	Westphalia	1815	86.716	1800 - 1874
Hannover	Hannover	1866	135.340	1800 - 1898
Frankfurt (Main)	Hessen-Nassau	1866	81.699	1805 - 1875
Nürnberg	Kingdom of Bavaria		138.237	1810 - 1900
Mannheim	Grand-Duchy of Baden		45.835	1800 - 1898
Heidelberg	Grand-Duchy of Baden		131.857	1800 - 1898
Lübeck	Hanseatic City of Lübeck		46.509	1818 - 1875

Table B.2: Popular national first names

Name	Observations	Top 30 Germanic Names based on Seibicke (1996, 1998, 2000, 2002, 2007)	Top 30 Germanic Names based on Khull (1909)
Men			
Carl	51941	✓	
Friedrich	35642	✓	✓
Heinrich	33409	✓	✓
Wilhelm	22550	✓	✓
Karl	13283	✓	✓
Ernst	10836	✓	✓
Hermann	10340	✓	✓
Ludwig	8356	✓	✓
Otto	6550	✓	✓
Albert	6177	✓	✓
Theodor	4667		✓
Adolph	4151	✓	✓
Richard	3869	✓	✓
Robert	3185	✓	✓
Hubert	2824	✓	✓
Women			
Caroline	13083	✓	
Wilhelmine	12572	✓	✓
Emma	10544	✓	✓
Bertha	7591	✓	✓
Friederike	6645	✓	✓
Ida	5444	✓	✓
Mathilde	4184	✓	✓
Carolina	3358	✓	
Wilhelmina	3276	✓	✓
Gertrud	3273	✓	✓
Minna	2575	✓	
Hedwig	2510	✓	✓
Karoline	1978	✓	
Karolina	1958	✓	
Friedrike	1621	✓	✓

Table B.3: Flexible Diff-in-Diff

	(1)	(2)
Dep. Var.	Dummy National First Name	
Treated × 1805	-0.0219	-0.0280
	(-0.62)	(-0.36)
Treated × 1806	-0.0304	-0.0290
	(-0.62)	(-0.13)
Treated × 1807	-0.0115	0.0678
	(-0.34)	(0.38)
Treated × 1808	-0.00703	0.0337
	(-0.18)	(0.37)
Treated × 1809	-0.0126	-0.0394
	(-0.29)	(-0.40)
Treated × 1810	0.00113	0.109
	(0.07)	(0.81)
Treated × 1811	0.00997	-0.0424
	(0.79)	(-0.43)
Treated × 1812	0.0173	0.0401
	(1.62)	(0.51)
Treated × 1813	0.0137	0.131*
	(1.03)	(2.10)
Treated × 1815	0.0773***	0.460***
	(9.16)	(16.46)
Treated × 1816	0.0605	0.205
	(1.29)	(1.18)
Treated × 1817	0.0712***	0.268*
	(2.64)	(2.12)
Treated × 1818	0.0806***	0.376***
	(2.90)	(7.75)
Treated × 1819	0.0855***	0.263
	(3.67)	(1.66)
Treated × 1820	0.0837***	0.211
	(7.15)	(1.68)
Treated × 1821	0.0849***	0.294**
	(17.32)	(2.89)
Treated × 1822	0.0779***	0.305
	(10.16)	(1.35)
Treated × 1823	0.112***	0.170
	(4.85)	(1.88)
Treated × 1824	0.0886***	0.0892
Year FE	✓	✓
City FE	✓	
Family FE		✓
Observations	97316	4224

Notes: Reference year: 1814. * p<0.1, ** p<0.05, *** p<0.01. Treated cities: Aachen and Hagen. Control cities: Frankfurt (Main), Hannover, Heidelberg, and Mannheim. Results based on equation 2. We only include parents without national first given name. Clustered standard errors at the city level.

Table B.4: Diff-in-Diff, Logit

	(1)	(2)	(3)	(4)
	Dep. Var. Dummy National First Name			
DiD Coefficient...				
...overall	1.666**		1.875***	
	(2.37)		(4.71)	
...Aachen		2.099***		2.369***
		(4.46)		(4.54)
...Hagen		1.457***		1.558***
		(2.68)		(2.67)
After 1815	1.379***	1.376***	1.207***	1.207***
	(4.00)	(4.05)	(2.74)	(2.74)
City FE	✓	✓		
Family FE			✓	✓
Observations	77328	77328	2422	2422

Notes: Coefficients display odds ratios; z statistics in parentheses. * p<0.1, ** p<0.05, *** p<0.01. Treated cities: Aachen and Hagen. Control cities: Frankfurt (Main), Hannover, Heidelberg, and Mannheim. Results based on equation 3. We only include parents without national first given name. Clustered standard errors at the city level.

Table B.5: Treatment Analysis 1815, Robustness

	(1)	(2)	(3)	(4)
	Panel 1: National conditional on family tradition			
DiD Coefficient...				
...overall	0.0787***		0.278***	
	(4.89)		(5.32)	
...Aachen		0.0595***		0.348***
		(20.35)		(15.99)
...Hagen		0.100***		0.223***
		(34.39)		(10.27)
After 1815	0.0329***	0.0329***	0.103***	0.103***
	(11.25)	(11.25)	(4.74)	(4.74)
City FE	✓	✓		
Family FE			✓	✓
Observations	77328	77328	2308	2308
	Panel 2: Excluding Wilhelm			
DiD Coefficient...				
...overall	0.0723***		0.255**	
	(5.46)		(3.02)	
...Aachen		0.0568***		0.373***
		(17.10)		(16.35)
...Hagen		0.0901***		0.160***
		(27.11)		(7.02)
After 1815	0.0248***	0.0248***	0.0484*	0.0484*
	(7.46)	(7.46)	(2.12)	(2.12)
City FE	✓	✓		
Family FE			✓	✓
Observations	76720	76720	2330	2330
	Panel 3: Berlin as control group			
DiD Coefficient...				
...overall	0.111***		0.322*	
	(6.65)		(4.24)	
...Aachen		0.0925***		0.416***
		(3.90e+10)		(1.55e+15)
...Hagen		0.131***		0.240***
		(7.77e+11)		(1.15e+15)
After 1815	0.0113***	0.0113***	0.0657	0.0657***
	(2.86e+10)	(7.44e+10)	(5.43e+10)	(2.80e+14)
City FE	✓	✓		
Family FE			✓	✓
Observations	64631	64631	2332	2332

Notes: * p<0.1, ** p<0.05, *** p<0.01. Panel 1 and 2: Treated cities: Aachen and Hagen. Control cities: Frankfurt (Main), Hannover, Heidelberg, and Mannheim. Panel 3: Treated cities: Aachen and Hagen. Control city: Berlin. In panel 1, we use national first names conditional on family tradition as dependent variable, i.e. we exclude families that choose a middle name from the parents as first name of the child. In panel 2, we exclude all families that choose "Wilhelm" as first name. Results based on equation 3. We only include parents without national first given name. Clustered standard errors at the city level.

Table B.6: Treatment Analysis 1815, Adjustment Aachen

	(1)	(2)	(3)	(4)
	Dep. Var. Dummy National First Name			
DiD Coefficient...				
...overall	0.0368		0.105	
	(0.72)		(1.11)	
...Aachen		-0.0256***		-0.0105
		(-9.49)		(-0.44)
...Hagen		0.107***		0.226***
		(39.84)		(9.48)
After 1815	0.0339***	0.0339***	0.0843**	0.0843**
	(12.56)	(12.56)	(3.54)	(3.54)
City FE	✓	✓		
Family FE			✓	✓
Observations	77328	77328	2454	2454

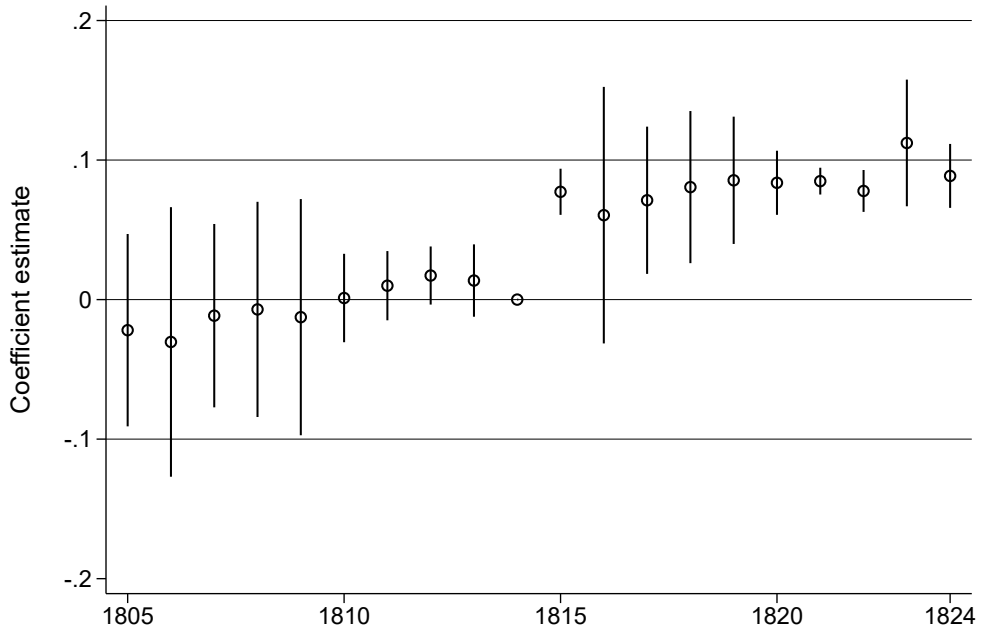
Notes: * p<0.1, ** p<0.05, *** p<0.01. Treated cities: Aachen and Hagen. Control cities: Frankfurt (Main), Hannover, Heidelberg, and Mannheim. Here, we assume that all French first names were "francicised" and thus, use the German translations. Results based on equation 3. We only include parents without national first given name. Clustered standard errors at the city level.

Table B.7: Treatment Analysis 1866

	(1)	(2)	(3)	(4)
	Dep. Var. Dummy National First Name			
DiD Coefficient...				
...overall	0.0000723		-0.0181	
	(0.01)		(-0.49)	
...Frankfurt		0.00622		0.0564***
		(1.12)		(5.07)
...Hannover		-0.0238***		-0.0694***
		(-4.27)		(-6.24)
After 1866	0.0193***	0.0220***	0.0705***	0.0639***
	(3.87)	(3.95)	(5.14)	(5.75)
City FE	✓	✓		
Family FE			✓	✓
Observations	66016	66016	6408	6408

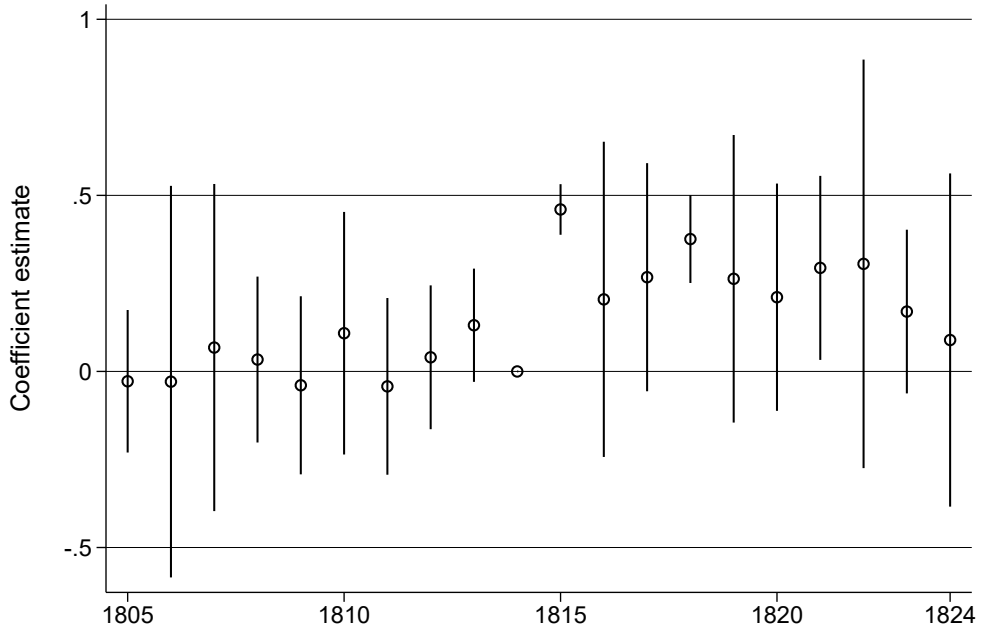
Notes: * p<0.1, ** p<0.05, *** p<0.01. Treated cities: Frankfurt (Main) and Hannover. Control cities: Heidelberg, Luebeck, Mannheim, and Nuernberg. We only include parents without national first given name. Clustered standard errors at the city level.

Figure B.1: Event Study, 1805-1824, City FE



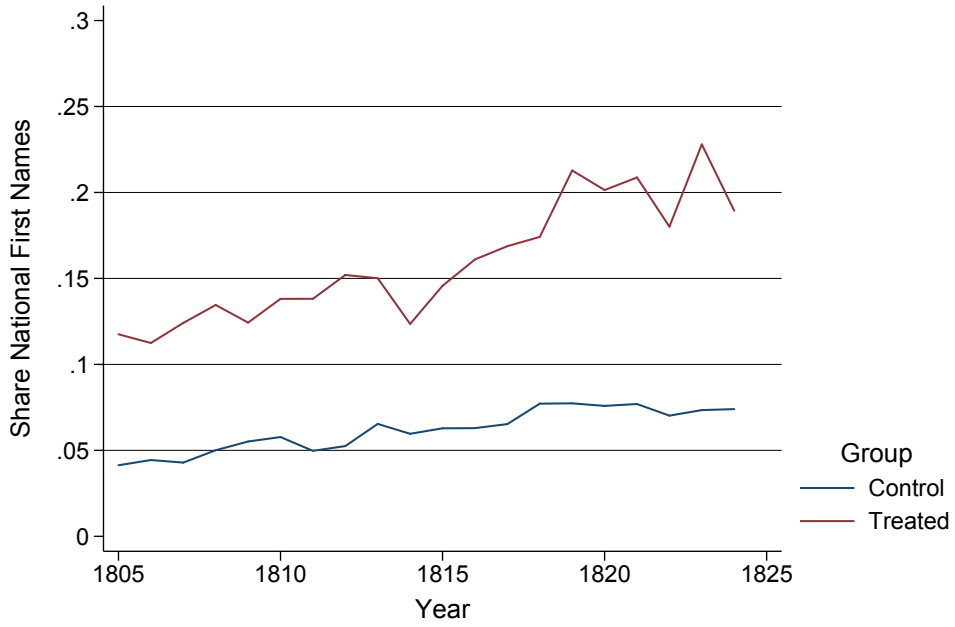
Note: The plot captures the coefficients from equation 2 based on column 1, Appendix Table B.3. Reference year: 1814.

Figure B.2: Event Study, 1805-1824, Family FE



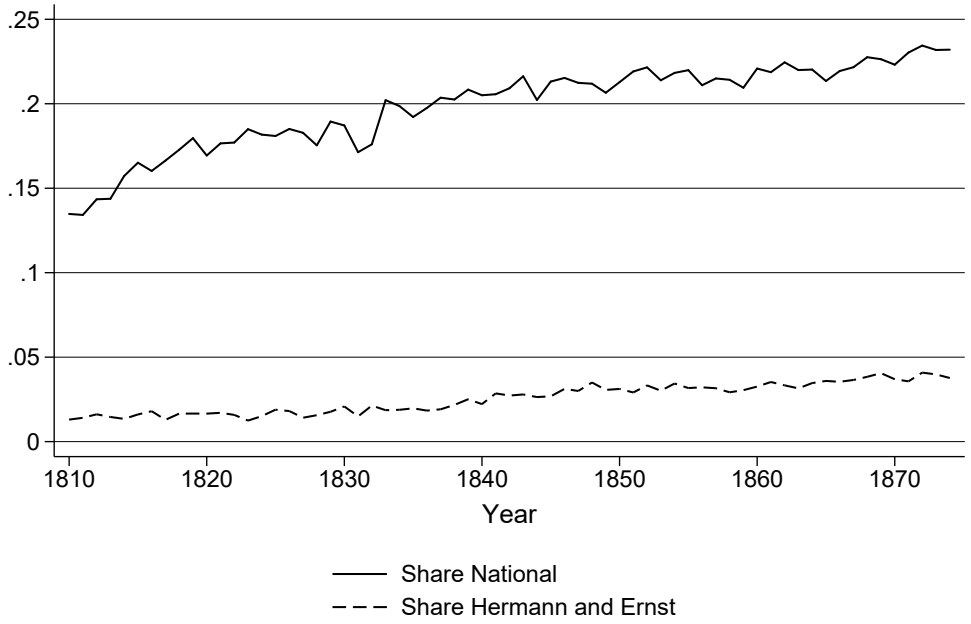
Note: The plot captures the coefficients from equation 2 based on column 2, Appendix Table B.3. Reference year: 1814.

Figure B.3: Development National First Names, 1805-1824



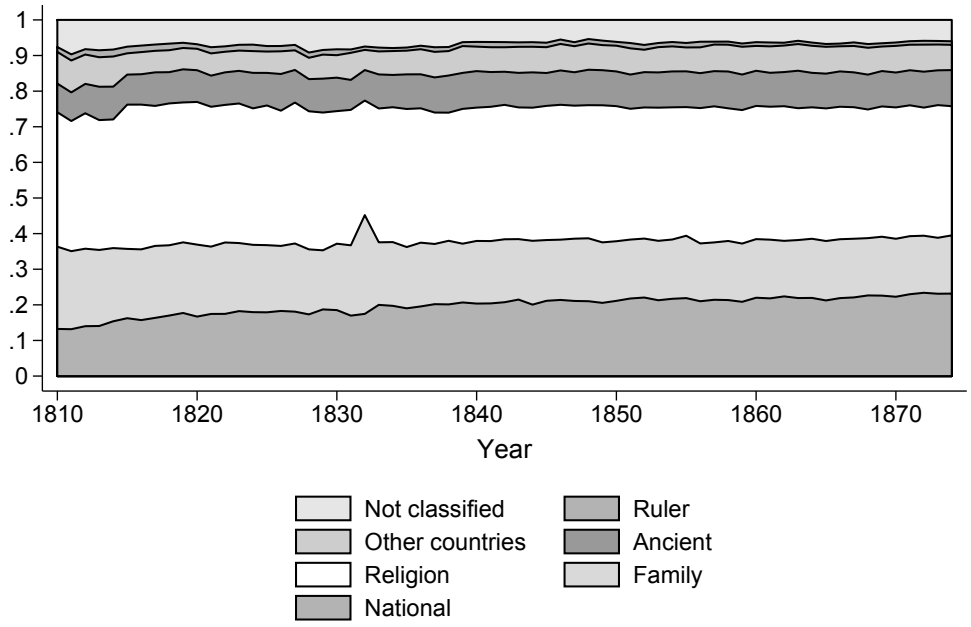
Note: Treated cities: Aachen and Hagen. Control cities: Frankfurt (Main), Hannover, Heidelberg, and Mannheim.

Figure B.4: Development National First Names, 1810-1874



Note: Balanced panel for 1810-1874. Includes: Frankfurt (Main), Hannover, Nuernberg, Hagen, Berlin, Aachen, Mannheim and Heidelberg.

Figure B.5: Development First Names, 1810-1874



Note: Balanced panel for 1810-1874. Includes: Frankfurt (Main), Hannover, Nuernberg, Hagen, Berlin, Aachen, Mannheim and Heidelberg.