University of Groningen

Historical Roots of Cities

Nguyen, Duc

DOI:
[10.33612/diss.201215971](https://doi.org/10.33612/diss.201215971)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2022

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Nguyen, D. (2022). *Historical Roots of Cities: Reexamining and Disentangling First and Second-Nature Geography*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen, SOM research school. <https://doi.org/10.33612/diss.201215971>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Historical Roots of Cities:
Reexamining and Disentangling First and Second-Nature
Geography

Duc Anh Nguyen

university of
 groningen

Historical roots of cities

Reexamining and Disentangling First and Second-Nature
 Geography

PhD thesis

to obtain the degree of PhD at the
 University of Groningen
 on the authority of the
 Rector Magnificus Prof. C. Wijmenga
 and in accordance with
 the decision by the College of Deans.

This thesis will be defended in public on

Thursday 17 February 2022 at 14.30 hours

by

Duc Anh Nguyen

born on 8 April 1991
 in Ratzeburg, Germany

Supervisors

Prof. S. Brakman

Prof. J.H. Garretsen

Co-supervisor

Dr. T. Kohl

Assessment Committee

Prof. H.J. de Jong

Prof. E.M. Bosker

Prof. D.R. Davis

Acknowledgement

I am eternally grateful to the people who had a tremendous influence on the development of not only this thesis, but my whole path within academia. Honestly, without the words of encouragement and support of some, I doubt that I would even be here and write this acknowledgement right now. Since this thesis is pretty much about path dependence, I will start chronologically: I credit Gonzague Van-noorenberghe for initially invoking my interest in trade through lectures, ultimately pulling me towards urban and regional economics. Very vague ideas that led to the first two chapters of this thesis were also a brainchild that sprung out of your office about 6 years ago. For those who feel addressed here, you can also largely blame Gonzague for my academic writing style. Daphne, thank you for causing my sleep deprivation during that project that led me to come up with the main idea of chapter 3, which in turn also heavily influenced the other chapters as well as my decision to apply to Groningen.

This leads me to Groningen, which was on my radar because some papers by Steven and Harry were often mentioned in the literature I was interested in. Apart from that, I honestly had little idea what to expect. As it turns out now, I ended up on the 'good path' and it was mostly due to the people I encountered in Groningen.

Thank you to my supervisors Steven, Harry and Tristan for teaching me the ropes of navigating through academia as well as the many fun but also productive meetings. Steven, thank you for the many constructive comments over the years, as well as not shying away from sometimes being the devil's advocate. I realize now that the latter is indeed a rare quality to have within academia as it takes much more effort to do it properly for it to ultimately lead to better work.

Harry, things would have been a lot more difficult with the repetitive data work I spent many of my days with without the enthusiasm you showed when I occasionally lost mine and this helped me to keep going. Thank you for the many invaluable suggestions over the years that pushed each chapter to be much better.

Thank you Tristan for often helping me to see things that were not immediately obvious to me as well as the many great and helpful comments throughout the years not only on my manuscripts. I appreciate the time and hard work to go the extra mile to meticulously scan through my manuscripts sentence by sentence, not missing even the smallest of details. Thank you also to Dimitrios who supervised me during the early months, and especially for the encouragement that further solidified me to pursue my ideas. I thank the members of the reading committee Maarten Bosker, Herman de Jong and Donald Davis for the careful review of my thesis and the very insightful comments that will keep me busy thinking about for a while. Thank you to the Gemmies for their support and the ability for always

managing to find a common slot in a bunch of busy schedules, as well as Kristian and Ellen and other SOM people behind the scenes for their support over the years. Thank

The close-knit and supportive community of PhD students also contributed to the typically high spirit throughout my time in Groningen. Although I am older, Romina behaved more like a nagging older sister and being far away from home, I think this familiar dynamic kind of made sense to the both of us. I was always looking forward to the highlight of the day which were the many coffee breaks to catch up (since I was absent half of the day) and if we were feeling very adventurous, to the 'fancy' coffee machine or the Chinese food shop. It was also the highlight of the week the rare times both me and Eda were present at the same time.

Maite, Femke, Daan, Ibra, Juliette, Stefan, Johannes, Timon, Nikos, Tobi, Gianmaria, Jos, the 9th floor and other amazing people: Thanks for the good times at lunch, dinners, snacks, bars and Vrijmibos, pub quizzes, Karaoke(!), movies, Game of Thrones season 8 watch parties (well, the meet-up was nice) and other gatherings where free food was involved. Thanks for the memories.

Finally, my parents and my brother for the unconditional support over the years through all the failures and hardships, for the food and for moving my stuff around Europe the past decade, and to Nicole, even though you are way too young to understand, things were always refreshingly simple when you were around. When you are finishing up your PhD, you will understand.

cheers (to you)

Duc

Bochum, January 2022

Contents

1	Introduction	1
1.1	Mean Reversion of the German City System	2
1.2	What's in a Name? Initial Geography and Urban Development . . .	5
1.3	From Bulwarks and Bombs to Boulevards	9
2	What's in a Name? Initial Geography and Urban Development	11
2.1	Introduction	12
2.2	Conjecture and Variety	16
2.3	Data	20
2.4	Zipf Distribution	24
2.5	Geographical Characteristics and Local Economic Activity	34
2.6	Conclusion	45
2.A	Data Appendix	46
3	Mean Reversion of the German City System After WWII: To What Mean?	61
3.1	Introduction	62
3.2	Non-linear Mean Reversion in the 20th Century	65
3.3	Synthetic City Construction	67
3.4	Data	69
3.5	Descriptive Statistics and Inverse U-shaped Growth	73
3.6	Empirical Strategy	74
3.7	Results	76
3.8	City Characteristics Conditional on Impact	87
3.9	Link between the Evidence and Theory	92
3.10	Conclusion	96
3.A	Data Appendix	97
4	From Bulwarks and Bombs to Boulevards	101
4.1	Introduction	102

4.2	Context	104
4.3	Empirical Strategy	115
4.4	Results	124
4.5	Robustness	127
4.6	Conclusion	135
4.A	Appendix: Data Digitization	136
4.B	Appendix: Amalgamation	136
4.C	Appendix: Individual Cities	143
5	Closing Remarks	145
	Bibliography	153
	Nederlandse samenvatting	165

List of Tables

2.1	Local Economic Activity and Geography: Similar Studies	19
2.2	Descriptive Statistics	25
2.3	Zipf and Toponyms	28
2.4	Zipf and Difference of City Rights pre-1700 and First Record Year . . .	33
2.5	Impact of Geographical Characteristics on 1910 Population: Baseline . .	38
2.6	Impact of Geographical Characteristics on 1910 Population: Robustness	39
2.7	Impact of Geography on Early Geographical Advantage: Baseline	41
2.8	Impact of Geography on Early Geographical Advantage: Robustness . .	43
2.9	Interaction between Characteristics and Early Cities	44
2.10	Description of Aggregated Characteristics	53
3.1	Descriptive Statistics	74
3.2	Individual City Examples	75
3.3	Unit Root Test on City Size (Donor Pool: <0.002 bombing deaths p.c.)	78
3.4	Unit Root Test on City Size (Donor Pool: <0.001 bombing deaths p.c.)	82
3.5	Actual and synthetic city equality of distributions	84
3.6	Unit Root Tests on Relative City Size (Donor Pool: <0.002 bombing deaths p.c.)	86
3.7	Unit Root Tests on Relative City Size (Donor Pool: <0.001 bombing deaths p.c.)	88
3.8	City Characteristics Conditional on Impact (Cutoff year 1980)	90
3.9	List of all sample cities	97
3.10	Unit Root Test on Log City Size (Donor Pool: <0.002 bombing deaths p.c.)	99
4.1	Descriptive Statistics	122
4.2	OLS estimates: Effect of a Change in the Street Infrastructure	127
4.3	First Stage: Effect of a Change in the Street Infrastructure	128
4.4	Second Stage: Effect of a Change in the Street Infrastructure	129
4.5	IV Estimates: Finer Patterns over Time	131

4.6	IV Estimates: Street Infrastructure and Hinterland Attractiveness . . .	132
4.7	IV Estimates: Effective Street Density	134
4.8	IV Estimates: Comparison of Different City Population Measures	142
4.9	Sample Cities	144

List of Figures

2.1	Toponym Dataset	20
2.2	Distribution of Year of First Record and City Rights Awarded	23
2.3	Distribution of Year of First Record and City Rights Awarded	26
2.4	Zipf Frequency of Geographical Characteristics (1. Specification)	29
2.5	Zipf Gap (City Rights pre-1700)	34
2.6	Zipf Gap (City Rights ≤ 1700 , First Record ≥ 1150)	35
2.7	Niemeyer (2012) excerpt	47
3.1	Individual City Gap of Population and Synthetic Population (Donor Pool: < 0.002 bombing deaths p.c.)	77
3.2	≥ 0.05 Weight Comparison City Gap	79
3.3	Individual City Gap of Population and Synthetic Population (Donor Pool: < 0.001 bombing deaths p.c.)	83
3.4	City Location by Relative City Size Impact (Cutoff 1980)	91
4.1	Ideal Street Density: Middle Age to Late Modern Period	111
4.2	Ideal Boulevard Density: Middle Age to Late Modern Period	112
4.3	Ideal Street Density: around WWII	113
4.4	Ideal Boulevard Density: around WWII	114
4.5	Boulevardization of Walls	124
4.6	Boulevardized Walls Rate and Approx. Earliest Defortification Year . .	125
4.7	Kassel Boulevards	137
4.8	Ulm Street Grid	138
4.9	Ulm Boulevards	139
4.10	Bundesfestung Ulm	140
4.11	Orientation Histogram	141

Chapter 1

Introduction

The rapid development in many emerging economies in the recent years also came along with a rapid movement of population from the rural to the urban areas, which emulates the process that happened about a century before within the now developed economies. This suggests an intimate connection between economic and urban growth. However, whereas the natural question in the study of economic growth is how this development came about, the corresponding urban growth questions and of this thesis would both be how and where this development exactly happened.

Initially, this chapter was written to introduce the three chapters of this thesis that deals with spatial economic topics on urban growth, geographical determinants of such growth in the long-run, and the urban form, with a common theme that binds them all together. However, instead of a common theme, there is a way of thinking about the literature that led me to the main ideas of each chapter; and this way of thinking was developed after critically reading the seminal paper by Davis & Weinstein (2002).¹ In a sense, after reading that paper, the mostly methodological criticism motivated the three chapters of this thesis, which will become apparent in the following introduction to each chapter, in section 1.1 to 1.3.

The motivation of this thesis is primarily to contribute to the regional and urban economics literature. However, one will soon realize that for certain topics in this literature, one will have to look back in time and deal with history as well. In this spirit, the following very brief overview of each chapter will be given in a chronological order: In chapter 2, we are interested in the geographical founda-

¹I was assigned to write a mock review as a graduate student. I suppose that reading that paper with a perhaps over-critical eye helped in imagining how I would answer, or re-answer some main questions they asked. Therefore, the following may, for some, read more as a very elaborate referee report on that paper, however providing my own take on their approach as well as that of the related literature on questions surrounding urban growth and the location of economic activity.

tions of the past and how they shaped the urban landscape of the past and (still) shape the urban landscape of today. For that, we have to look back at the initial days of cities that can be found during the Middle Ages for most European cities. However, the typical challenge of looking that far back into the past is the lack of good information or data. In chapter 2, we consider initial settlement names as a source for such information, because they contain information about the initial geographical circumstances of settlements and preserved them throughout the many centuries.

Turning to the mid-20th century in chapter 3, we look at whether WWII had any lasting impact on German cities and its urban landscape. One central theme of this chapter is that of structural changes that occurred during the 20th century, such as deindustrialization, the decline of coal mining and the motorization/suburbanization within cities. These developments ought to be separated from the developments caused by the WWII destruction as otherwise, we may for instance misattribute developments that actually stem from structural changes as an after effect of WWII.

Finally, we arrive in the 1960s and 70s in chapter 4, where we take a look at German post-WWII urban transportation infrastructure projects. Urban planners used the post-WWII destruction as an opportunity to accommodate the urban form towards more car-friendliness. We estimate the impact of the post-WWII understanding of a modernized inner city on suburbanization.

The following overview and state of the literature will however not be given in this chronological order, but in the order of how this thesis came about. This is because each subsequent chapter is primarily built upon follow-up questions, ideas or thought experiments from the previous chapter. In this sense, the following will also be in a chronological order with the aim of showing the links between each chapter.

1.1 Mean Reversion of the German City System

Within the literature, Davis & Weinstein (2002) stands out by introducing several approaches to test some very different theoretical approaches on the location of economic activity. Notably, there was a lack of empirical evidence on the ideas following the seminal paper by Krugman (1991) that introduced to economists the idea of historical accidents and lock-in effects through cumulative causation effects. For example, the development of Silicon Valley is usually attributed to the 'historical accident' of Hewlett and Packard founding the company of the same name in a garage in Palo Alto. This is followed by many other similar companies being drawn or founded in the same location due to other similar companies being

around, which is the cumulative causation process in this case, as they for instance want to take advantage of the highly specialized workforce or suppliers around that cater to this industry. Other industries, such as the movie industry in Hollywood or the financial industry around Wall Street seem to have sprung out in a similar manner in the past, driven by similar advantages that the proximity to similar firms or specialized workforce offers. One idea that is implied here are strong lock-in effects of these industries in the particular location, reinforced by these cumulative causation effects. Another implication is that if there is an event that even temporarily makes many firms leave Silicon Valley, another high technology cluster could spring out somewhere else in this world.

We can think of cities as clusters that exploit these advantages of proximity as well, for instance by citizens sharing amenities that are typically found in larger agglomerations i.e. universities, cinemas, malls, or markets. Duranton & Puga (2004) for instance introduce the taxonomy of sharing (e.g. through large indivisible facilities among citizens and firms); that of matching in which agglomerations (e.g. through raising either the expected quality/chances of matching); and finally that of learning (e.g. through an exchange or accumulation of skills, knowledge and information). Although the mechanisms presented here are very different and for the sake of tractability are typically introduced individually within urban economic models, these mechanisms all follow the general idea that agglomeration causes more agglomeration, whether it is through sharing, learning and/or matching. Hence fittingly, these mechanisms are together summarized as increasing returns theories.

This is where one of the main contributions of Davis & Weinstein (2002) lies by introducing an empirical test on the question: what if we shock a system of cities enough such that it could 'initiate the reshuffling' of the economic activity in a region, would the spatial distribution of economic activity return to some previous long-run equilibrium or would some place benefit while others lose? Importantly, if we would detect a permanent effect of such a shock that would lead to a shift of the economy towards another long-run equilibrium, thus implying the existence of multiple equilibria, this finding would be most in line with the aforementioned increasing returns theories. On the other hand, they argue that no permanent effect through such a shock would be most in line with a theory that emphasizes natural advantages of a location such as rivers or sea-access, as these did not disappear so that we would return back to the previous long-run equilibrium. A third theory posits that urban growth is random, drawn from a distribution with a common mean and variance and allowing for a permanent effect through a change of the relative city size. This implies that the shape of the distribution should remain stable; contrary to increasing returns theory that does not impose such a

restriction. Important here to the literature is that the outcome of this empirical test gives us clues on the relevance of these three different theories, which would in turn inform the theoretical side of the urban growth literature going forward.

For the large shock, Davis & Weinstein (2002) consider the WWII bombings on the Japanese urban system and whether this would induce a permanent change in the Japanese relative city size, to the detriment of cities that were bombed more heavily. However, they concluded reversion to the mean as the effect of this large shock dissipated in 1965. They explain this result as being highly consistent with a theory of locations possessing fundamentals, such as river or sea access and other natural advantages, which would make it difficult to move away from a location even after large shocks on the city itself. Similarly until 1963, Brakman et al. (2004) found that the large WWII shock had a temporary effect. However, Bosker, Brakman, Garretsen, & Schramm (2008) increased the years of analysis until 1999 found a permanent effect for most cities.

This seems puzzling, as a permanent impact of WWII was somehow detected with a much longer time span of investigation, even though there was no permanent impact detected initially until the mid 1960s. We would instead expect the opposite, in that we may still find an impact until the mid 1960s, which did not last until 1999. It was also unsatisfying that the literature itself did not offer an explanation for this puzzle.

One way to explain this puzzle is through the suburbanization and the decline of coal regions in the 1960-70s that would not matter if we were to only look until the 1960s: For the German highly industrialized and coal rich Ruhr area in particular, coal production becoming not as profitable anymore may have also driven the decline of Ruhr area cities that has nothing to do with WWII itself. Similarly, better transportation drives away inhabitants at the center to the suburbs, which may also contribute to the decline and which development may be picked up if we look at the development of German cities from the 1970s onwards. One then needs to find a way to disentangle or decompose the effect of suburbanization and that of other simultaneous developments, with the WWII bombing shock itself. One way to do this is to consider counterfactual German cities that were not subject to the extensive WWII bombings, which we then use as a benchmark for the recovery of German cities after WWII. We do this in chapter 3 of this thesis, while aiming to leave everything else almost unchanged as compared to the study of Bosker, Brakman, Garretsen, & Schramm (2008) such that we can mostly attribute any difference to the consideration of a counterfactual.

Interestingly, we find no impact of WWII bombings on most German cities, neither until the 1960s nor the 1990s, which starkly differs from Bosker, Brakman, Garretsen, & Schramm (2008) conclusion of a permanent impact on all cities due to

WWII. Importantly, this result itself can reconcile the different findings of the literature as mentioned above, where the results until the 1990s may indeed be driven by developments that have nothing to do with WWII, such as the aforementioned decline of coal regions and widespread suburbanization. Hence, the puzzle in the literature does not seem puzzling anymore. Importantly, our result of no impact of WWII here also points at the aforementioned fundamentals of a location being the most relevant theory about where the economic activity is spatially distributed i.e. around natural advantages. As mentioned before, it is most consistent with the result of a large shock on the city system that does not show any permanent effect, provided that these fundamentals remain unaltered.

1.2 What's in a Name? Initial Geography and Urban Development

The question that may immediately arise is what these locational fundamentals exactly are. Davis & Weinstein (2002) gave the example of rivers, coasts, plains instead of mountains or deserts, however without measuring the importance of each of those features. Surveying one part of the literature, similar examples are often given (Redding, 2010; Redding et al., 2011) about fundamentals without actually being specific about their relative importance to another. Instead, they are mentioned together as fundamentals as important determinants that explain the persistence of economic activity around a location after other explanations have been exhausted, whereas the importance of individual geographical characteristics were at most discussed descriptively at this point.

Recently, a more systematic approach measures the contribution of a few individual geographical characteristics, e.g. Henderson et al. (2017); Bosker et al. (2013); Bosker & Buringh (2017). This may reflect a parallel branch of the literature that is more interested in the relative importance of individual locational characteristics to another, instead of these locational fundamentals altogether in relation to other drivers of local economic activity, such as the aforementioned increasing returns. This branch also looks through the lense of history and historical urbanization patterns, and is hence also close to the traditional economic history and development literature. This stands in contrast to the other urban growth literature branch with strong theoretical foundations that takes advantage of structural models. However, these models typically consider fundamentals as a feature of the model while not specifying what these fundamentals exactly are. This also means that with the exception of the more recent papers that measures the contribution of very few individual geographical characteristics, relatively little is known about what these fundamentals entail.

Furthermore, within the literature that primarily focuses on natural advantages and its impact on local economic activity, there is the idea of path dependence; in this context as natural advantages that determined the spatial distribution of economic activity very early on, which persist until today even if these natural advantages became obsolete. This idea was pursued by Bleakley & Lin (2012) with portage sites as a convenient place to carry boats, making them good places for trade. The advantage of these sites disappeared through transportation innovations, but major cities still persist around these places, which points at the relevance of path dependence even today. Hence the path dependence story is very similar to the story of historic accidents and subsequent increasing returns that explain Silicon Valley, Wall Street or Hollywood.

There is an equivalent and comprehensive explanation for this within the literature through the conceptionalization of first-nature and second-nature geography: First-nature geography reflects favorable geography that came first in determining the location of economic activity, such as important early natural resources (copper, tin), river or sea-access, or the aforementioned portage sites. It is then when second-nature geography kicks in, which describes the interaction of people with each other, drawing other people to the location, leading to more people to locate there and so on. With these concepts in Davis & Weinstein (2002), one assumption here is that first-nature geography came first in determining the location, with second-nature geography becoming more important over the many centuries of human settlement. If we look at the large agglomerations of today around New York, Tokyo or London, it would seem difficult today to attribute any of this development to first-nature geography. Still, we think that first-nature geography was an important determinant at least very early in time during the initial settlement period. For New York, we can think of its historically important port, which was the first piece of land of the new continent that many migrants first saw. For London, it was known as the Roman colony of Londinium and with its bridge an important crossing over the river Thames. For Tokyo however, it was initially a fishing village until it became the center of political rule in Japan during the Edo period at the beginning of the 17th century.

Yet, apart from this political determinant of being a capital city, which does not apply to most cities, we may typically think of something exceptional that a place has which makes it easier to imagine a site 'snow-balling' out of control. We may not think of a heath land without access to water, which barely supports life with low fertile land as a place with a potential for 'snow-balling' in the past. We may then conveniently think of 'initial conditions' determined by first-nature geography which made the second-nature snowball roll or not.

For London, this initial condition may have been the Roman bridge over the

Thames, while it may have been the port for New York. This also points at some man-made advantages to be similar to natural advantages, as they can be considered as first-nature geography as long as they were present during the initial years of the settlement.

Jedwab et al. (2020) sees a lack of man-made advantages being considered as compared to natural advantages, at least within the empirical literature on the impact of first-nature geography (Henderson et al., 2017; Bosker et al., 2013; Bosker & Buringh, 2017). This may be due to the likely reverse causality inherent to man-made advantages, which makes it hard to distinguish from second-nature geography. For example, the bridge over a river may attract settlers, but it may also have been built because of many settlers already having been present around the location. Hence, we may summarize two challenges in measuring the importance of different first-nature, natural or man-made, advantages: First, it is difficult to disentangle the individual effects of first-nature from second-nature effects as determinants of the economic activity of a location and second, it is difficult to identify man-made advantages that are first-nature i.e. those that are considered as initial conditions as well.

It is then no surprise that the literature on first-nature geography focused on natural advantages given these challenges, as they were clearly present at the beginning of the settlement i.e. are considered initial condition, and are on top of that easily measurable if these persisted throughout time. That is, one can simply consider the rivers, mountains and seas of today as being the same as when these natural characteristics were first settled at; however with the disadvantage of severely reducing the scope of first-nature characteristics that are being considered to only natural ones that are easily observable today.

However, an ideal study on the initial conditions that determines the most important earliest path within the idea of path dependence would seem to need a larger scope or variety of geographical characteristics, that also considers man-made characteristics besides natural ones, and which were clearly already present very early in a settlement's history; thus emphasizing the first in first-nature.

In chapter 2 of this thesis, we will consider the name of a place as the name conveys exactly this kind of information: The initial name of a city typically describes the immediate surroundings of a place, which is typically kept throughout history. This gives us a glimpse of the geographic features, man-made or natural, at the time of the very beginning of a settlement. With developed cities of today, we may not exactly know around which initial geography the early settlement grew around, so it would be hard to isolate its initial geographical condition. Another advantage of considering place names here is that place names survive, while some characteristics may not, such as forests or castles, while it will still survive in its

name. Hence, place names as a source of information may give us a closer look at the initial conditions, and as such the initial path taken that is important within the concept of path dependence; at least as compared to the alternative of considering geographic informations recorded in modern times. Another advantage is as a long-term effect may indeed only materialize in a very long-term manner, the geographic information drawn from place names are likely the furthest away in a settlement's history that we can obtain.

Continuing with Davis & Weinstein (2002) and locational fundamentals, they posit that geographical advantages are in line with what is known as Zipf's law for cities, in that the city size distribution follows a remarkable pattern: In a given region, the second largest city is about half as big as the largest in terms of population, while the third largest city is about a third as big as the largest, the fourth largest about a fourth as big as the largest and so on. Following Krugman (1996) one can then see that if there is a Zipf's law for geographical advantages as well, then it would give a powerful explanation for the emergence of Zipf's law for cities.

The issue is that so far, no paper attempted to test the simple idea of a Zipf's law for geographical advantages, although there had been approaches to measure such advantages in the past. One core issue surrounding such a test is that, as we want to essentially measure first-nature geography effects, one needs to abstract from the assumed second-nature geography. In modern times, with the idea of second-nature assumed to dominate first-nature geography effects, this seems to be another challenge without a straightforward approach. That is, one approach is to boil down the effects to only first-nature geography, which could be done by looking at the development during the early history of a settlement when second-nature geography is assumed to not play much of a role yet.

One benefit of the place name data is that it also provides us with a measure on the year of the first foundation of a settlement through the year the name of a place was first mentioned. Together with the city rights year awarded, we can construct a very simple measure of geographical advantage measured very early on in the history of a settlement. The time between the foundation date and the city rights date is when a settlement is in its primitive, early state where markets and walls, courts and basic rights not granted to peasants that are subject to a feudal lord are not present. Only when a place is of a certain size and can afford to buy itself out from their feudal lord via city rights can we think of significant second-nature geography effects to take place, mostly through the introduction of a basic market. Hence, the time it takes from foundation date until the city rights date is then mostly determined by first-nature geography effects, which gives us a simple proxy measure for first-nature advantages. We indeed find evidence for a

Zipf's law for geographical advantages by using such a measure, which is the first of such result presented in the literature.

1.3 From Bulwarks and Bombs to Boulevards

The focus of the previous section was on first-nature geography, expanding the definition beyond natural features and emphasizing the first in first-nature. In particular, we have considered man-made features in the same category of first-nature as natural ones, as they both may as well be considered as an initial condition of a location. A typical assumption made within the literature as well as by Davis & Weinstein (2002) is that natural features are fixed and not subject to changes, so that if the spatial equilibrium of economic activity is uniquely determined by natural features, it will be unchanged as well throughout history. However, with man-made features also considered within the realm of first-nature, not allowing the possibility of first-nature to change would seem peculiar, especially when considering the long time horizon and the possibility of humans creating and changing man-made features.

This seems especially relevant for cases of initial first-nature advantages becoming increasingly detrimental. For instance, city centers being very close to a river may have been great for trade when the cost of shipping was low relative to horse-powered trade. When horses were replaced by steam and combustion engines, and with growing urbanisation, the proximity to a large river may have become detrimental, especially on the other side of the river due to the need to cross this river to reach the center. Hence, first-nature geography may not be fixed in its presence, but also not in effect over time; and we may want to somehow reduce the detrimental effect through e.g. more and wider bridges.

Similarly, transportation within a city were initially done by hand-carriages which did not require particularly wide streets. With the exception of a few cities that were planned by a monarch and its many pompously planned wider boulevards, the typical European city sprung out of a humble village first which did not have the foresight that it may become a larger urban area. Thus, it may need wider streets to accommodate its increasing traffic needs, either through more pedestrians or later through horse-carriages and cars. An extensive network of narrow streets that were created around the Middle Age may not be advantageous anymore in light of increasing urbanization and urban transport innovations.

To put it another way, the demand for such infrastructure is likely to be endogenous with urban economic activity or transportation technology. However, similar to a city being stuck around river locations that are rather disadvantageous today, the city may be stuck with too narrow streets that would require very (po-

litically) costly renovations. As these changes tend to not happen very frequently at a significant scale throughout history, the supply of such infrastructure may lag the endogeneously determined demand.

Chapter 4 considers the case of WWII again, but this time we consider the long-run impact on infrastructure as a man-made first-nature that is difficult to change. Here, the difficulty to change is due to considerable cost of razing older, perhaps disadvantageous infrastructure laid down during the Middle Ages within a city, and which were razed during the bombings that made it easier to replace them with infrastructure more in line with today's infrastructure demands.

In particular, we take a look at the supply of streets and wider boulevards and whether a change of these after WWII led to a decentralization of population outside the immediate city. We find suburbanization effects through more boulevards in the city center around the 1970s, but no such effects through changes in the density of streets.

Putting this chapter in the context of other chapters, first, the street network is not natural, but man-made and we have considered this case in chapter 2. However this case here is in contrast to chapter 2 where we rule out reverse causality of man-made features that existed very early in a settlement's life as suggested by place names. Chapter 2 thus measures the traditional idea of what we think first-nature geography to be as initial condition of a settlement. Within a historical context, chapter 2 also considers the case the farthest away in history, which allows us to abstract from anything else other than this traditional view of first-nature geography.

To keep a chronological order, chapter 2 is first presented in this thesis, as we consider the first-nature historical roots of a city that is given by its initial conditions as suggested by the place name. Chapter 3 primarily deals with a disruption of second-nature effects abstracting from anything else i.e. a disruption of human interaction with another at some location. As stated before, with the idea that first-nature geography were the primary determinant of urban growth of early settlements, larger cities thrive on second-nature geography going forward throughout history. Conveniently, we may think of the historical roots determined by second-nature causes to have occurred primarily in recent history. Lastly, chapter 4 with its street infrastructure considers the historical roots that were added throughout the Middle Ages and were only changed many years after WWII.

This introduction also showed that each chapter in this thesis is connected in a common framework, where the main object of analysis can be conveniently categorized as either first or second-nature. However, the central idea from which each individual chapter evolved can be summarized in the question: What if first-nature geography/location fundamentals are allowed to change?

Chapter 2

What's in a Name? Initial Geography and Urban Development*

Abstract: Place names, so-called toponyms, offer a unique historical glimpse of the immediate surroundings of the very first settlers of a settlement. Toponyms convey the original geographical characteristics that include characteristics that have not been considered in the literature, were forgotten throughout history, or are simply unobservable today. We construct a unique dataset of 3738 mostly German toponyms, their first ever recorded mention as well as their city rights year. This allows us to see which geographical characteristics were important determinants of a settlement's growth during its infancy. Our results indicate that land-based trade characteristics and security were important determinants very early in development, with characteristics associated with water-based trade being more important in more recent periods. Our unique dataset also allows us to test for geographical Zipf's law. We show that the frequency of geographical toponyms, as well as a proxy for local geographical advantage both follow a Zipf distribution.

Keywords: First-nature geography, city size distribution, Zipf's law, path dependence

*This chapter is co-authored with Steven Brakman, Harry Garretsen and Tristan Kohl.

2.1 Introduction

Most cities in Europe can find its origins during the Middle Ages, where the role of geography is often emphasized to have played a crucial role in this initial stage of urban development (e.g. Bosker & Buringh (2017)). Yet, relatively little is known about which geographical circumstances, or 'first-nature' characteristics as they are also called, were crucial during the earliest development stage.

Furthermore, early geographical circumstances may still shape the modern urban landscape today, even after many centuries or after becoming obsolete (e.g. Bleakley & Lin (2012)). This is an example of path dependence, not through temporary historical events that have a permanent impact on the spatial distribution of economic activity, but through initial conditions as determined by early geographical circumstances (see e.g. Allen & Donaldson (2020)). Hence, knowing about the geographical circumstances as close as possible to the beginning tells us about the origins of the cities of today. However, a rule of thumb is that historical data is sparser the further we go back in time (Nagy, 2021); data which is often non-existent for the earliest settlements.

In this paper, we aim to empirically show *which* types of geographical characteristics were more advantageous for urban development during the early development stage and which types (still) matter today. Addressing the issue of finding suitable historical data, we have created a dataset that exploits the information contained in mostly German place names, so-called 'toponyms', providing us a unique glimpse at the initial conditions of early settlements. With this dataset, we contribute to both sides of equations that estimates the impact of geographical characteristics on historical and modern urban development:

We contribute to the 'left-hand side' of the equation, where the variables may be early history local population, or some proxy for early economic activity through the number of archeological sites (Davis & Weinstein, 2002; Bakker et al., 2019), ancient cuneiform records (Barjamovic et al., 2019), or ancient Roman tableware (Flückiger et al., 2019). However, with the exception of a few Roman founded cities, almost all German settlements that still exist today were first seen after 800AD, as seen in Figure 2.3a. Yet, medieval data for settlements in its early development stage are lacking. For example, Bairoch et al. (1988) dataset is widely used in the medieval urban economics literature (Hanlon & Heblich, 2020). Their city population estimates mostly consider the larger European cities, which means that population data of city's at the early stage of development are often missing.¹

¹The data is also often based on educated guesses, which are prone to large errors (Bairoch et al. (1988) pp. 298). Typically, only after the introduction of regular censuses at the beginning of the 19th century is the population in cities systematically and centrally recorded and thus reliable. For example, census city level population data became available for France, the U.K.,

To get as close as possible in time to these early settlements, we have collected data on which year a place's name was first mentioned in some record, as well as on the date when the place received its city rights. Together, we propose a novel proxy specifically to measure the inherent early geographical advantage of a location around the time of its establishment, which we define as the 'gap' in the years between when the city rights were received and when the toponym was first mentioned. Here, we exploit the fact that only larger places in feudal Europe were able to receive city rights as a developmental milestone, from which on they were allowed to hold markets, create courts, or build walls. These are examples of early indivisible facilities for the just promoted city and we regard them as the initial source for increasing returns through sharing (Duranton & Puga, 2004). We can then isolate geographical advantage effects before these city rights were received as we abstract from increasing returns effects. Hence, this data allows us to estimate which characteristics provided settlements with an early geographical advantage.

We contribute to the 'right-hand side' of the equation, where the typical measures are geographical characteristics as measured *today*. Using present data as a proxy for the geographical circumstances of the past may be inappropriate due to depletion and destruction that happened throughout the many centuries. For example, if forests, fords or castles vanish over time, and if they were important drivers of early human settlements, we may not consider them simply because they are not physically present anymore. The possibility of characteristics not being fixed throughout history adds to the challenge of historical circumstances that is not addressed by considering today's geographical circumstances.² Furthermore, the location of the earliest settlement within the larger city may be unknown today, so that we may misassign present characteristics as having been relevant very early on. That is going back to the beginning, we would want to measure both the *local* and the *historical* geographical circumstances of the early settlement around the original houses.

These local and historical characteristics survived in the toponym of a place, providing us a glimpse at the geographical circumstances from the very origins of a place's establishment. Toponyms convey geographical features that span not only natural features that we can observe today, but also those that have disappeared throughout history, as well as man-made ones such as castles, churches and

and the U.S. around 1800 and later, while they became available only later for Germany.

²See e.g. (Combes et al., 2010) for an exception, as they include soil quality in their analysis.

monasteries that were around from the very beginning.^{3,4}

As compared to existing research in the economics literature which is naturally inclined to analyze the impact of geography from a trade-based perspective by emphasizing the role of rivers, sea or road access of locations (see e.g. Bosker, Brakman, Garretsen, De Jong, & Schramm (2008), Henderson et al. (2017), Barjamovic et al. (2019), Bakker et al. (2019)), and look at geography through a narrow lense, toponyms allow for a broader menu of different geographical features that can be considered. In our dataset of 3738 mostly German toponyms, we categorize 242 different types of characteristics, many of which are rare.⁵

Toponyms enlarge the potential set of first-nature characteristics with a focus on the very first characteristics of settlements that may have disappeared, while the individual inclusion of characteristics is not driven by the preferences of the researcher. We thus consider various characteristics that were not considered before in the literature surrounding the relative importance of various geographical characteristics (Bosker et al. (2013), Bosker & Buringh (2017), Henderson et al. (2017)).

Our estimates of a large 'gap' indicate that water-based trade characteristics, such as rivers or fords, were not advantageous characteristics very early in a settlement's development. This is contrary to other papers on the role of the access to a river or the sea for very early development (Bosker, Brakman, Garretsen, De Jong, & Schramm, 2008; Bakker et al., 2019). However, those places are on average more populated in 1910, which indicates that these characteristics only became more advantageous later on. On the other hand, bridges and the security of castles were advantageous very early on, with castles having a lasting positive impact on population even in 1910 when they have already become long obsolete. This is consistent with the notion of the decline of land-based trade towards water-based trade e.g. Bosker et al. (2013).

³The literature typically distinguishes between first-nature and second-nature geography, with the former typically being represented as natural features existing before people i.e. exogenously given, while second-nature arises through interaction of people with another (Redding, 2010). In that sense, man-made characteristics are typically categorized as second-nature characteristics as a consequence of human interaction. However, if these man-made characteristics existed before a significant number of people were present for these characteristics to arise through interaction, we may then consider man-made characteristics as first-nature as well.

⁴Jedwab et al. (2020) sees a lack of man-made characteristics that are considered in the literature. This may be partly due to the inherent reverse causality with population that makes it difficult to distinguish it from second-nature effects. Arguably, man-made characteristics that are considered as first-nature characteristics have the advantage that they are less susceptible to reverse causality.

⁵However, the number of different types depend on the level of aggregation. The 242 different characteristics here can be roughly thought of as belonging to the same toponym with different spelling, but the same *meaning*. This corresponds typically to the 'suffix' of a toponym, while ignoring the 'prefix', as we will explain in data section 2.3. Also considering the 'prefix' of a toponym would further increase the different characteristics that may describe the local geography with too much detail for analysis.

Lastly, a complementary question to *which* first-nature characteristics mattered for urban development is whether first-nature advantages are even relevant enough to explain some empirical regularities found by the literature. One important regularity is the size distribution of cities, that has been empirically shown to follow a Zipf distribution (Giesen & Südekum (2011), Ioannides & Overman (2003), Soo (2005), Nitsch (2005), Dittmar (2011)). Hence, Zipf’s law of cities is an attractive statistical property to emerge out of urban growth models in order to show the relevance of the mechanisms in conforming with this law.⁶ Krugman (1996) and Gabaix & Ioannides (2004) proposed a simple mechanism where Zipf’s law of cities is a consequence of a Zipf’s law of local geographical advantages, as the city size is determined by the city’s local geographical quality. Finding the same empirical regularity in the geographical quality as well would indicate a more intimate relationship between geography and the city size; a relationship that has so far not been explored empirically due to the lack of a suitable proxy for geographical advantages as it is difficult to distinguish increasing returns from first-nature effects (Redding, 2010).

Our early geographical advantage proxy, defined as the gap in the years between city rights and first record offers a simple way to measure this advantage without imposing additional structural assumptions. Instead, the crucial assumption here is that increasing returns effects were not relevant before settlements became *de jure* cities, allowing us to isolate the geographical advantage effect. We show that this measure for early geographical advantage follows a Zipf distribution as well.

Additionally, we look at the early settlement patterns to first-nature characteristics, capturing the settlers choice where to settle first. Concretely, our dataset contains a large variety of first-nature types that enables us to look at the frequency of occurrence of different first-nature types. We show that this frequency of occurrence of different types also follows a Zipf distribution.⁷ Together, we take this as support for theories that rely on geography to generate a Zipf’s distribution.

This paper is related to development economics literature surrounding geography. Some studies looked at a few characteristics, such as terrain ruggedness and its impact on long-term development by inhibiting slave trade (Nunn & Puga, 2012), the impact of abnormal floods (Chaney, 2013), the Tsetse fly on popula-

⁶This distribution has generally been generated through some form of Gibrat’s law of proportionate growth (Gabaix, 1999a,b) within some urban growth models (Gabaix, 1999b; Eeckhout, 2004; Rossi-Hansberg & Wright, 2007; Duranton, 2006). However, as Gibrat’s law by itself has no natural economic explanation, an economic explanation to Gibrat’s law is typically imposed. This in turn may add too much freedom in the choice of the economic mechanism that governs Gibrat’s law.

⁷Villette & Purves (2018) provide evidence for a Zipf distribution for ‘field names’, not names of places settled by people. They also do not individually interpret and categorize the names for each place manually into the given types, but categorize them based on common and similar names.

tion density (Alsan, 2015), and the economic impact of droughts (Hornbeck, 2012). Toponyms provides us a very long perspective on economic development.

Some other papers used names as a data source, e.g. other toponyms (Lee & Lin, 2018; Villette & Purves, 2019) or street names (Oto-Peralías, 2017). The distinguishing feature of our dataset is that it is made by interpreting individually researched entries of place names. In our case, we rely on a toponym encyclopedia for (mostly) German places, the *Ortsnamenbuch* by Niemeyer (2012), referred to as “the book” henceforth. This approach allows us to reduce the misidentification of names, and especially of names that occurred many centuries ago and were subject to many iterations of small changes. As far as we know, our dataset is the first that looks directly at the etymology of toponyms on a large scale.

The motivation for the underlying primary data source is discussed in more detail in section 2.2. We discuss the original purpose of toponyms in section 2.2.1, as well as elaborate more on why toponyms may be better suited than the ‘conventional’ data sources for our purposes in section 2.2.2 and 2.2.3. Section 2.3 shows how we have interpreted the original data to create the dataset, and we elaborate on details of the dataset itself. Section 2.4 presents the result on the Zipf distribution, while section 2.5 presents our results on the impact of geographical characteristics on local economic activity. Section 2.6 concludes. The data appendix 2.A presents more details on the book and some example entries, as well as toponym changes throughout history, as well as details on the first recorded mention/city rights year.

2.2 Conjecture and Variety

2.2.1 Purpose of Toponyms

During the Middle Ages, the vast majority of people were illiterate. This is a problem in terms of orientation, as the typical person could not read road signs that mark the location of some place.

Toponyms fulfill the crucial part of conveying the geography of the place of interest without the need to read. They acted as orientation points between places and within the landscape and fulfilled a similar role as postcodes today. That is, toponyms are not chosen to represent some key characteristic of early settlers, but rather offer a glimpse of their immediate surroundings.

The need for this orientation becomes apparent in the case of toponyms that originally shared the same name as they are named after the very same or a similar local geographic characteristics. Typical here is then to change the name with a distinguishing prefix, typically referring to the relative age (old-, new-) or position of the place (north-, south-, high-, low-), thus creating local individualisation of

the toponym. This whole development is described as 'toponymische Raumorganisation' ('toponymical space organisation') (Brendler, 2008).

Even though there is no official convention in the naming of a settlement, it is noteworthy that 'fantastical' names in early settlements that do not describe the geographic characteristic further, such as Friedrichsgabe (present of Friedrich), Glückstadt or Freudenberg (joy place and joy castle) were not common at all, especially not for early settlement. Also, when new names were needed after places were merged or annexed – a process which mostly occurred in the 20th century – a typical suffix is the rather encompassing -tal (valley), suggesting that an accurate description of the local geography through toponyms became not important anymore when most of the population are literate.

It is for this reason that the study of toponyms in our context is best done in the Old World. For instance, it is relatively common to find toponyms that refer to cities in the Old World, such as New York or New Orleans, or honorary names such as Bismarck, or the many Washingtons or Lafayettes. This signals that accurate geographical description embodied in the toponym are no longer needed with the adoption of widespread, sophisticated cartography around the time of settlement. Second, planned cities and as such, names as chosen by a planner or ruler, were not common in the Middle Ages where the vast majority of toponyms were first mentioned (see Figure 2.3a).⁸ A ruler deciding on the name from afar did not happen as often and can be considered a modern phenomenon, especially when a new name is needed after a fusion or annexation of places. As such, we consider the majority of the names that mention some geographical characteristics to be accurately reflected in the local geography.⁹

2.2.2 Interpreting Path Dependence

The existing literature often employs geographical characteristics that are observed today. Although rivers and mountains do not change much over time, some geographical characteristics include characteristics that humans can change, disappear or are characteristics whose importance has altered throughout history. One example are forests that we cannot observe today because they disappeared, clearings of forests to make room for arable land that led to a rapid deforestation, but also an abundance of wood, or castles that were once important for security but became obsolete.

Only focusing on currently measured characteristics ignores these important *historical* aspects and may underestimate the role of first-nature geography on

⁸In our dataset of 3738 places, we count 11 planned settlements.

⁹The individual entries of the book (Niemeyer, 2012) also checks a toponym on plausibility with the local geography.

path dependence. Toponyms give us a unique glimpse on the *original*, i.e. the local and historical, geographical characteristics and are better suited to study path dependence.

We can also be certain that man-made structures already existed around the beginning of settlement, so we do not restrict on buildings that only persisted until today. Moreover, the advantage of restricting ourselves to the earliest buildings is that we can avoid a potential simultaneity bias, which may be the reason why there are few man-made characteristics as compared to natural characteristics considered so far in the existing literature (Jedwab et al., 2020).

2.2.3 Variety of Geography

Previous papers investigated characteristics with a particular emphasize on the role of ports (Acemoglu et al., 2005), (Bosker, Brakman, Garretsen, De Jong, & Schramm, 2008), or portage sites (Bleakley & Lin, 2012), and rivers or sea-access (Bosker & Buringh, 2017) (Henderson et al., 2017) (Rappaport & Sachs, 2003). (Armenter et al., 2014) looked at the effect of bridges on local economic growth.

Nevertheless, the variety of characteristics considered so far is rather limited as compared to the variety of nature. The use of toponyms will therefore not only introduce some non-obvious characteristics, such as hedges, fords, floodplains or clearings, it will also add some more distinction within certain characteristics. For instance, we do not only consider the typical river, but also fords, river islands, river mouths, floodplains or streams. The same goes for altitude, where we can for instance distinguish between mountains, stone, hills or simply 'elevation'. In essence, due to the large variety of meaning in toponyms, we can compare geographical characteristics in a more disaggregated manner, and also substantially expand on the range of first-nature geographic characteristics beyond the current scope of the existing literature.

Table 2.1 presents a list of studies which study the role of geographic characteristic, whether natural or man-made, on local economic activity. For one, most characteristics considered in previous studies are easily observable ones, such as river access, sea access or altitude.

Table 2.1: Local Economic Activity and Geography: Similar Studies

Paper	Sample	Time	Natural	Man-made	Results
Rappaport & Sachs (2003)	USA	2000	ocean coast, great Lakes navigable river		Coast, great lakes and navigable river positive effect on log population density and log employment density
Acemoglu et al. (2005)	Europe	1300-1850		Atlantic ports	Effect of Atlantic Ports after 1600 on log city population points towards importance of Atlantic trade
Bosker, Brakman, Garretsen, De Jong, & Schramm (2008)	Italy	1300-1861	navigable waterway, mountains	roman road, roman road hub, seaport	Seaport, navigable waterway positive and significant, mountains negative and significant
Bleakley & Lin (2012)	USA	1790-2000	portage sites		Portage site places are still population centers today, even though portaging became irrelevant
Bosker et al. (2013)	Latin west and middle eastern cities	800-1800	sea, river	roman road, caravan, roman road hub, caravan hub, bishop, archbishop, capitol, university	Positive effect of sea-access and roman road hub in Latin West log city population, negative effect of sea on middle eastern cities. No effect of rivers.
Armenter et al. (2014)	USA	19th century		bridges	Places with bridges developed more than no-bridge places after bridge construction
Bosker & Buringh (2017)	Europe	800-1800	sea, river, ruggedness, elevation, cultivation probability	roman road, roman road hub	Water based transportation increases in importance over the centuries
Henderson et al. (2017)	World	2010	various biomes, temperature, malaria, precipitation, land suitability, elevation, river, coast, lake, ruggedness	harbor	As for the trade covariates, harbors, rivers, lake positive effect on log light per pixel, while distance to coast has a negative effect

2.3 Data

Our main source Niemeyer’s (2012) *Deutsches Ortsnamenbuch*, which is an encyclopedia of mostly German toponyms, henceforth referred to as ‘the book’. However, the book does not only contain entries from current border Germany, but also includes former German places. Figure 2.1 shows the spatial distribution of each entry in the book. The dataset is better described as containing places of the German Reich right before WWII, as well as Alsace-Lorraine and the German part of Switzerland i.e. the German cultural sphere of influence. The book also only considers places that are above about 7500 inhabitants today.¹⁰

Figure 2.1: Toponym Dataset

Note: Places marked as red have received city rights, while blue places have no city status.

A typical toponym consists of a prefix and a suffix. For example, Gladbach has the suffix “-bach” describing a stream, as well as the prefix “Glad-” describing a smooth (stream). Due to the structure of the German language, typically the suffix conveys the characteristic, while sometimes a prefix further describes this

¹⁰7500 is not a hard cutoff, as a few places are below that cutoff. The book also does not provide any exact date about that population figure.

particular characteristic in more detail.¹¹ Streams alone could for instance be further described as fast, slow, bendy, cold, rich in fish etc. within the prefix.

We may then distinguish between suffix-level toponyms, and further description of the suffix at the prefix-level.¹² At the most disaggregated level, each of these prefixes have their own category. We then have 533 different categories in the dataset and this needs to be reduced for our analysis.¹³ In this paper, we will only consider suffix-level characteristics i.e. we will not consider the prefix of a specific geographic characteristic suffix, which will reduce the different *types* of characteristics to 242. That is for Gladbach, we ignore the 'smooth' of 'smooth stream'.

Further, for Mönchengladbach, the Mönchen (Monk, refers to a monastery) part was later added to differentiate with other "smooth stream" places, which we, as discussed in previous section, also ignore as we are interested in the first recorded name.¹⁴

The same remark applies to types, which are not readily apparent and could be a geographic characteristic itself, not a description of the suffix. For instance, do we consider Ochsenfurt (-furt suffix with ox prefix, "ox ford") as a ford with the type ox, or a ford, where oxes live nearby? Does the ox describe the ford, or are they separate geographic entities? Note that we do not consider the prefix ox as a way to differentiate with other, nearby or not, -ford places, as long as the ox prefix is mentioned in the earliest mentioned record.¹⁵

There also are combination of prefix and suffix which describe two different geographical characteristics, such as Querfurt (-furt (ford), Quer (mill), so "mill ford") or Steinhagen (-hagen (hedge), stein (stone), so "stony place hedge"). Note that stone prefix of Steinhagen could also refer to a mountain, instead of a stony place referring to the ground. The book typically provides further clarification and otherwise, we did additional research ourselves. In these clear and rare cases, we assign both geographic meanings.

We aggregate names that have the same meaning in terms of their geographi-

¹¹This is typically not the case for old, Roman cities, such as Köln (colony) or Bonn, which has no such prefix or suffix. We treat these as standalone suffix toponyms without any prefix.

¹²We may here think of this distinction similar to product harmonization in international trade, where a phone would be at the suffix-level, while a blue phone would be at the prefix level. We can then think of toponyms, due to its purpose of offering descriptions to the illiterate, as providing a similar harmonization standard.

¹³Note here that we have not given individual categories for smooth rivers or smooth streams for example. If we have done so, we would have even more than the 533 categories.

¹⁴Note that the typical German toponyms are such that the suffix carries the geographical meaning most of the times, while the prefix also often carry a personal name e.g. Tübingen (-ingen suffix for the personal name Tuwo, "with the people of Tuwo").

¹⁵Therefore, we also assign the ox- prefix rather than ignoring it. However, this and other similar examples are very rare in the data and do not enter our empirical analyses due to use of a cut-off frequency.

cal characteristics. This typically involves aggregating names which have different origins, but the same meaning e.g. the celtic -dunum, -stein (as in "stone"), -berg/burg ("castle" itself), or -grad/grod ("castle", but slavic e.g. Graz, Belgrad). However, there was little distinction in early times between a castle and a Schloss, which nowadays tends to be translated as "palace". The issue here is whether we should aggregate Schloss to castles in this case. Throughout the aggregation process, issues similar to these occurred, such that we relied on our own judgement which characteristics are similar enough, and which are entirely different characteristics.

Table 3.8 presents each aggregate characteristic, and the name and description/translation of the individual characteristic aggregated. The typical toponym suffix is shown in the second column, and the book often cross-references these suffixes with another in each entry which reveal their relatedness. It was then mostly straightforward to aggregate them into the different types, as the book has done the most work already. However, we only consider the types in this study, but we also describe the uncommon characteristics which we will not consider. Figure 2.2a shows the kernel estimates of the first recorded mention of the 5 most common aggregated variables with a geographic meaning, as well as of the places with no geographic meaning in Figure 2.2b.¹⁶

We find a bimodal distributions especially for toponyms without geographic meaning, streams and rivers; as well as castles, mountains and forests. The likely explanation for the two peaks can be found in the Frankish colonisation, which took place in the three centuries following the Battle of Zulpich in 496AD. It seems that Frankish colonisation settlements primarily have names without geographic meaning, or have the names of streams and rivers.

The kernel estimates in Figure 2.2b illustrates this in a sense that typical Frankish colonisation settlements without geographic meaning, such as -heim or -hausen, are first mentioned around that time. Also notice the gallo/roman -acum suffix ('place of'), which also occur before 500AD, and the -dorf suffix, which is not associated with the Frankish colonisation.

Furthermore, to measure modern urban development, we consider the 1910 population census data as the size of these German places in 'modern' (early 20th century) times, and not today's population data. The single most important reason for this decision is the widespread mergers and annexation of places in the course of the 20th century, which was a relatively rare occurrence until 1910. A consequence

¹⁶For those, we only consider places within the current borders of Germany as of now in order to see this development. The reason for the restriction on Germany here is that the book, or rather the researchers who contributed the articles to the book, are biased towards places within the current German border, and not those within the border of former German territories. As such, we should be more consistent at least in some given geographic border.

Figure 2.2: Distribution of Year of First Record and City Rights Awarded

(a) First Record Distribution of Common Geographic Characteristics

(b) First Record Distribution of Some Non-Geographic Meaning

Note: These graphs reflect the first recorded mention of the 3210 current border Germany places in our dataset. The -acum suffix is typically Roman.

of this merger and annexation wave is the lack of population data today on very small, already annexed or merged places. That is, the 1910 population census data is more complete than today's population data.

There was however not a single census which encompassed the whole of the German Reich in 1910, but many single regional ones. Fortunately, the 1910 population data can be found on Gemeindeverzeichnis.de, which provides comprehensive population data for almost all of the places in our dataset, as they compiled the population data on a municipality level within the German Reich borders of 1910.^{17,18}

Lastly, city rights are either taken from the book which is not always provided there, so that the rest have been gathered by ourselves.¹⁹ Similarly, the first recorded mention is also found in the book. However, the book is not precise sometimes with the exact year, but only provides the century. For those cases, we have also tried to get the exact year.²⁰ Figures 3a and 3b provide the distribution of the first recorded mention, as well as the first mentioned city rights.

Table 2.2 provides the descriptive statistics of the population and year data.²¹

2.4 Zipf Distribution

If first nature is relevant for a city size distribution, it should be consistent with a rank-size curve or Zipf's law. Gabaix (1999b) and Gabaix (1999a) provide an explanation why city size distributions may behave according to Zipf's law. If city growth follows Gibrat's law with a common mean and variance, then a Zipf distribution follows, at least in the upper tail. Krugman (1996) conjectured that if we assume that the variation of a static landscape is random, it implies a Zipf distribution in the upper tail.

Our dataset allows us to consider the relationship between first-nature geography and Zipf's law. In section 2.4.1, we look at whether the frequency of geographical characteristics as suggested by its toponym follows a Zipf distribution.

¹⁷The list of individual 1910 censuses can be found on the website: www.gemeindeverzeichnis.de/gem1900/gem1900.htm?quellen/quellen.htm.

¹⁸As for the Austrian, Swiss and Czech, Belgian and Luxembourg places, the 1910 data can be readily found from their respective Statistical division.

¹⁹Typically, we can find those city rights from the official website of the given place. For each city right not found in the book, we have provided the source in the dataset.

²⁰The exact year is typically found on the official website of the given place.

²¹Important to note here is as the researchers in the book are primarily German, it is shown by the relatively low rate of places with city rights within the German border in our dataset, while the rate is generally higher for outside of Germany. It may reflect that it would take a bigger city outside of Germany for some researcher to be interested in interpreting the name. It may also reflect the migration wave eastwards, characterized by a higher rate of city rights given to these places to attract new settlers. We should then think of individual, current border country dummies as not only accounting for regional differences, but also the tendencies of German toponym researchers in the book to not interpret toponym in certain countries as frequently.

Table 2.2: Descriptive Statistics

Variable	mean	std. dev.	min	max	observations
Full Sample					
Log Population (1910)	7.908	1.311	2.303	14.550	3738
Log Urban Potential (1910)	10.905	1.190	7.684	19.243	3738
First Record Year	1098.551	236.382	-40	1910	3738
First City Right Year	1518.885	319.132	50	2019	2118
Austria					
Log Population (1910)	8.807	1.038	6.596	14.550	127
Log Urb. Pot.l (1910)	11.133	1.005	9.170	16.747	127
First Record Year	1036.228	248.524	50	1352	127
First City Right Year	1593.619	340.566	1014	2001	113
Belgium					
Log Population (1910)	8.115	0.566	7.715	8.516	2
Log Urban Potential (1910)	10.936	0.540	10.554	11.317	2
First Record Year	889.5	341.533	648	1131	2
First City Right Year	1175	247.487	1000	1350	2
Czech					
Log Population (1910)	9.785	0.636	8.563	10.874	40
Log Urb. Pot.l (1910)	12.280	0.704	11.120	13.566	40
First Record Year	1221.375	128.526	993	1506	40
First City Right Year	1471.25	261.534	1223	1952	40
France					
Log Population (1910)	8.414	1.340	6.180	12.095	44
Log Urb. Pot.l (1910)	11.313	1.327	9.127	15.111	44
First Record Year	981.477	266.259	350	1629	44
First City Right Year	1283.8	136.168	1150	1580	15
Switzerland					
Log Population (1910)	8.363	0.944	6.153	12.281	146
Log Urban Potential (1910)	11.154	0.927	9.017	15.200	146
First Record Year	997.253	253.457	150	1788	146
First City Right Year	1453.116	326.820	1045	2019	43
Germany					
Log Population (1910)	7.749	1.275	2.303	14.544	3210
Log Urb. Pot.l (1910)	10.829	1.200	7.684	19.243	3210
First Record Year	1099.802	234.445	-40	1910	3210
First City Right Year	1530.733	323.894	50	2017	1754
Italy					
Log Population (1910)	8.769	0.991	7.751	10.323	8
Log Urb. Pot.l (1910)	11.127	0.984	10.025	12.639	8
First Record Year	816.5	349.276	150	1256	8
First City Right Year	1345.2	124.719	1200	1536	5
Lithuania					
Log Population (1910)	9.974		9.974	9.974	1
Log Urb. Pot.l (1910)	11.660		11.660	11.660	1
First Record Year	1253		1253	1253	1
First City Right Year	1328		1328	1328	1
Luxembourg					
Log Population (1910)	8.368	0.946	7.072	10.718	17
Log Urb. Pot.l (1910)	11.177	0.923	9.955	13.516	17
First Record Year	997.118	210.390	634	1274	17
First City Right Year	1357.8	305.886	952	1907	10
Poland					
Log Population (1910)	9.424	0.869	6.797	13.146	133
Log Urb. Pot.l (1910)	11.601	0.880	9.190	15.368	133
First Record Year	1255.624	117.768	970	1645	133
First City Right Year	1376.532	195.665	1187	1959	126
Russia					
Log Population (1910)	9.574	1.326	8.481	12.413	9
Log Urb. Pot.l (1910)	11.381	1.272	10.408	14.185	9
First Record Year	1351.889	105.292	1255	1580	9
First City Right Year	1555.333	199.080	1286	1772	9

Figure 2.3: Distribution of Year of First Record and City Rights Awarded

(a) First Record

(b) City rights

Note: Graph (a) reflects the distribution of the first recorded mention, full sample. Graph (b) reflects the distribution of the first mentioned city right/mentioning as a city as elaborated in section 2.3, full sample.

In section 2.4.2, we introduce a proxy measure for early geographical advantage, defined by the year gap of city rights received and the year of the first recorded mention, and whether this proxy follows a Zipf distribution

2.4.1 Frequency Distribution of Geographical Characteristics

With toponyms, we look at early settlement patterns to some specific geographical characteristics, capturing the choice where to settle first. They provide a straightforward way to assign local characteristics to settlements, without having to choose a maximum distance between the settlement and the characteristic. That is, one advantage of using toponyms is that the implied characteristics are measured in a binary, which simplifies quantification. As each characteristic is measured in the same manner, it allows us to consider the frequency distribution of a large variety of characteristics.

We consider our approach to have a more direct link between geographical characteristics and the Zipf's distribution for settlements.²²

Following (Gabaix, 1999a), the probability of the frequency $freq_l$ of a geographic characteristic l being larger than some F is equal to

$$Pr(freq_l > F) = \frac{k}{F^\zeta} \quad (2.1)$$

for some positive constant k and $\zeta = 1$. In this context, we can then regress the following through ordinary least squares (OLS):

$$\log(rank(freq_l)) = \alpha - \zeta \log(freq_l) \quad (2.2)$$

for some constant α . When the negative slope ζ being estimated is close to 1 and with a straight line i.e. high R^2 , this is to be taken that the frequency of geographic characteristics follows a Zipf distribution.

In panel (A) of Table 2.3, we only consider places with geographic meaning i.e. no -ingen, -heim, -dorf etc places. In both samples we approach a Zipf coefficient ζ of 1 as we increase the minimum frequency of a given geographic characteristic, which is a similar pattern to what the literature on city population Zipf finds. As noted by Brakman et al. (1999), between a certain frequency in our case, the negative correlation between rank and frequency at the upper tail may break down for a lower frequency. Specifically for our case, one explanation for this is that integers may be an issue for lower frequencies, which combined with many unique names

²²As compared to Villette & Purves (2018), who show a Zipf distribution of microtoponyms i.e. names given to natural features, in the Canton of St. Gallen, our dataset has the advantage that we consider a large part of Central Europe. Ours also uses the actual interpretation of the earliest known name by linguists, while theirs is based on a lexicon of typical toponyms that is subject to misinterpretations discussed in Appendix A2.

that cannot be aggregated easily, hence leading to a lower estimated coefficient. Figure 2.4 shows the typical Zipf plot for the full sample first specification of Table 2.3, with no minimum frequency. We have added random noise to make the many overlapping points at the lowest frequencies more visible.

In panel (B) of Table 2.3, we additionally restrict ourselves to only place with natural geographic meaning i.e. we do not consider man-made geographic characteristics anymore. Again, we approach a Zipf distribution as we increase the minimum frequency. As such, this settlement behavior around geographical characteristics exhibits a Zipf distribution in the upper tail.

However, although these results are intriguing and support of the idea that Zipf's law for cities may have its origins in nature, as suggested by (Krugman, 1996), more evidence is warranted, ideally by measuring an individual location's geographical advantage or land quality early in time.

Table 2.3: Zipf and Toponyms

minimum frequency	Full Sample			Germany		
	coefficient	number of geographies	R^2	coefficient	number of geographies	R^2
Panel A: Geographic Meaning						
≥ 0	-0.659 (0.008)	194	0.971	-0.673 (0.009)	177	0.972
≥ 5	-0.826 (0.014)	75	0.980	-0.813 (0.014)	75	0.978
≥ 10	-0.887 (0.022)	46	0.974	-0.884 (0.022)	46	0.974
≥ 15	-0.934 (0.028)	35	0.971	-0.938 (0.027)	35	0.973
≥ 20	-0.952 (0.034)	30	0.967	-0.969 (0.031)	30	0.972
≥ 25	-0.972 (0.046)	23	0.954	-1.007 (0.041)	23	0.966
≥ 30	-1.027 (0.062)	18	0.945	-1.048 (0.055)	18	0.957
Panel B: Additionally No Man-made Geographies						
≥ 0	-0.632 (0.010)	142	0.965	-0.645 (0.011)	130	0.964
≥ 5	-0.823 (0.017)	56	0.978	-0.811 (0.018)	56	0.975
≥ 10	-0.889 (0.021)	39	0.980	-0.879 (0.022)	39	0.978
≥ 15	-0.918 (0.029)	29	0.973	-0.915 (0.029)	29	0.973
≥ 20	-0.928 (0.037)	24	0.967	-0.941 (0.036)	24	0.968
≥ 25	-0.932 (0.052)	18	0.952	-0.971 (0.052)	18	0.957
≥ 30	-0.961 (0.073)	14	0.936	-0.982 (0.072)	14	0.939

Note: The coefficients show the OLS regressions of the log rank with the log frequency of geographical characteristics i.e. ζ in equation (2). Standard errors are in parenthesis. We consider Germany in its current borders.

2.4.2 Distribution of Early Geographical Advantage

One shortcoming of the above exercise is that it only addresses that the frequency of settlements along geographical characteristics follows a Zipf distribution. It does not directly link these characteristics to a city size distribution. The inherent characteristics of a location are best seen as an index of geographical characteristics, which determines a location's economic activity (see Gabaix & Ioannides (2004)).

Figure 2.4: Zipf Frequency of Geographical Characteristics (1. Specification)

Note: We consider the full sample of Table 2.4 panel A, "Geographic Meaning" in this figure with no minimum frequency, on the full sample. We have added random noise (in STATA, option jitter(2)) to make overlapping points a little more visible.

We measure this implied index of geographical advantage itself through the gap of the year between the first city rights given and the first recorded mention of a city.

If the time between the first mentioning of a city and the year in which city rights were granted is small, the place has undergone rapid development and we interpret this as an indication of favorable first-nature characteristics. First, this link should be relatively stronger in the absence of increasing returns effects or when first-nature geography is a relatively stronger determinant of local economic activity i.e. when population levels are not very high. This would be the case when a location was initially settled, as there was little human interaction or build up of endogenous amenities early on.²³

²³Even around the year 1500 in the late Middle Age within the Holy Roman Empire, around 90 to 95 percent were below a population of 2,000. Furthermore 85 percent of these cities were considered dwarf cities with a population of below 1,000 (Schilling (2004) p. 8), which would make the bulk of the cities in our dataset. For these dwarf cities, the transition between rural and urban is considered 'smooth' (Stoob (1985) p. 152) here. Within the literature, Bosker, Brakman, Garretsen, De Jong, & Schramm (2008) and Bosker et al. (2013) measured both the impact of first- and second-nature geography on economic activity throughout time. While the first paper on Italian cities of above 10,000 between 1300AD to 1861AD find no significant impact of second-nature geography as proxied by an urban potential measure, the latter using the 'Bairoch et al. (1988)' dataset with above 10,000 inhabitants in Western European, as well as a unique dataset on North African and Middle Eastern cities does. However, our dataset includes much smaller dwarf cities as well, which is reflected by the larger number of cities that we can consider

However, we can think of some characteristics to be predestined as natural second-nature places early on. For example we may think of choke points such as fords, bridges and hedges as a natural hub for merchants, where trade facilities developed relatively early on. Following Duranton & Puga (2004), increasing returns effects would then come from the sharing of indivisible facilities. Marketplaces were previously conjectured to play a role in second-nature geography (see Cronon (1991) p.56, Konishi (2000) and Berliant & Konishi (2000), and Duranton & Puga (2004) for a short discussion on their microfoundations).

However, these studies did not explore the relevance of market places empirically. Now the point around which markets were allowed to take place in the Middle Age were after obtaining either market rights or city rights. In the book, they were mentioned as 'Markt toponym' if they only have known market rights, and typically as 'Stadt toponym' if they have city rights. This makes it difficult to see whether places received market rights before receiving city rights. However, the number of mentioned market rights that we have recorded in our dataset is low with 288 cases, so that for the majority, market rights were likely skipped to obtain city rights directly. As such, we consider city rights to be a more comparable proxy for when markets were allowed to take place.

The years between the first recorded mention until city rights can then be considered a time where important indivisible facilities such as a marketplace, criminal courts or walls did not exist yet, thus restricting the second-nature impact through sharing. Another source may come from the sharing of a relatively wider variety of intermediate inputs that were imported. We may measure this through the variety of guilds that were present in a city. For example, Lübeck as the de facto capital of the Hanseatic League and had around 36 different guilds at the beginning of the 15th century, while Braunschweig had around 13 in year 1445 (Tuckermann, 1906), despite their roughly equal population of 25,000 and 18,000 respectively in 1400 (Bairoch et al., 1988). However, we again consider these sources to be mostly relevant to the largest cities after receiving city rights, whereas the majority of cities in our dataset would still be classified as dwarf cities.²⁴

In any case, we will proxy the inherent, inverse early geographical advantage by the difference of the year when city rights were first issued and the year of the first recorded mention for our place i , that is:

around Germany. For example, the number of cities in the Bairoch et al. (1988) database includes 244 cities within the current German border.

²⁴In a similar line of argument stemming from intermediate inputs as well as from matching and learning, there were designated streets for some professions, as indicated by its street name. In the city of Lübeck for instance, bell founders were found around the Glockengießerstraße, butchers around the Fleischhauerstraße, and hucksters around the Huxstraße. This may be of interest for future research in a similar vein.

$$\log(\text{gap}_i) \equiv \log(\text{cityright}_i - \text{firstrecord}_i) \quad (2.3)$$

A small gap corresponds to a large inherent advantage very early in its settlement, and vice versa. Given the previous discussion, we thus assume that second-nature causes are not relevant while first-nature causes are, at least until the city receives its city right. For our purposes, the advantage to a 'traditional' historical economic activity measure, such as the population estimates of Bairoch et al. (1988), is that this proxy allows us to consider all cities in the data set and it is not limited by the availability of data in a given year. A quite similar rationale to this measure was made by Ioannides & Overman (2004), who indirectly measured variations in first-nature geography through the age of places in the US, as they reasoned that places that were settled earlier were better places in terms of first-nature characteristics. On the other hand, our rationale is that good places in terms of first-nature were settled faster from the point of a primitive rural settlement, measured by first recorded mention, until a threshold was reached where it was considered a city *de jure*. Arguably, this would be a more direct measure for a location's strength, as we look at first-nature around the time of first settlement with a clear threshold point.

City rights were issued in two waves (see Figure 2.3b): The initial wave started in the High Middle Age after 1150 until around 1700, typically for a fee to the feudal lord, cities were given rights to hold markets, to fortify, to create courts or for staple rights that made traders hold their products in warehouses for around 3 days that will then be sold to the citizens first. Hence, a city typically needs to be of a certain size to afford city rights. The second wave started during industrialization and were typically given if the place reached a certain size, or there was merely the wish of the place to be called 'city'. To abstract from increasing urbanization during industrialization, which has less to do with some inherent locational advantage, we will therefore only consider city rights that were issued before 1700 in Table 2.4, where we estimate the Zipf coefficient of our geographical advantage proxy.

The probability of the inverse early geographical advantage gap_i for some city i being smaller than some A is equal to

$$\Pr(\text{gap}_i < A) = \frac{c}{A^\eta} \quad (2.4)$$

for some positive constant k and $\eta = 1$. In this context, we can then regress the equivalent log-linear version through ordinary least squares (OLS)

$$\log(\text{rank}(\text{gap}_i)) = \alpha + \eta \log(\text{gap}_i) \quad (2.5)$$

We now take a positive slope η being estimated close to 1 and with a straight line i.e. high R^2 , as the gap of city rights and first record year following a Zipf distribution.

(Gabaix & Ioannides, 2004) showed with a simple model that, if η_{pop} is the power law exponent of the population distribution across, in our case, places, and we show $\eta_{pop} = \eta = 1$, then geographical advantages can be an explanation of Zipf's law for cities/places.^{25 26}

Table 2.4 shows that our measure for geographical advantages of early settlements indeed exhibits a Zipf distribution. Furthermore, similar to the frequency of characteristics in section 2.4.1, we observe that the behaviour of the coefficient approaches unity as we decrease the maximum difference. Figure 2.5 shows the relation of the rank of the gap and the gap. Notice that for high values of the gap i.e. low geographical advantage places, the points do not fit the linear fit line perfectly. A simple explanation is as the book only considers places above around 7500 population today, we would have excluded relatively more low geographical advantage places in our sample. Simply expanding the sample would likely create a better picture in the upper tail. Another explanation for this deviation is those represent cities that were founded much earlier, before the first modern city rights around 1150 spread throughout the region.²⁷ We can control for this by adding bicentennial dummies, starting from year 1 in panel B of Table 2.4. The estimates are now even closer to unity as compared to the specification without controls of panel A. In Figure 2.6 for the purpose of illustration, we show a more radical restriction than adding bicentennial dummies, where we instead only consider cities which were first mentioned in and after 1150. There is now no noticeable deviation from the linear fit at the upper tail anymore, which is in line with our suspicion that this deviation may partly be driven by cities which were founded before the first modern city rights. We will hence control for this in the rest of this article.

If we take our 'gap' proxy as appropriate to measure early settlement geographical advantage, the results in Table 2.4, together with the results of section 2.4.1,

²⁵We may additionally require that this equality holds at the given time, so that we require $\eta_{pop,t} = \eta_t$, if $\eta_{pop} \neq 1$ at the time of interest t . We do not have this population data for this many places around the Middle Ages, but the literature suggests that Zipf's law is a stable empirical regularity throughout history. Hence, we consider an estimated $\eta = 1$ as not the definitive answer on whether geographical advantages explains Zipf's law, but a strong suggestion that it does.

²⁶(Barjamovic et al., 2019) inferred population sizes from ancient Assyrian trade data and a structural model, and found a Zipf distribution for the estimated population of some of the largest cities during the Bronze Age in Anatolia. Our approach does not directly measure actual population sizes as well; our proxy can be interpreted as the growth rate between a primitive settlement until a mature city.

²⁷1160 can be considered the year where the first modern city rights spread first by the Hanseatic city of Lübeck in 1160, which was derived from the city rights of Soest. Due to our way of dealing with sources that merely mention centuries by assigning it the middle of the century i.e. 1150, and not the exact or approximate year, we consider 1150 as the 'start' of the modern city rights in our sample.

support the conjecture of geographical advantages being itself Zipf distributed as proposed by Krugman (1996). These results may also open up alternative explanations for Zipf's law that do not rely on some form of Gibrat's law with a reflexive lower bound (Gabaix, 1999b; Eeckhout, 2004; Rossi-Hansberg & Wright, 2007; Duranton, 2006). For example Lee & Li (2013) show a static model where the city size is determined by a product of multiple random factors which are i.i.d., a subset of these factors being represented by heterogeneous exogenous physical features. Importantly, they generate a Zipf distribution without the need for Gibrat's law. In any case, we have showed that the empirical regularities pertaining to city-size distributions also have a bearing on the first-nature geography of places or cities. In the next section we will investigate which specific geographical characteristics help to explain the size of German cities in modern times.

Table 2.4: Zipf and Difference of City Rights pre-1700 and First Record Year

	Full Sample			Germany		
maximum difference	coefficient	observations	R^2	coefficient	observations	R^2
Panel A: no additional controls						
no restrictions	0.822 (0.004)	1279	0.977	0.868 (0.004)	1042	0.982
≤ 700	0.846 (0.003)	1228	0.986	0.887 (0.003)	1007	0.989
≤ 600	0.852 (0.003)	1199	0.986	0.892 (0.003)	983	0.989
≤ 500	0.867 (0.003)	1094	0.988	0.906 (0.003)	890	0.990
≤ 400	0.894 (0.003)	933	0.990	0.930 (0.003)	741	0.991
≤ 300	0.921 (0.003)	796	0.993	0.958 (0.003)	616	0.993
≤ 200	0.950 (0.003)	607	0.993	0.986 (0.004)	461	0.993
≤ 100	0.980 (0.005)	343	0.990	1.020 (0.007)	252	0.990
Panel B: with bicentennial dummies						
no restrictions	0.886 (0.004)	1279	0.989	0.919 (0.004)	1042	0.990
≤ 700	0.893 (0.003)	1228	0.991	0.927 (0.004)	1007	0.992
≤ 600	0.895 (0.003)	1199	0.991	0.930 (0.003)	983	0.993
≤ 500	0.901 (0.003)	1094	0.992	0.936 (0.004)	890	0.993
≤ 400	0.912 (0.003)	933	0.992	0.947 (0.004)	741	0.993
≤ 300	0.931 (0.003)	796	0.993	0.966 (0.004)	616	0.993
≤ 200	0.955 (0.003)	607	0.994	0.989 (0.004)	461	0.994
≤ 100	0.981 (0.006)	343	0.991	1.020 (0.007)	252	0.990

Note: The coefficients show the OLS regression of the log rank with the log year gap of city rights received and first record mention of the toponym. Panel A shows the estimates of η for equation (2) without any additional controls, while panel B shows the estimates with additional bicentennial dummies, starting from year 1. Intercept estimates are not shown here. Standard errors are in parenthesis. We consider Germany in its current borders.

Figure 2.5: Zipf Gap (City Rights pre-1700)

Note: We consider the full sample of Table 2.4, Panel A in this figure with no additional restrictions.

2.5 Geographical Characteristics and Local Economic Activity

2.5.1 Early Characteristics on Modern Population

We will now turn to our second main question as to which of the initial geographical characteristics help to explain the variation in German city sizes. We will estimate the impact of these geographical characteristics, as suggested by the toponym, on the population of 1910. That is, we write:

$$\log(pop_i)_{1910} = \alpha + \beta_1 nomeaning_i + \gamma_2 X + \beta_3 firstrecord_i + \beta_4 UP_i + \delta_5 Z + \epsilon \quad (2.6)$$

where, *nomeaning* is the variable which aggregated the no geographical feature places (-ingen, heim, dorf), where X is a vector of the first-nature types we have described above. Z is a vector of control variables, that could contain a current border country dummy; latitude, longitude, their squared and interaction terms; and a dummy whether a place has city rights. *firstrecord* is a variable for which we have the very first recorded mention of the name. Lastly, UP is a simple urban potential control on the 1910 population data:

Figure 2.6: Zipf Gap (City Rights ≤ 1700 , First Record ≥ 1150)

Note: We consider the German sample in this figure. We restrict the sample to cities that were first mentioned in 1150 and after. This figure corresponds most closely, but not directly, to Table 2.4 panel B with bicentennial dummies starting from year 1 and no additional restrictions. The slope of the linear fit is 0.967, with 364 observations.

$$UP_i = \sum_j^n \frac{pop_{i,1910}}{\tau_{i,j}} \quad (2.7)$$

where we, the same as Bosker, Brakman, Garretsen, De Jong, & Schramm (2008); Bosker et al. (2013) set distance to linear affect UP. However, we simplify the measure by not giving different distance weights depending on the implied mode of transportation to reflect estimated differences in Medieval transport cost. One reason is the emergence of railroads that made such discounting less relevant in 1910. That is, $pop_{i,1910}$ is the population in 1910, while τ is the simple euclidean distance measure in kilometers from place i to place j .

Table 2.5 presents some results of the OLS regression of log population in 1910 on some variables of separate geographic characteristic groups; grouped by water, buildings, elevation, plants, land and resources, as well as a current border country control. For instance in column I, we find that places which refer to some river are around 50.8 percent and those that refer to a river mouth, which is a confluence of a water body into another, even by around 151.7 percent. Generally, the estimates stay the same when we consider all groups in column VI.

Our treatment assignment, which in this case are geographical characteristics as implied by its toponym, are clustered regionally. Some regions may be more mountaintaneous for example especially in the south of Germany, while some can be considered swampy or bog like as is in the region of Friesland. Hence, the treatment assignment may be clustered. Following Abadie et al. (2017), the appropriate level to cluster now is at the individual place level and not, for instance, at a district level. Note that due to our cross-sectional data structure, individual level clustered standard errors are the same as heteroskedasticity robust standard errors.²⁸

We add more controls in Table 2.6: a first record year control, a city right dummy, longitude and latitude controls as well as the urban potential measure. Following Giesen & Suedekum (2014), US settlements that were founded earlier are likely larger now. Hence, some of the coefficients may be biased upwards by earlier settlements being founded disproportionally around some characteristics. Therefore, we add the first recorded mention in column VII to control for this. To a certain degree, this will also control for some regional differences, as settlements were initially in the regions West of the Rhine, with a wave of settlements towards the North East only happening after around 800AD with the Frankish Colonisation and later the Ostsiedlung (Eastern Settlement). The sign for the first record is expectedly negative, which is consistent with the findings of US city age being positively correlated with city population (Giesen & Suedekum, 2014)

Except for streams, all statistically significant estimates are now slightly attenuated as compared to column VI. Adding latitude and longitude, and their squared and interaction terms in VIII-XI also do not change the estimates much. It is only once we add the log urban potential in 1910 measure (7) in X that the coefficients attenuate.

Nevertheless, the results indicate that places named after rivers are on average much larger than stream places even today. Furthermore, places with a ford on the other hand that may serve natural trading hubs through land and water show an even larger effect. The same holds true for river mouths as crossing point or confluence between two water bodies together, which together indicate the importance of characteristics associated with trade as compared to those associated merely with

²⁸However, the use of dummy variables where the assignment is sparse i.e. unique and seldom toponyms, could lead to overly deflated robust standard errors. Interestingly, we find lower heteroskedasticity robust standard errors as compared to the conventional OLS standard errors of the first and third columns at least for the dummy variables, while the opposite is the case for the non-dummy, control variables. A simple explanation for it is if we were to aggregate all the dummy variables, the 0's could have a higher variability than the 1's. This may stem from the many unique and rare names in the dataset, which are not included in the specification of Table and would simply be assigned a 0 in this case. A large population size variability between the unique names is not implausible, as it could be driven by rare desirable and undesirable locations. For example, the name Stuttgart is derived from 'stable', or Stralsund as a 'strait' are respectively the two only instance of its kind in our dataset. The four places we assign a 'bay' are the relatively big cities of Hamburg, Kiel, Flensburg and Schleswig.

water-access. In contrast the purpose of hedges was not only to defend against attackers as part of the system of the so called 'Landwehr', and as such can also be artificial and not natural; but it can also act as a funnel to collect tolls (Küntzel, 2004). That is, we may think of hedges as important to land-based trade, where we find an overall negative lasting impact.

Castle places as a form of security are still on average larger today, even though the need for this kind of security has long disappeared. In column IX, this relation has disappeared after we added the 'before 1910 city rights dummy'. This is not surprising, as places with city rights are typically much larger than those without, thus attenuating positive and significant coefficients more here.

In conclusion, the results overall support the importance of characteristics that we associate with water-based trade. However, to say something about the persistence of these effects, we would need to estimate their importance also early in a settlements life.

2.5.2 Impact of Early Characteristics on Early Geographical Advantage

We will repeat the exercise of section 2.5.1, however with the dependent variable being the early settlement geographical advantage measure, as introduced in previous section 2.4.2. The idea is to uncover the determinants of early geographical advantages as proxied by our 'gap' variable.

That is, we estimate:

$$\log(\text{gap}_i) \equiv \log(\text{cityright}_i - \text{firstrecord}_i) = \alpha + \beta_1 \text{nomeaning}_i + \beta_2 X + \beta_3 \text{firstrecord}_i + \delta_4 Z + \epsilon \quad (2.8)$$

where, cityright_i is the year a place i has first received the city right, firstrecord_i the year the place was first mentioned in a record, nomeaning is the variable which aggregated the no geographical feature places (-ingen, heim, dorf), and X is a vector of the first-nature types we have described above. Z is a vector of control variables, that could contain a current border country dummy; latitude, longitude, their squared and interaction terms; and a bicentennial dummy.

Table 2.7 shows the baseline results. Notably, water-based trade characteristics, such as rivers and fords, were apparently not important very early on in time. At first, it seems to deviate from the literature that stresses trade as important determinants of early development of settlements, such as Bakker et al. (2019) for the ancient Phoenician trade in the Mediterranean Sea. In contrast, these results somewhat support the findings of Henderson et al. (2017) that characteristics that are suitable for water-based trade became *relatively* important over time, as compared to characteristics associated with agriculture, which in our case would be the

Table 2.5: Impact of Geographical Characteristics on 1910 Population: Baseline

Type (frequency)	I	II	III	IV	IV	V	VI
no geographic meaning (1389)	-0.156 (0.045)***	-0.146 (0.043)***	-0.183 (0.042)***	-0.192 (0.042)***	-0.184 (0.042)***	-0.174 (0.042)***	-0.158 (0.055)***
Water							
stream (388)	-0.164 (0.069)**						-0.175 (0.075)**
river (249)	0.411 (0.082)***						0.373 (0.084)***
floodplain (92)	0.053 (0.106)						0.065 (0.113)
ford (39)	0.789 (0.242)***						0.769 (0.243)***
source (37)	-0.319 (0.200)						-0.326 (0.204)
lake (29)	-0.249 (0.201)						-0.303 (0.201)
water (20)	0.303 (0.224)						0.329 (0.249)
riv. mouth (17)	0.923 (0.312)***						0.923 (0.314)***
riv. island (15)	0.017 (0.272)						0.015 (0.274)
Buildings							
castle (281)		0.336 (0.081)***					0.297 (0.085)***
church (77)		-0.125 (0.120)					-0.118 (0.123)
farm estate (67)		-0.541 (0.112)***					-0.557 (0.115)***
monastery (45)		0.010 (0.186)					-0.006 (0.189)
bridge (25)		0.307 (0.282)					0.256 (0.268)
fence (24)		0.246 (0.162)					0.217 (0.167)
mill (20)		0.540 (0.338)					0.561 (0.335)*
Elevation							
mountain (174)			-0.137 (0.098)				-0.115 (0.102)
elevation (35)			0.131 (0.217)				0.114 (0.218)
stone/cliff (34)			0.123 (0.221)				0.096 (0.231)
hill (26)			-0.085 (0.169)				-0.125 (0.184)
valley (22)			0.112 (0.217)				0.129 (0.216)
Plants							
tree (148)				-0.151 (0.104)			-0.095 (0.105)
forest (137)				0.063 (0.114)			0.072 (0.117)
hedge (77)				-0.680 (0.155)***			-0.652 (0.158)***
clearing (82)				0.103 (0.132)			0.145 (0.135)
bush (17)				0.162 (0.213)			0.158 (0.222)
Land							
field (108)					-0.075 (0.122)		-0.072 (0.124)
swamp (45)					-0.036 (0.229)		-0.031 (0.234)
bog (29)					-0.786 (0.219)***		-0.782 (0.210)***
land (28)					0.052 (0.215)		0.030 (0.220)
Resources							
salt (24)						0.357 (0.187)*	0.348 (0.184)*
reed (18)						0.125 (0.236)	0.211 (0.255)
Controls							
country FE	yes	yes	yes	yes	yes	yes	yes
Observations	3738	3738	3738	3738	3738	3738	3738
R ²	0.1306	0.1244	0.1158	0.1208	0.1179	0.1155	0.1486

Note: Dependent variable is log population in 1910. *, **, *** denotes significance at the 10%, 5%, 1%. Robust standard errors are given in parenthesis.

Table 2.6: Impact of Geographical Characteristics on 1910 Population: Robustness

Type (freq.)	VII	VIII	IX	X	XI
no geographic meaning (1389)	-0.236 (0.054)***	-0.222 (0.053)***	-0.017 (0.010)*	-0.113 (0.046)**	-0.011 (0.011)
Water					
stream (388)	-0.226 (0.074)***	-0.219 (0.073)***	-0.001 (0.013)	-0.108 (0.063)*	0.005 (0.013)
river (249)	0.269 (0.083)***	0.240 (0.082)***	0.048 (0.015)***	0.129 (0.071)*	0.041 (0.013)***
floodplain (92)	0.106 (0.114)	0.104 (0.104)	0.034 (0.020)*	0.028 (0.094)	0.028 (0.019)
ford (39)	0.684 (0.227)***	0.705 (0.227)***	0.096 (0.030)***	0.393 (0.196)**	0.079 (0.025)***
source (37)	-0.317 (0.195)	-0.278 (0.188)	-0.022 (0.024)	-0.172 (0.152)	-0.017 (0.024)
lake (29)	-0.362 (0.202)*	-0.367 (0.184)**	0.002 (0.041)	-0.324 (0.156)**	-0.003 (0.040)
water (20)	0.172 (0.237)	0.122 (0.229)	0.058 (0.037)	0.103 (0.191)	0.058 (0.037)
riv. mouth (17)	0.825 (0.276)***	0.784 (0.254)***	0.050 (0.037)	0.500 (0.235)**	0.039 (0.036)
riv. island (15)	0.041 (0.297)	-0.200 (0.244)	0.070 (0.019)***	-0.445 (0.189)**	0.038 (0.017)**
Buildings					
castle (281)	0.294 (0.084)***	0.238 (0.084)***	0.049 (0.017)***	-0.087 (0.077)	0.020 (0.012)*
church (77)	-0.131 (0.126)	-0.019 (0.123)	0.075 (0.018)***	0.038 (0.105)	0.079 (0.018)***
farm estate (67)	-0.555 (0.113)***	-0.490 (0.116)***	-0.001 (0.023)	-0.348 (0.109)***	0.002 (0.023)
monastery (45)	-0.015 (0.187)	0.002 (0.184)	0.043 (0.017)***	-0.044 (0.162)	0.038 (0.018)**
bridge (25)	0.297 (0.258)	0.355 (0.279)	0.057 (0.032)*	0.065 (0.233)	0.035 (0.029)
fence (24)	0.195 (0.184)	0.282 (0.192)	0.111 (0.037)***	0.216 (0.177)	0.108 (0.037)***
mill (20)	0.486 (0.329)	0.502 (0.334)	-0.007 (0.047)	0.326 (0.290)	-0.012 (0.046)
Elevation					
mountain (174)	-0.052 (0.100)	-0.076 (0.096)	0.028 (0.016)*	-0.144 (0.083)*	0.018 (0.015)
elevation (35)	0.038 (0.221)	0.113 (0.206)	0.035 (0.027)	0.018 (0.170)	0.027 (0.027)
stone/cliff (34)	0.068 (0.241)	0.033 (0.225)	0.015 (0.025)	0.034 (0.188)	0.016 (0.023)
hill (26)	-0.038 (0.189)	0.001 (0.189)	0.058 (0.026)**	0.049 (0.155)	0.062 (0.027)***
valley (22)	0.184 (0.231)	0.076 (0.243)	0.019 (0.030)	0.172 (0.250)	0.030 (0.033)
Plants					
tree (148)	-0.119 (0.105)	-0.128 (0.100)	0.025 (0.015)*	-0.098 (0.087)	0.024 (0.015)*
forest (137)	0.137 (0.115)	0.140 (0.103)	-0.023 (0.052)	0.190 (0.090)**	-0.014 (0.049)
hedge (77)	-0.553 (0.156)***	-0.505 (0.147)***	-0.071 (0.024)***	-0.441 (0.127)***	-0.075 (0.024)***
clearing (82)	0.195 (0.135)	0.173 (0.120)	0.071 (0.023)***	0.166 (0.104)	0.073 (0.023)***
bush (17)	0.064 (0.223)	0.090 (0.218)	0.026 (0.042)	0.034 (0.181)	0.021 (0.036)
Land					
field (108)	-0.129 (0.120)	-0.110 (0.120)	-0.009 (0.017)	-0.092 (0.097)	-0.010 (0.017)
swamp (45)	-0.123 (0.237)	-0.186 (0.220)	-0.009 (0.021)	-0.087 (0.186)	-0.003 (0.021)
bog (29)	-0.313 (0.193)	0.119 (0.196)	-0.008 (0.052)	0.141 (0.183)	-0.003 (0.054)
land (28)	0.044 (0.225)	0.155 (0.205)	0.045 (0.032)	0.063 (0.141)	0.038 (0.030)
Resources					
salt (24)	0.275 (0.187)	0.339 (0.186)*	0.033 (0.028)	0.144 (0.190)	0.021 (0.026)
reed (18)	0.190 (0.233)	0.025 (0.215)	0.071 (0.020)***	-0.068 (0.202)	0.061 (0.022)***
Controls					
first record/100	-0.117 (0.010)***	-0.132 (0.011)***	-0.013 (0.005)**	-0.072 (0.010)***	-0.009 (0.004)**
city right dum.				1.286 (0.038)***	0.132 (0.036)***
ln urbpot (1910)			1.023 (0.023)***		0.998 (0.029)***
country FE	yes	yes	yes	yes	yes
lat, long, sq. & int.	no	yes	yes	yes	yes
Observations	3738	3738	3738	3738	3738
R ²	0.1884	0.2435	0.9607	0.4314	0.9622

Note: Dependent variable is log population in 1910. *, **, *** denotes significance at the 10%, 5%, 1%. Robust standard errors are given in parenthesis. The bicentennial dummy takes the value 1 if the place was first recorded in the period, starting from the year 1. The lat & long, sq. & int. control of the latitude, longitude, their squared terms and interaction terms with a total of 8, equivalent to the polynomial $(latitude^2 + latitude + 1)(longitude^2 + longitude + 1) - 1$. The city right dum. variable is a dummy that takes the value 1 if the city received its city right before the year 1911.

nutrients provided by floodplains and other land characteristics. However, most places in our dataset were founded during the Middle Age and include many smaller places that still exists today. Furthermore, due to the relative lack of coasts, trade within the Holy Roman Empire likely took place relatively more through land routes as compared to the Ancient World. This is evident by bridges being on average faster in receiving its city rights by around 44%.

However, as compared to the impact of bridges on the population of 1910 through the estimation of equation (6) there seem to be a positive, but no statistically significant impact, suggesting that these bridge places lost its importance over time. On the other hand, we find the opposite pattern for rivers and fords on the population of 1910. This could indicate, at least in our sample, a change from land-based trade towards river-based trade throughout time. We also find that swamps were negative characteristics early on, but we found no lasting impact of swamps on today's population. This could simply reflect that large scale draining of swamps only happened later on, for instance to create farm land which especially happened in the northern part of Germany.

We see that characteristics that can be associated with security were important early on, such as mountains and castles. We may worry that there is significant positive reverse causality between man-made characteristics and local economic activity i.e. due to endogenous, man-made amenities as in Brueckner et al. (1999). However, the man-made building variable coefficients are not negative and significant throughout, which we would expect if simultaneity bias is strong. We take this as a reassurance that second-nature geography is not as relevant compared to first-nature geography as a reasonable assumption, at least until cities received their city rights.

Table 2.8 reports a few robustness checks. In VIII, adding latitude and longitude and their squared and interaction terms slightly attenuate the estimates, except for bridges. In IX, some characteristics may be settled earlier relatively more often (c.f. Figure 2.2a), and as the wave of the first popular city rights started around 1160, we may find that these older settlements before 1160 may disproportionately take longer to receive its city rights. Hence, not controlling for this may bias the estimates for these places upwards, so we add a bicentennial dummy starting from year 1 on the first recorded mention to control for this. Again, the estimates are quite robust to this change. Lastly, we have so far only considered cities that we observe to have city rights today, which ignores the places that do not have city rights and could be considered a city today. Although rare, some large cities such as Offenbach do not possess city rights, as there are no legal differences anymore today. In column X, and in order increase the number of comparison units between characteristics, we additionally consider the whole sample

by assuming that those 1719 places without a city right would have received it in 2012, which is the year the book (Niemeyer, 2012) was published. Recall here that 'soft' threshold of the book were places with above around 7500 inhabitants. By increasing the comparisons, we find that the river estimate is highly significant and negative now, indicating that river places took 17% less years to receive its city rights.

Table 2.7: Impact of Geography on Early Geographical Advantage: Baseline

Type	I	II	III	IV	V	VI	VII
no geo. meaning	0.391 (0.057)***	0.270 (0.056)***	0.310 (0.053)***	0.330 (0.053)***	0.342 (0.054)***	0.331 (0.053)***	0.320 (0.066)***
Water							
stream	0.400 (0.079)***						0.347 (0.086)***
river	0.052 (0.089)						0.043 (0.091)
floodplain	-0.174 (0.188)						-0.174 (0.184)
ford	0.145 (0.207)						0.087 (0.208)
source	0.548 (0.199)***						0.481 (0.193)**
lake	-0.424 (0.310)						-0.405 (0.291)
water	0.284 (0.358)						0.226 (0.367)
riv. mouth	-0.110 (0.356)						-0.146 (0.361)
riv. isl.	0.133 (0.309)						0.073 (0.310)
Buildings							
castle		-0.490 (0.093)***					-0.429 (0.097)***
church		0.295 (0.117)**					0.356 (0.123)***
farm est.		0.477 (0.149)***					0.554 (0.155)***
monastery		-0.067 (0.242)					0.008 (0.235)
bridge		-0.492 (0.255)*					-0.444 (0.259)*
fence		0.369 (0.172)**					0.458 (0.178)***
mill		0.003 (0.295)					0.004 (0.311)
Elevation							
mountain			-0.339 (0.142)**				-0.320 (0.145)**
elevation			0.332 (0.201)*				0.394 (0.191)**
stone/cliff			-0.213 (0.288)				-0.249 (0.301)
hill			0.043 (0.294)				0.104 (0.310)
valley			0.141 (0.361)				0.154 (0.367)
Plants							
tree				0.084 (0.116)			0.058 (0.116)
forest				0.037 (0.146)			0.060 (0.149)
hedge				-0.400 (0.302)			-0.399 (0.305)
clearing				0.072 (0.193)			0.068 (0.194)
bush				0.078 (0.595)			0.034 (0.602)
Land							
field					0.283 (0.119)**		0.235 (0.126)*
swamp					0.600 (0.161)***		0.572 (0.163)***
bog					-0.149 (0.377)		-0.068 (0.444)
land					-0.258 (0.298)		-0.247 (0.286)
Resources							
salt						0.168 (0.230)	0.140 (0.224)
reed						-0.192 (0.262)	-0.183 (0.275)
Controls							
country FE	yes	yes	yes	yes	yes	yes	yes
Obs.	2019	2019	2019	2019	2019	2019	2019
R ²	0.1108	0.1190	0.1027	0.1111	0.1030	0.0984	0.1403

Note: Dependent variable is log of the difference of city right and first record year for places that received city rights. *, **, *** denotes significance at the 10%, 5%, 1%.

2.5.3 Early and Late Cities

From Figure 2.3b, we see that most city rights were awarded in the first wave before the 16th century, with a second wave of city rights being awarded from the 19th century onwards. Very stylistically, we can sum up these periods as the extensive margin of population being the driver in the first wave, for example with the Frankish Colonisation and later the Ostsiedlung (Eastern Settlement) as two concrete examples. On the other hand, the intensive margin would be the driver during the second wave, as there were very few new settlements founded during that period, but another wave of city rights was awarded from the start of industrialization and increasing urbanisation.

Similar to Henderson et al. (2017), we can think of early, first wave and late second wave cities in this case, and examine their differential effect depending on the characteristic. Older cities are likely to be larger today, for example as shown by the negative first record control coefficient of table 2.6, so that we are more likely to find a positive differential effect of early cities. However, these may depend on the characteristics itself, with differences here telling us how different characteristics make use of population, and thus second-nature effects over time.

In Table 2.9, we interact the characteristics with a dummy of early cities, defined by city rights being awarded before the year 1801. We consider 1800 specifically as being right before the industrialization and urbanisation wave, as well as the fall of the Holy Roman Empire. In table 9, column (I) and (III) report the main effect, with column (II) and (IV) reporting the interaction effect. Adding bicentennial year dummies and the latitude, longitude, their squared terms and interaction in column (III-IV), does not change the estimated coefficients much. As expected, the interaction effect of being an early city almost always outweighs the main effect, suggesting agglomeration. For example, the difference between main and interaction effect for streams, rivers and settlements without geographic meaning is about 0.5, and about twice as large for fords and bridges. This suggests that early cities along fords and bridges, two characteristics also associated with land-based trade, tend to be larger than their late city counterpart of water-access without the immediate early ability to cross rivers. That is, land-based trade characteristics that are bigger early on are relative to their smaller counterparts, relatively more relevant later on. This seems, especially compared to the 'no meaning' baseline, less to be the case for those associated with security such as river islands, hedges, castles, or mountains.²⁹

²⁹One guess is that these land-based trade characteristics are relatively rare, as compared to the relatively frequent stream and river places. Land-based traders and merchants may agglomerate more in these places, leading to overproportional second-nature effects due to population. Another way to put it is that these places may need more population to thrive, while places associated with security may not benefit as much from population.

Table 2.8: Impact of Geography on Early Geographical Advantage: Robustness

Type	VIII	IX	X
Dependent Variable	$\log(gap_i)$	$\log(gap_i)$	$\log(gap_i \text{ or } 2012 - firstrecord_i)$
no geo. meaning	0.285 (0.062)***	0.129 (0.058)**	0.173 (0.039)***
Water			
stream	0.206 (0.082)**	0.194 (0.079)**	0.161 (0.051)***
river	0.022 (0.082)	-0.104 (0.074)	-0.170 (0.065)***
floodplain	-0.122 (0.181)	-0.095 (0.161)	-0.074 (0.118)
ford	0.020 (0.209)	-0.143 (0.189)	-0.258 (0.162)
source	0.398 (0.194)**	0.356 (0.215)*	0.310 (0.110)***
lake	-0.118 (0.246)	-0.345 (0.203)*	-0.067 (0.178)
water	0.128 (0.247)	-0.168 (0.234)	-0.154 (0.189)
riv. mouth	-0.099 (0.311)	-0.207 (0.288)	-0.519 (0.287)*
riv. isl.	0.183 (0.205)	0.365 (0.203)*	0.294 (0.213)
Buildings			
castle	-0.376 (0.092)***	-0.377 (0.082)***	-0.525 (0.077)***
church	0.169 (0.130)	0.242 (0.133)*	0.198 (0.077)***
farm est.	0.440 (0.146)***	0.388 (0.126)***	0.216 (0.069)***
monastery	0.004 (0.220)	0.013 (0.213)	-0.083 (0.156)
bridge	-0.566 (0.257)**	-0.458 (0.243)*	-0.456 (0.222)**
fence	0.262 (0.190)	0.258 (0.187)	-0.015 (0.133)
mill	-0.067 (0.261)	-0.202 (0.251)	-0.274 (0.214)
Elevation			
mountain	-0.305 (0.142)**	-0.243 (0.134)*	-0.161 (0.094)*
elevation	0.247 (0.196)	0.042 (0.162)	0.025 (0.111)
stone/cliff	-0.280 (0.273)	-0.298 (0.269)	-0.104 (0.192)
hill	0.137 (0.307)	0.210 (0.313)	0.150 (0.171)
valley	0.207 (0.341)	0.245 (0.355)	0.123 (0.219)
Plants			
tree	0.090 (0.114)	0.103 (0.110)	-0.007 (0.076)
forest	0.098 (0.124)	0.196 (0.120)	0.083 (0.082)
hedge	-0.441 (0.283)	-0.241 (0.281)	-0.059 (0.158)
clearing	0.082 (0.181)	0.170 (0.177)	0.077 (0.117)
bush	0.107 (0.512)	0.038 (0.467)	-0.061 (0.307)
Land			
field	0.161 (0.128)	0.070 (0.118)	0.050 (0.081)
swamp	0.502 (0.149)***	0.343 (0.167)**	0.278 (0.102)***
bog	-0.259 (0.515)	0.616 (0.397)	0.172 (0.116)
land	-0.345 (0.259)	-0.195 (0.253)	-0.304 (0.213)
Resources			
salt	0.038 (0.208)	0.013 (0.196)	-0.210 (0.182)
reed	0.023 (0.290)	-0.181 (0.283)	-0.006 (0.205)
Controls			
country FE	yes	yes	yes
bicentennial dummy	no	yes	yes
lat, long, sq. & int.	yes	yes	yes
Obs.	2019	2019	3738
R ²	0.2572	0.3650	0.2904

Note: Dependent variable is log of the difference of city right and first record year for places that received city rights. *, **, *** denotes significance at the 10%, 5%, 1%. The bicentennial dummy takes the value 1 if the place was first recorded in the period, starting from the year 1. The lat, long, sq. & int. control of the latitude, longitude, their squared terms and interaction terms with a total of 8, equivalent to the polynomial $(latitude^2 + latitude + 1)(longitude^2 + longitude + 1) - 1$.

Table 2.9: Interaction between Characteristics and Early Cities

column	I	II	III	IV
Type	main effect	interaction	main effect	interaction
no geo	-0.558 (0.056)***	1.122 (0.064)***	-0.518 (0.054)***	0.976 (0.060)***
meaning				
Water				
stream	-0.548 (0.073)***	1.121 (0.132)***	-0.513 (0.071)***	0.976 (0.128)***
river	-0.128 (0.104)	0.721 (0.139)***	-0.178 (0.103)*	0.670 (0.137)***
floodplain	-0.226 (0.138)	0.330 (0.195)*	-0.165 (0.129)	0.346 (0.171)**
ford	-0.043 (0.295)	1.252 (0.423)***	0.024 (0.284)	1.140 (0.409)***
source	-0.818 (0.181)***	1.596 (0.398)***	-0.705 (0.176)***	1.436 (0.398)***
lake	-0.776 (0.216)***	0.850 (0.331)***	-0.654 (0.208)***	0.486 (0.340)
water	-0.140 (0.256)	0.862 (0.403)**	-0.174 (0.268)	0.568 (0.450)
river mouth	-0.013 (0.217)	1.318 (0.414)***	0.202 (0.255)	0.930 (0.399)**
river island	-1.220 (0.354)***	1.515 (0.430)***	-1.273 (0.461)***	1.443 (0.517)***
Buildings				
castle	-0.353 (0.114)***	0.809 (0.143)***	-0.301 (0.111)***	0.689 (0.141)***
church	-0.226 (0.144)	0.351 (0.203)*	-0.137 (0.145)	0.296 (0.197)
farm estate	-0.676 (0.133)***	0.271 (0.223)	-0.532 (0.130)***	0.117 (0.234)
monastery	-0.332 (0.157)**	0.866 (0.433)**	-0.256 (0.154)*	0.683 (0.446)
bridge	-0.789 (0.234)***	1.597 (0.362)***	-0.618 (0.240)***	1.488 (0.377)***
fence	0.026 (0.237)	0.342 (0.295)	0.114 (0.239)	0.135 (0.288)
mill	0.591 (0.519)	-0.261 (0.643)	0.517 (0.499)	-0.122 (0.629)
Elevation				
mountain	-0.640 (0.118)***	1.046 (0.176)***	-0.549 (0.117)***	0.927 (0.173)***
elevation	-0.096 (0.259)	0.648 (0.402)	-0.030 (0.248)	0.534 (0.403)
stone/cliff	-0.269 (0.186)	0.833 (0.393)**	-0.203 (0.204)	0.699 (0.380)*
hill	-0.097 (0.179)	0.110 (0.379)	0.033 (0.205)	0.063 (0.344)
valley	0.226 (0.216)	-0.168 (0.503)	0.298 (0.202)	-0.657 (0.563)
Plants				
tree	-0.305 (0.124)**	0.557 (0.207)***	-0.289 (0.123)**	0.523 (0.195)***
forest	-0.201 (0.120)*	0.925 (0.236)***	-0.100 (0.111)	0.782 (0.216)***
hedge	-1.091 (0.171)***	1.439 (0.290)***	-0.853 (0.159)***	1.188 (0.276)***
clearing	-0.254 (0.148)*	0.983 (0.239)***	-0.134 (0.139)	0.729 (0.220)***
bush	0.416 (0.269)	-0.881 (0.376)**	0.299 (0.273)	-0.654 (0.359)*
Land				
field	-0.510 (0.123)***	1.069 (0.229)***	-0.465 (0.120)***	0.965 (0.223)***
swamp	-0.635 (0.158)***	1.241 (0.386)***	-0.624 (0.156)***	0.986 (0.389)**
bog	-0.788 (0.212)***	1.216 (0.814)	-0.336 (0.226)	0.831 (0.873)
land	-0.558 (0.201)***	1.320 (0.317)***	-0.351 (0.203)*	1.130 (0.311)***
Resources				
salt	0.331 (0.319)	-0.042 (0.374)	0.321 (0.350)	0.013 (0.403)
reed	0.094 (0.318)	0.126 (0.508)	-0.010 (0.313)	0.104 (0.435)
Controls				
country FE	yes		yes	

bicentennial dummy	no	yes
lat, long, sq. & int.	no	yes
Observations	3738	3738
R ²	0.2999	0.3592

Note: Dependent variable is log population in 1910. *, **, *** denotes significance at the 10%, 5%, 1%. Robust standard errors are given in parenthesis. The interaction term is the city rights dummy with 1 if city rights were received before 1800, and 0 if city rights were received later than 1800, or if no city rights were awarded. The bicentennial dummy takes the value 1 if the place was first recorded in the period, starting from the year 1. The lat, long, sq. & int. control of the latitude, longitude, their squared terms and interaction terms with a total of 8, equivalent to the polynomial $(latitude^2 + latitude + 1)(longitude^2 + longitude + 1) - 1$.

2.6 Conclusion

We started by considering place names, so-called toponyms, as a novel source for early settlement geographical characteristics. Toponyms have the advantage that we can glimpse at both the local and historical; in other words, the *original* geographical characteristics of a place. For this, we have constructed a unique dataset that contains 3738 mostly German places, where we have individually categorized their toponyms to 242 different geographical characteristic. We have also digitized the first recorded mention of a place, as well as the city rights year.

This unique dataset allowed us to show that characteristics, as suggested by its toponyms as well as a proxy of early geographical advantage, defined by the year gap of city rights received and first recorded mention of the place, are both exhibiting a Zipf distribution. These results together support the strand of the literature that attempts to explain Zipf's law for cities through variations in early geographical advantages, which itself are considered Zipf distributed (Krugman, 1996; Gabaix & Ioannides, 2004).

Next, we have estimated the impact of individual characteristics on local economic activity. We showed that places with characteristics associated with security and land-based trade, such as mountains, castles and bridges, received their city rights relatively early. However, we also estimate a positive impact of rivers, river crossings and fords on the modern population of a place, with fords in particular exhibiting strong path dependence as they are long gone and not relevant anymore. That is, water-based trade seems to have become important only later in history. At least for most of the Holy Roman Empire, land-based trade seemed to be more important than water-based trade initially, with the importance reversing throughout history. That is, contrary to a vast literature that emphasized the role of water-based trade characteristics (e.g. Bosker, Brakman, Garretsen, De Jong,

& Schramm (2008), Henderson et al. (2017), Barjamovic et al. (2019), Bakker et al. (2019)), we do not find a particular importance of those characteristics at least very, very early in a settlement's history.

We see future avenues of research in the estimation of how individual characteristics change in importance over time. This would however require better population estimates throughout time; better in terms of how far the data can go back in time, as well as how many places are considered. One promising upcoming dataset by Chaney (2020) provides a proxy for population estimates, going back further in time, as well as for more cities than the rather restrictive dataset by Bairoch et al. (1988).

2.A Data Appendix

2.A.1 Data Source

As the main source of our data, we make use of an encyclopedia of toponyms, 'Deutsches Ortsnamenbuch' by Niemeyer (2012), which we will in the following refer to as the 'book'. As compared to the 'Städtebuch', which gives a chronical overview of German cities, the book provides the etymological development of a given place from the first known recorded mention of the place. Figure 2.7 shows an excerpt of the book. Each entry of the book is categorized into five parts, (I) a summary of the place, which in general provide current population, the legal status of the place as a municipality or as a city, and a description of the location of the place which is especially useful to identify which place is actually meant in case of a common toponym. Not always is the year of the first city or market rights provided. Here, the legal status of the place provides guidance to whether the place had obtained city rights in the past, and if the year is missing from the entry, that we add it to the dataset. Typically, we found those years on the official website of the place. (II) provides an overview of the names of the place over time, as well as the year of the recorded mention. (III) provides the toponymy, that is, the origin and meaning of the toponym of interest. On occasion, it is not immediately clear what the exact meaning of a toponym is, often due to similar sounding suffixes i.e. -berg (mountain) and -burg (castle), or because it is a rather unique name (such as Dortmund) or variant of a toponym (Dörpen, instead of the common -dorf). As such this category provides more of an interpretation of the most plausible meaning by today's research on the toponym, rather than a definite answer. (IV) provides similar toponyms in this encyclopedia. Lastly, (V) shows the source of this entry, as the book can also be considered as a collection of summaries of research articles in toponymy by various authors.

Figure 2.7: Niemeyer (2012) excerpt

135

zur Ortsnamenkunde und Sprach- und Siedlungsgeschichte des Gebietes zwischen mittlerer Saale und Weißer Elster. Berlin 1984; SNB. KH

Dornstadt I. Gem. und gleichnamige VVG im Alb-Donau-Kreis, 13 185 Ew., etwa 8 km nnw Ulm und ca. 42 km ssö Göppingen in der flachen Mulde der Flächenalb in der Kuppenalb gelegen, Reg.-Bez. Tübingen, BW. 1334 durch Kaiser Ludwig an Graf Berthold von Graisbach verpfändet, 1347 an das Kloster Lorsch, 1465 an Württemberg durch Tausch, 1803 an Bayern, seit 1810 württembergisch. Business Park Exit 62, Martinskirche in Tomerdingen, Kirche St. Stephanus (Bollingen), Altar von Scharenstetten. II. 1225 *Dorneconstat*, 1330 *Dorgenstat*, 1334 *Dorgunstat*, 1643 *Dornstat*. III. Der Name ist als 'Wohnstätte auf dem dornigen Platz' zu deuten und gehört zu mhd. *dornec* 'dornig' und dem Gw. ahd. mhd. *stat* 'Stelle, Ort, Wohnstätte' (†-statt). Die Anlehnung an nhd. *Stadt* ist sekundär, der Ausfall des *g* ist alem. Namegebend waren die alten Dornhecken im Muschelkalkgebiet des Heckengäus, die auf den Lesesteinwällen der Feldraine wachsen. IV. †Dornstetten, Lkr. Freudenstadt, BW. V. Bach DNK 2; LBW 7. JR

Dornstetten I. Stadt im Lkr. Freudenstadt, 8045 Ew., ca 7 km ö Freudenstadt und etwa 22 km nnw Sulz am Neckar im Freudenstädter Graben im Haupttal der Glatt zwischen zwei zur Glatt fließenden Bächen gelegen, Reg.-Bez. Karlsruhe, BW. In fränk. Zeit entstanden, zu Beginn des 12. Jh. Edelfreie von Dornstetten bezeugt, im 12. Jh. an die Grafen von Urach, Stadtgründung Mitte 13. Jh. durch die Grafen von Fürstenberg, seit 1759 Sitz eines württembergischen Amtes, 1807 aufgehoben und an Freudenstadt. Früherer Silberbergbau, Barfußpark, Rundfachwerk, Himmlisch Heer, Geburtsort von Jacob Beurlin. II. 768 (Kop. 12. Jh.) in *Tornigesteter marca*, 771 (Kop. 12. Jh.) in *pago Tornegovve in villa Stedden*, 779/783 (Kop. 12. Jh.) in *Tornigestat*. III. Wie †Dornstadt urspr. zu mhd. *dornec* 'dornig' und dem Gw. ahd. mhd. -*stat* 'Stelle, Ort, Wohnstätte' (†-statt), hier im Dat. Pl. -*stetten*. IV. †Dornstadt, Alb-Donau-Kreis, BW. V. Bach DNK 2; LBW 5. JR

Dörpen I. Gem. und gleichnamige Samtgemeinde im Lkr. Emsland, 15 680 Ew., 30 km n Meppen, NI. II. 890 *Dorpun*, 10. Jh. *Thorpun*, um 1000 *Dorpun*, 1350–1361 in *villa Dorpen*. III. Der Name besteht aus dem Simplex asä. *thorp*, mnd. *dorp* im Dat. Plur (Lok.), †-*dorf*. IV. Dörfen in Bayern und Baden-Württemberg. V. H. Abels; Casemir/Ohainski. MM

Dorsten I. Stadt im Kr. Recklinghausen, 77 975 Ew., an der Lippe, Reg.-Bez. Münster, NRW. Kirchdorf im 12. Jh. auf einem Hofe des Stiftes Xanten, 1251 befestigte Stadt im kftl.-kölnischen Vest Recklinghausen,

Zugehörigkeit zur Hanse, Schiffsbau für die Lippeschifffahrt, 1803 zum Hztm. Arenberg, 1811 zum Ghzmt. Berg, 1813 zu Preußen, 1992 Eröffnung des „Jüdischen Museums Westfalen“, Kohlebergbau bis 2001. II. Um 900 *Durstinon*, 10. Jh. *Durstina*, 1251 *Durstene*, 1293 *Dorsten*. III. Nd. *dorst*, heute eingeeengt auf 'Durst', hat, wie lat. *torrere* 'trocknen', 'dörren' und hd. *dörren* und *dürr* zeigen, eine allgemeinere Bedeutung im semantischen Bereich von 'trocken'. Das Suffix -*ina* wird in Siedlungsnamen gern als 'Gewässer' gedeutet, kann aber auch als Verkleinerungssuffix verstanden oder für die Adjektivbildung gebraucht werden. Motiv für die Namengebung mag eine für die Haltbarmachung bestimmter Früchte unabdingbare Darre (Dörranlage) gewesen sein. Die Wortform mit -*st*- lässt auch den Gedanken an einen „alteuropäischen“ Gewässernamen zu (Krahe). Die ungewöhnliche Beibehaltung des -*o*- an Stelle des wegen des -*i*- in der Folgesilbe zu erwartenden Umlauts -*ö*- (†Datteln, Kr. Recklinghausen, NRW) hat Parallelen in den benachbarten Niederlanden. IV. Dorstfeld, Stadt Dortmund, NRW; †Dorslon, Hochsauerlandkreis, NRW. V. Werdener Urbare A; WfUB VII; Krahe, H.: Über einige Gewässernamen mit *st*-Suffix. In: Beitr. zur Namenforschung 10 (1959). *schü*

Dortmund I. Kreisfreie Stadt, 584 412 Ew., Reg.-Bez. Arnsberg, NRW. Bis 1802 einzige Reichsstadt Westfalens, mit eigener Stadtrechtsfamilie, an der Fernhandelsstraße „Hellweg“ von Köln nach Leipzig und Krakau, Mitglied der Hanse, 1802 zum oranien-nassauischen Ftm. Fulda, 1808 zum Ghzmt. Berg, 1813 preußisch, ab 1892 Bau des Dortmund-Ems-Kanals, 1928 und 1929 u.a. Erweiterung durch Teile der Landkreise Dortmund und Hörde, größte Flächenstadt des Ruhrgebiets, Kohlebergbau bis 1987, Schwerindustrie, Bierproduktion, 1968 Universität. II. Um 890 *Throtmanni*, 947 in *Throtmennia*, zwischen 1033 und 1050 *moneta Thrutminensis*, 1074 *Drutmunne*. III. Nach vielen, zum Teil dilettantischen Deutungsversuchen des Namens hat P. Derks 1987 eine philologisch vertretbare Erklärung bestätigt, die schon 1919 vorgeschlagen worden war: Ae. *throtu* und ahd. *drozza* 'Kehl', 'Gurgel', 'Schlund' (vgl. nhd. *erdrosseln*) belegen die Existenz eines für das And. zufällig nicht belegten **throt*. Das Grundwort -*menni* ist in mehreren Siedlungsnamen als Gewässerwort belegt. Das Schluss-*d* ist, wie die Belege zeigen, erst spät in Anlehnung an *mund* 'Schutz' hinzugetreten. Beim ON Dortmund handelt es sich demnach um eine Bezeichnung für einen in einer kehlenartigen Geländeform (vgl. „Hohlkehle“) fließenden Bach. IV. †Holzminden (< *Holtesminne*), NI; Opmünden (< *Upmenni*), OT von Sassendorf, Kr.

Dortmund

To translate an example page, the entry Dornstadt in Figure 2.7 reads 'I. municipality and Vereinbarte Verwaltungsgemeinschaft in the Alb-Donau-district, 13185 population, 8 km north-north-west of Ulm and ca. 42 km south-south-east of Göppingen [...] II. 1225 Dorneconstat, 1330 Dorgenstat, 1334 Dorgunstat, 1643 Dornstat III. The name is to be interpreted as 'habitation on the thorny place' and belongs to middle high-german dornec 'thorny' and the basic word, old high-german, middle high-german stat 'site, place, habitation'. The basis on lower high-german is secondary, the dropping out of the g is alemmanic. Eponymous were the old thorny hedges of the Muschelkalk area of the Heckengäu, which grow on the clearance cairn of the field border.

We assign this place to the '**hedges**' type on the 'suffix-level' (assign a 1 to the hedges dummy), with the additional description 'thorny hedge' with 'thorny' on the 'prefix-level' (assign a 1 to the prefix-level thorny dummy here, but we do not consider this level for our analysis). Note that thorny is a reference to a hedge in this case, so it does not fit into the traditional suffix/prefix dichotomy.

Similarly, the next entry Dornstetten has the same meaning, and is written as such. Therefore we assign both places to the 'statt/stetten' dummy variable, as well as the 'thorny hedges', which is a subgroup of the 'hedges' dummy. The next entry Dörpen is rather short, which typically indicates that the interpretation is straightforward. Here, 'II. 890 Dorpun [...] III. the name consist of the simplex, old saxon thorp, middle lower-german dorp, in the dative plural case, -dorf'. The next two entries however are not as straightforward. The city Dorsten reads 'II. around 900 Durstinon [...] III. Lower German dorst, today narrowed down to 'thirst', like latin torrere 'drying, parching', and show high german parching and dry, a general meaning in the semantic area of 'drying'. The suffix -ina is, in settlement names, often interpreted as a 'waterbody', can also be understood as a diminutiv suffix or can be used to construct adjectives. Motive for the naming could be, for the preservation of certain fruits, indispensable Darre (drying facility). The word with -st- also leaves the possibility of an 'old european' water body name. The unusual retention of the -o- instead of the -i- in the following syllable to be expected umlaut -ö- has parallels to the neighboring Netherlands'. For this case, we do not assign Dorsten to any type, as it is unclear and no favorite or likely interpretation is provided in the entry. Furthermore, it is mentioned that Dorsten became a fortified city in 1251, which is unclear whether the city got simply fortified in 1251, or received the city rights in 1251 with the typical rights to build walls. We double-checked these particular cases where it was unclear, and they indeed also received city rights in 1251. Unfortunately, because there were more than 80 authors involved in the creation of the book, it also introduces non-standardization of the entries in terms of what kind of information, and how these informations

are being provided. As such, we cannot guarantee that each entry is consistent with each other, so that we do not assign this place any characteristic (every geographical characteristic dummy is assigned a 0). Second, we assign the first recorded mention 'around 900' as simply 900. The same goes for other words of approximation. Also, whenever we encounter a century approximation e.g. 15th century, we simply assign the middle of the century i.e. 1450, unless we find another source which provides a more precise year.

Lastly, Dortmund is 'II. around 890 Throtmanni [...] III. After many, partly amateurish attempt to interpret the name, P. Derks 1987 confirmed a philologically justifiable explanation, which was suggested already in 1919: Old estonian throtu and old high-german drozza 'throat, gullet, maw' (cf lower high-german choke) [...] The basic word -menni is in many settlement names documented as waters. The -d at the end came about, as documented, only later as a reference to mund 'protection'. Therefore, the place name Dortmund is a description of a throat like shape of terrain (cf. hollow throat) flowing stream'. We therefore assign Dortmund to the '**stream**' type suffix-level (assign it to the stream dummy with value 1), with the second assignment at the prefix-level of terrain having the shape of a throat (assign 1 to the unique 'throat' shaped stream prefix dummy, but we do not consider this level for our analysis). Again, it is reference to a stream in this case, so it does not fit into the traditional suffix/prefix dichotomy very often found with stream places that end in -bach or -ach most of the time.

Note that at least in this example page, the interpretation and assignment are straight-forward. However it can happen that the interpretation of the text and assignment is not straight-forward at all, which is well in the nature of translating a text into data as each entry is unique. For instance, it can happen that there are competing interpretations of a name. Often in these cases, the book indicates which one is more likely to be the case, while it presents the less likely ones as also a possibility. Here, we assign the most likely one in our dataset and ignore the competing, less likely interpretations. For those parts or the entire name of a place being unclear, we do not assign anything. For example, the place Dörnhagen (1253) can be translated as 'thin hedge' or 'thorn hedge'. That is, we assign the place to the 'hedges' dummy as it is clear, but we do not assign either 'thin' or 'thorn'. For our purposes however, this would not matter as we simply ignore the type of a hedge in this case, and merely consider 'hedges' for our analysis. As such, the approximately 5% places for which at least a part of the name is unclear is considerably less, especially as the 'type' of some geographic characteristic are typically more varied, and therefore less standardized and harder to interpret correctly.

2.A.2 Changing Names Over Time

It is often the case that the toponym of the first record is not exactly the same as the toponym of today, due to centuries of changes driven simply by changes in language, or mistranscriptions over time that accumulated. For instance, Dortmund went from Throtmanni (890) to Throtmennia (947) to Drutmunne (1074). Those cases are typically not a problem for our purposes, as the interpretation of the meaning did not change over the centuries. On the other hand, mistranscriptions that change the meaning could introduce some ambiguity, which may or may not be resolved. For example, Friedberg (Bavaria) was first mentioned as ciuitatem Fridberch (1264). However, a castle was already built in 1260, with peace in mind, to secure the Bavarian border to Augsburg. It is therefore more likely that the name of the place actually means 'peace castle', and not 'peace mountain'. This brings us to castles being often found on mountains, such that -berg (mountain) and -burg (castle) were often interchangeable, and hence the true meaning of a place may be ambiguous and may often change back and forth over the centuries. In the case of Friedberg (Bavaria), it is clearer which was meant. As for other -berg/-burg places, the book often describes which of the two is actually meant, so that it is not often the case where we have to assign none of the two due to ambiguity. Also, to further reduce ambiguity, we have done our own research by looking at whether there had historically been a castle in the area, or a mountain.³⁰ Only if still remains unclear, then we assign the place to none of the two. We count 56 of those places where it was too ambiguous to assign anything, or the literature has no idea about the meaning, among 4207 entries which we have digitized without reducing it for the analysis. This constitutes around 1.3 percent. Similar to mistranscription is that the people themselves over time misinterpret the original meaning of the place. For example, Neuenhagen bei Berlin was originally nyenhoue (1367) 'new farm', which somehow changed over time to 'new hedge' of today without any conscious decision. Since we only consider the very first recorded mention, we always assign very first mentioned characteristic in our dataset. However, in the case of Hirschaid, which was mentioned first as Hirzheide (1097) and which would then and now be translated as 'deer heath'; it is likely not the first toponym, but rather the first recorded one, as heaths are not the natural habitat of deers. That is, we never really know if the first recorded mention is the first name given to the place.

However, the authors of each entry typically discuss the plausibility of the meaning e.g. if there is actually a mountain around if the place is named after a mountain, so that the name makes sense. As such, even if the first recorded

³⁰For that, the website alleburgen.de, which lists all castles, was particularly useful.

mention is different from the first name given, we should have at least a plausible name most of the time. Indeed, the Hirschaid example is the only case we know of from the sample with a name which does not make sense, and for which an earlier, more plausible meaning is not available.

Note that we do not count the change of Gladbach to Mönchengladbach and that of another Gladbach to Bergisch Gladbach as a change in their name, as these clarification prefixes were added to distinguish the two places. This is often the case when we have places with the same name in close proximity to each other. The other typical case is when two places used to be one, but for some reason drifted apart, split and/or became significant places on their own to warrant the distinguishing prefix. For instance, the typical distinguishing prefixes are about their respective direction to each other (west, east etc), their respective age difference (old, new etc), their respective altitude (high, below etc), or simply a geographic characteristic (church, castle, mountain). Not adding more characteristics over time would make it consistent with other places which did not feel the need to add any distinguishing prefixes.

However the most common cases of changes in names were deliberate. We count 325 entries with a change in name out of 4207, before any reduction for the analysis was made. Out of those, we count 127 individual places which changed its name not due to fusion of places where they needed a new name. For instance, Wasserburg am Inn used to be named Hohenowe 'high floodplain' (1137), but was renamed to Wazzerpurch 'water castle' (1310). However we only consider the first record, and ignore the 'am Inn' river, as we are interested in the very first geography. One reason for this decision is that we want to know the path dependence from the beginning. Another, non-minor reason, is that it is sometimes difficult to say whether it is a deliberate name change or whether it is simply a distinguishing prefix. By only considering the first recorded mention, we should for the most part avoid this problem, which relates to the point above.

2.A.3 First Record and City Right Year

We consider the first recorded mention as the year the place became significant enough to be recorded somewhere. Therefore, it is not exactly an official 'foundation' date, such as the date of incorporation of incorporated places in the US. Hence, with the few exceptions of planned places and cities such as Karlsruhe, the first recorded mention does not correspond to the date of foundation or the date of first settlement.

Typically, the place itself is mentioned in a manuscript or document, with the date given. Occasionally, not the place itself is mentioned, but a person named after

the place, such as 'Wernherus de Menkemere' from Maikammer (1260) or Waltherus de Aken from Aken (1227). The notion that the a place is only mentioned if it became significant enough would not apply in these cases. In any case, the first recorded mention is most of the time not the year the first inhabitants populated the location. However, it is the closests possible estimate we can think of. Figure 2.3a shows the distribution of first recorded mentions for our full sample of 3738 places in the book.

Furthermore, we have very incomplete market rights data, but complete city rights data for our sample. The idea is that a place receiving city rights obtained a higher level of significance, which is correlated with population. Therefore, we can exploit this information as a proxy for population growth.

Historically, the two give a place a different level of autonomy, in which the historical market right granted a place the right to hold its own markets. On top to the rights to hold a market, city rights also grant the rights to build walls, as well as its own court and jurisdiction. Therefore, market rights are typically granted before city rights; unless this additional step is skipped. In this sense however, we can consider the year of granting these right as the year the place became significant enough.

Note here that we also assign places as having city rights if they resemble or are considered as similar to city rights. Traditionally, we think of the Magdeburger Stadtrecht and the Lübecker Stadtrecht as the prototypical city rights. However, we also consider the not so typical places as having de facto city rights if they possess the Freiheit or are named as an oppidum or as a Wigbold at some point in time. A place may receive 'proper' city rights after it was named an oppidum years later, but being named an oppidum signals that the place already resembled a city with city rights, or is at least big enough to be named as such. Nevertheless, we should also note here that an oppidum is not a clear-cut city right in itself, and that a place with 'proper' city rights may not even build walls. Also, some places may never receive or apply for city rights, even if they are considered as a city by modern standards. In terms of population the biggest example is Offenbach, which in 1910 already had a population of 75,583 and is by all means considered a city today, and yet apparently never possessed anything which resemble city rights.

The date of receiving city rights are typically recorded in an official document. Yet, the first document may not have been available anymore, maybe because it got lost in a fire. Then it can happen that city rights needed to be confirmed again, which is also documented. However, there are other ways to approximate the date, for example, by the year walls or courts were first mentioned somewhere; or simply even by walls depicted in their coat of arms or seal. Today, municipalities can also apply to be named as a city or be given city rights, but it is merely

symbolic whether a place is called a city or a municipality. In general, places which became big in recent history applied for these symbolic city rights; and this notably happened after sudden increases in population due to annexations or fusions. That is, we typically see a spike in city rights received especially after 1960 due to a series of reforms (Gebietsreform) which merged places. Another reason for the increase in city rights given is a general growth in population after the advent of industrialization. Figure 2.3b shows this spike, starting around the middle of the 19th century for the full sample of 3738 in the book. Most of the city rights were issued however before the 16th century.

2.A.4 Miscellaneous Data

Furthermore, if we have first mentioned names on the same date, for instance if the place received both a slavic and a old-german name, then we consider the most likely one to be there earlier. If that cannot be determined, then we consider the old-german name to be closer to the majority of old-german names in the sample.

Table 2.10: Description of Aggregated Characteristics

Type	typical toponym	Aggregated which fits the following description
A. Common geographic characteristics		
Water		
stream	-ach/aha, bach/beke/beck, A-	- Small, flowing, natural water body. Often not navigable.
river	not applicable	Large, flowing, natural water body. Likely more navigable than streams. Toponyms often refer to the river and the hydronyms of rivers are typically much older and retained even if the name is from a very different language family. As such, there are no typical, obvious and Germanic river suffixes. The hydronyms themselves however often translate to 'move, flow' or other water body references.
floodplain	-au(e)/ouwe, -oog, -sege, -lage, -grim	- Refers to the land adjacent to a water body that is regularly flooded, thus leaving nutrients behind valuable for agriculture. By far the most common suffix -au/-ouwe is very similar to the slavic suffix -ov/-ow 'belongs to, place of', which makes an etymological evaluation especially necessary for these cases.
ford	-furt/förde, -wedel, -brod	- wat, Shallow part of a river which is considered safe to cross, thus making it a natural bridge.
source	-born/brunn/bronn, font	- Spring or source typically provided by ground water.
lake	-see	Non-flowing, natural water body.
water (misc.)	-apa, -phe, ahwa, -isa	- opa, - We group every toponym which refers to a water body or water.

river island	-werth/werder	Ait (river island) or land between two rivers e.g. Werder Bremen.
river mouth confluence	-mund/gmund	A place where two rivers cross, or where a river flows into a different water body
Buildings		
castle	-burg/berg, -stein, -hain, -eck/egg, Kastell, -grad/grod, -dunum, -briga/brica, -ruhe	Due to its similarity of castles (Burg) and mountain (Berg), etymological evaluation is, similar to the flood-plains case, very necessary here. For castles and mountains, this similarity is found for other toponyms as well, such as -stein and -eck, due to the proclivity of castles being built on elevated positions.
church	-kirch/kirchen	Church, and not a chapel which are a category of its own.
farm estate	-hof/hofen	Farm, smallholding, estate. Very typically found around Bavaria. It however can have many different meanings, such that a clear categorization towards the no geographic meaning toponyms, such as -dorf (village) or -heim (home) is not straightforward. As such, we do not group it further.
monastery	-zelle/zell, -münster	Can actually have many functions, as monasteries sometimes provide further services to the community, such as education (monastic schools) or the copy of texts.
fence	-reke/rike, -hegel, -wik/wiek, -tun/zaun, -glinde, -hürde, -esch, -ham, -brüel	Typically described by the German word Einfriedung, which can be translated as fence or enclosure. As compared to the hedge, which may be a natural fence, a Einfriedung may be man-made.
bridge	-brück/bruck/brugg, -specke	Here, the suffix -specke describes a log bridge in particular, at least in this case. Often, a bridge name is accompanied by the name of the water body e.g. Osnabrück (Ausana river), Saarbrücken (Saar river) Innsbruck (Inn river), is named after a person e.g. Erndtebrück (bridge of Irmingard), Copenbrügge (bridge of Cobbo) or describes the type of bridge e.g. Delbrück (plank bridge)
mill	Mühl/mul, Quer-	Interestingly, both Mul and Quer- are often prefixes of a different toponym.
Elevation		
mountain	-berg, mons, gora, -sporn, -spitz, -eck	-berg is the suffix typically associated with mountains, while -sporn and -spitz describe the peak of a mountain. -eck here are typically mountain chains. As for -berg and -eck, they are also used as castle names.
elevation	not applicable	We consider all references to elevated places here.
stone/cliff	-stein, -fels, -klif, -brun	Place on a rock. Similar to -berg, -stein can sometimes refer to castles instead.
hill	-hügel, -bühel, -hübel	

valley	-tal/-thal	Note here that valley is a popular name after fusions, as they are often fitting in describing many originally separate places together. For example, Wuppertal did not exist as a name until 1930 after the fusion Elberfeld, Barmen and Vohwinkel. As a good rule of thumb, the old spelling of -thal indicates that the name is likely not due to fusions as compared to the modern spelling -tal.
--------	------------	---

Plants

forest	-wald, -lohe, -stock, -bere/-bire, -holz, -forst, -hart, -wedel	Forest or woods, any type of forest is found here. -hart is a special case as it describes a light, mountain forest but does not have the typical prefix type to the suffix.
tree	Buch-, Eich-, Birk-, Erl-, Lind-	Here, often only type of the tree is mentioned, such as Buch- (oak) or Eich- (oak), so that a typical suffix such as -baum (tree) does not exist.
hedge	-hagen, -hag, -hain, -heck, -wurt, -wört, -dorn, -bram	Hedge, or other natural fences. Often part of a Landwehr, which is a system of hedges either natural, or artificial or both for protection or for border control, also to collect tolls. -wurt should not be mistaken with Wurt, which are artificial mounds created to protect against floodings. Again, the book typically clarifies which case is actually meant here.
clearing	-rode/reuth/reut/ried/rathscheid, -schneise, -bracht, -brecht, -berth/ breth, -grün, -hau/gehäu, -walde	We consider all references to clearings of a forest. -scheide and -schneise mean borders or divide made through clearings, which are however not common as compared to the -rode places.
bush	-strauch, -busch	

Land

field	-feld/felde, -magen, -wang, -wiese	Considers a field, often used for agriculture.
swamp	-bruch/brook/broich, -bel/birl/berl, -mar/meri	Bruchland are typical swamp places and not void of tree as compared to marshes. As compared to bogs, peat does not develop here.
bog	-fehn/fen, -torf, -moor	-fehn/fen refers to the bog itself, while -torf refers to the peat typically found around bogs.
land	-land/lant, -grün, -kamp, -gau, -reich/rich	

Resources

salt	Salz-, -hall	-hall is traditionally translated as salt, which in some cases is controversial e.g. Halle an der Saale, which also could be translated as 'slope'.
reed	kulba, schilf, -ried	We consider reed places, and reference to it, as a resource here as it was widely used as building material due to its good properties and its use as a pen. Places with reed also act as a natural way to treat water.

Misc

No	Mean-	acum/ako, -um, -	We consider all places here that have no geographic
ing		orum, casa, (i)anum, treb, trabena, büll, beuren/beuern/büren/buror borstel/bostel, he- lid/hütte, büttel, fleth, leben, lev, lage, plan, heim, stat(t), stadt/stedt/stett(en), sal/saal, dorf/torf, weil/wil/weiler/wilen/weier/ville/wilari, wik/wiek/wick, wies/wis, villa, garden/gadum, both, ballig, haus/hus, ingi, ing/ingen/ungen hausen/husen, sel- ida/salida, issa, rege, teil, in, icz/itz/ici/ica/witz, ov/ow, ovik/owice/ovici j, jb, no, yna/ynia, ec, sk, k,inja, unja, inni, innia, jane, bn, y, je, any	meaning. Typically, they can be translated as belong- ing, place, inheritance, house/home, village, hamlet, with the people of a person. Most of the time, the name of the person is male.

B. Uncommon geographic characteristics not considered

bath	Bad, Aquae	References to baths. We do not consider the, typi- cally in the 19th and 20th century, added 'Bad' in the dataset, as they were added to indicate a bath, hot or cold spring, today. For example, the official name of Aachen is Bad Aachen. Aachen itself derives from the roman word for water, aquae, which refers to the bath, so that Bad Aachen would be mean 'bath bath'.
anabranh	twist	Anabranh, a twist of a river. The only entry is Weil- erswist, first mentioned as Swist in 1180.
bay	ham, wick/wik/wick, - förde/fjörd	We consider references to bays. Förde/Fjörd are spe- cial cases of bays that evolved due to glacier move- ments,
marsh	marsch	We consider references to marshes, that differ from swamps with a lack of trees, so that we do not consider marshes as swamps.
inner coast waters	mar/haff	The only entry is Mohrunen/Morag.
pond	-teich, -tümpel, -pfühl	-teich are typically man-made, while the latter are typ- ically natural.

strait	-sund, -belt	The only entry in the dataset is Stralsund (edge form area strait)
canal	-fleet	Mostly references to canals. Fleet are navigable canals regulated by tides.
port	-hafen/haven	
hithe	-hude	No direct translation, but equivalent to hithe places in England. -hude places are landing places for smaller boats on a water body, and can be considered as smaller ports. They also act as storage of wood.
boat	-floß, -plau, -kahn, -fähre	They each refer to different types of boats.
groyne	Buhne	Dam in a water body, for instance to regulate the flow, or to prevent erosion
weir	-wehr	
sluice	Schleuse, -siel, Arke	
dam	-damm	
dune	-lund	Only entry is Schafflund
portage	-walk	Only entry is Pasewalk
roadstead	Reede	Only entry is Rhede
aqueduct	-duct	Only entry is Vaduz
Elevation		
drainage divide	-scheide	Border of neighboring drainage basins
ridge	not applicable	All references to ridges
cape	-horn	A piece of land which protudes into the water
foothills	-ness, schachen	Transition of plains to elevated positions i.e. the border to mountains or hills
mudflow	Mure, Murgang	Only entry is Rum
elevation	Hoch, höhe-, Man-, barm	All references to elevated positions, or places that are located on elevated positions.
Land		
pasture	-weide, -las, -veluwe/velwa, -sweiga	All references to pastures, that is land designated for grazing of animals.
steep terrain	-bille, -wang, -hat/har/hitta, -schleide, -bra	
acre	-arten, ard/eard/acker	acre, or plowed land and references to it
cover	-hut	For example Landeshut and Landshut, which both mean land cover
slope/pile	-halde/helde, -haufen	Slope, which refers to a pile or heap.
edge/wedge area	-horn, -keil, -ecke, -winkel	reference to the form of the area itself
ear form	-ohr	Only entry is Oer-Erkenschwick
ravine	-tobel, -schlucht	
trough	-mulde	Similar to valley
remote, enclosed	-winkel	Reference to remote and enclosed areas around hills, forest etc

fenced	-krug	Only entry is Aukrug, mentioned in 1875 i.e. modern name
meadow		
reese form	rispe	Only entry is Rees
cut in area		Only entry is Riesa
flat surface	-telva	Only entry is Telfs
Building		
chapel	Kappel, cappeln	All references to chapels, which are smaller churches.
braid		Reference to wicker works
silo	-kast	
wall	-macher	Reference to walls, palisades, bullwarks
tower	-turm	
barn	-schuppen, -scheune	
stable		Only entry is Stuttgart
hall	-saal	
dovecote		Only entry is Colmar
court	-dingel, -gericht	
pole	-pfahl, pflock	Could simply refer to walls out of wood
door		Only entry is Düren
knights		Only entry is Dürrenberg, which is a reference to the manor at the top of the thin mountain
manor		
mayor	mayor	Only entry is Meran
estate		
outlook/vantage	gewecht/wacht	Is not necessarily a tower.
point		
care/hospital	-spiez	
mining		Only entry is Grund (Harz)
iron works	-hütte	All references to iron works, although -hütte is the most common.
glass works	-hütte	All references to glass works, which can also make use of the -hütte suffix.
saw mill		Only entry is Schneidemühl
smithery	-schmitt/schmiede	
toll	-zell	Only entry is Gröbenzell (which refers to a house at the Gröben stream)
hunting		All references to hunting
forestry		All references to forestry. We consider it man-made as it is inherently a human activity
road/path	-straße, -stig, -pfad	All references to pathways, streets, roads, boardwalks etc
protection	Schild, schützen	Reference to cover, or protection
shop	-kauf	
winery	Wein, Winzer	
market	-markt	
warehouse		
hub		Only entry is Pforzheim
post station		Only entry is Neumarkt (Egna), which changed names.
tavern	taverna	We consider all references to inns, or taverns
gate	Tor	Only entry is Solothurn

threshing		Only entry is Tennstedt
floor/place		
brick roof		Only entry is Tholey
house		
Plants		
'lar'	lar	Hard to explain. It typically refers to anything natural, so it is a nature reference
swamp forest		
below bork	bast	
tree stump	-stock, -stumpf	May refer to live stock.
twig		All references to twigs.
moss	Moos-	Moos- and reference to moss
bush	-strauch, -bush, Dorn-	All references to bushes, or thorns that refer to bushes.
heath	-heide	Can have different meaning, but most commonly means heath or vacant land.
weathering		
garden	-garten	Could also simply mean 'enclosed'.
geest	Geest/güst/gest	dry, infertile land
wasteland	brache, ledo	
swampy	-luh, -lug	
meadow		
wallow	lache, Lake, liunas, sul, sol, mor	
elevated, morasty	-donk, -dunk	Only entry is Wachtendonk
place		
scrubs in	Lobez	
swamp		
dirt/mud	-hor, -mul	
Resources		
clay	thohe, leim, kley	We consider all references to clay.
gravel	Kies, Gries, Sand	We consider all references to gravel or gritted sand.
copper	Kupfer	Only entry is Kupferzell.
iron	Eisen, Isern	They are typically mentioned with -hütte i.e. refer to iron works, not iron itself.
coal		The only entry is Weglewo, named in 1945. That is, it is a modern name.
gold		The only entry is Goldberg (Złotoryja), known as Aurum in 1201. Since Aurum means gold, it is not a modern name, but uncommon.
pea/bean		The only entry is Erwitte.
grain/wheat/corn	Äckern-, Gren	
hop		The only entry is Hoppegarten, first mentioned in 1797. That is, it is a modern name
onion, leek	Lock/Luck	
cabbage		The only entry is Kölleda

flint stone		The only entry is Flintbek. We consider it a resource and not a type of stream in this case.
chalk		The only entry is Kalkar
juice pro- cessing		The only entry is Kelter, named in 1972 first. That is, it is a modern name.
slate		The only entry is Kirn
pitch		The only entry is Schmölln
flax		The only entry is Flachsmeer, first mentioned in 1818. That is, it is a modern name.
weather		
fog		The only entry is Dissen am Teutoburger Wald
rain		We consider all references to rain.
sun		The only entry is Sonneberg, which was however first mentioned in 1734, which was renamed from 'daz Stetyln zu Roetin' first mentioned in 1317.
warm		The only entry is Tapiau (Gwardeisk)
winter	Winter-	Winterbach and Winterberg are the only entries. How not entirely clear whether it refers to winter, snow etc.
misc.		
name		Relatively often are places named after other places or people, such as Frankfurt (Oder) from Frankfurt (Main). We consider those places as having no clear geographic meaning, even if they do.
transfer		We consider all references to places where people pray.
praying place	Bet-	
cross place		The only entry is Kreuzlingen.
burial place		We consider all references to burial places, such as references to ghosts or grave stones.
place to dry things		One entry, Attendorn, could be mistaken with thorns which is similar. However, it means 'dörren', which is to dry things to conserve them better.
counting place of wood		The only entry is Bobrek-Karf, in this case the Karf part. Karf is first mentioned in 1845, so that this is a modern name.
bald spot	kahl, kalwa, kalewe, gola	We consider all references to bald spots, whether on land or mountain.
border re- gion	-mark/mark-, -lauche, -schnitt	We consider all references to borders, border regions, or divisions.
'ithi'/'ede'	-ithi/ede	A relatively common suffix, where the exact meaning is unknown.
desert(ed), waste	Wuste-	It is closer translated as wasteland, desolated place.
open, unin- habited		The only entry is the relatively big city of Eupen, in now Belgium.
crack	-riss	The only entry is Mömbriss
grind		We consider all references to grinding, which refers to grind stones etc.

Chapter 3

Mean Reversion of the German City System After WWII: To What Mean?

Abstract: This paper uses the synthetic control method to construct comparison units for West-German cities. We use these as counterfactuals to assess the long-run impact following WWII on the economic activity of 53 West-German cities, and the West-German city size distribution. We extend the literature in that we do not only look for whether cities following the WWII bombing shock are mean-reverting or not, but the use of counterfactuals allows us to distinguish whether individual cities experienced a positive or negative impact as well. We overall find mean reversion of the city share for around 50-70% of cities, as well as a roughly balanced ratio of positively to negatively impacted cities.

Keywords: urban growth, synthetic control method, World War II shock, city size distribution

3.1 Introduction

There is considerable interest in the study of the long-run impact of large and temporary shocks on a city's growth, whether they are caused by war or are a cause of nature. Researchers who try to explain economic activity across space exploit the findings of their long-run impact in order to gauge the relevance of some fundamental theories. For instance, the seminal paper by Davis & Weinstein (2002) looked at the WWII bombings on Japanese cities, while the focus of Brakman et al. (2004) and Bosker, Brakman, Garretsen, & Schramm (2008) were the WWII bombings on German cities. However, one issue researchers in the study of these shocks face is a potential change in the trend not caused by the shock, which may simultaneously occur in the aftermath of the shock. For instance, as Glaeser (2005) argues, there is no reason to remain in a city close to a productive natural resource once the demand for this particular natural resource declines. In a similar manner, the development away from a manufacturing to a service economy, the decline of the US rust belt or suburbanization of cities (Glaeser & Kahn, 2004) come to mind as underlying factors which could induce the change in trend. For Germany in particular, the 'Strukturwandel' which describes structural change through the decline of the coal industry in the German Ruhr area around the 1960s, is one example of such a factor, which could have induced a change in the trend. If we do not account for this potential change, then we could otherwise wrongly attribute the decline of those cities to the WWII bombing shock; a decline which may have happened regardless of the shock through the decline of coal for example.

The goal of this paper is to disentangle the effects of the WWII bombing shock on West-German cities from these other, underlying drivers of population growth which the literature so far did not explicitly account for. That is, in the evaluation of WWII bombings, we typically rely on methods which show whether city population is mean-reverting or not to a trend as estimated from the trajectory itself, pre and post-WWII.

However, if the counterfactual is non-linear, then we run the risk of wrongly concluding no reversion to the mean which may be driven by non-linear developments; yet we may attribute it to the shock itself even if the shock did not actually have any long-run impact. Vice versa, we may wrongly conclude mean reversion in the case of the shock actually having a long-run impact, which we do not detect as such because the non-linear counterfactual development counteracts this impact towards a linear trend. In other words, the question should not be whether cities are reverting to the mean, but whether they revert to a potential 'new mean'.

To disentangle WWII bombing shock from other factors, we will employ the synthetic control method (SCM) by Abadie & Gardeazabal (2003) and Abadie et

al. (2010, 2015) to construct the counterfactual/synthetic city for each sampled city, and use this synthetic city as the comparison unit for the bombed city. As compared to other methods which also use counterfactuals, the SCM is designed to ameliorate some issues of finding a suitable counterfactual: First, as the synthetic control method 'constructs' a comparison unit for the bombed city treatment unit by choosing a convex combination out of suitable non-treated comparison cities, we do in principle have infinitely many possible comparison cities.^{1,2} This ameliorates the problem of not finding a suitable comparison unit, which is especially important when the treatment unit are cities at the upper tail of the city size distribution for which suitable single unit counterfactuals are rarer. This is one reason why SCM is preferred over the typical difference-in-difference approach in our case since we can extend the number of potential counterfactual as we consider a convex combination instead of a single unit. Second, we add objectivity to the counterfactual selection as this convex combination is selected from a data-driven procedure, where we minimize a loss function.

As we will show later, the choice now is not about having enough comparison units to choose from, but the choice of matching covariates which will be minimized to obtain the most suitable counterfactual. Typical to the literature however is that the SCM is applied in a single treatment unit without much discussion on the choice of set of matching covariates. Yet the counterfactual and as such the results are sensitive to the inclusion or exclusion of certain covariates, with no apparent criteria as to which specification is the best. Since we apply the SCM with the very same set of covariates to many cities in the German city system, we implicitly add transparency to the whole covariate selection process. In this regard, the SCM method is especially suitable for our case.

Crucial for our purposes is that the synthetic control counterfactual allows for a non-linear city population path. That is because in the study of impacts stemming from large, temporary shocks on the location of economic activity, we need to consider a large time span due to the slow-moving nature of population; otherwise we cannot distinguish mean reversion from unit roots.³ However with a large time span considered, the likelihood of significant non-linearities occurring would be higher, or more pronounced over time. This implies that assuming a linear trend may not be appropriate for relatively long time spans and for these purposes, an explicit estimation of the counterfactual path may be more appropriate.

We will replace the deterministic trend term in the unit root model with the

¹Of course, due to computation reasons, we only consider the finite case.

²That is, instead of considering a single city, or a single point in a convex hull as a combination chosen by the researcher, we now consider the entire convex hull as a possibility, where the first two convex combinations are simply a special case within the convex hull.

³See (Perron, 1991) for the importance of the time span against frequency in unit root tests.

synthetic control term, so that this simple correction is not very invasive in terms of changing fundamental testing principles. Thus, we retain some level of comparability with previous results in the literature (Davis & Weinstein, 2002; Brakman et al., 2004; Bosker, Brakman, Garretsen, & Schramm, 2008) as our results should not be driven solely by the different method itself.

First, our results show an inverse-U shaped population trajectory throughout the 20th century for most of our synthetic cities, which justifies our approach that can deal with these non-linearities. Second, we find reversion to the synthetic mean for the city share of the West-German population for around 72% until 1970, and 64% until 1980. Given these results, adjusting for these non-linearities, we can reconcile the different results of the literature of finding mean-reversion after such a shock (Davis & Weinstein, 2002; Brakman et al., 2004) and of a permanent impact (Bosker, Brakman, Garretsen, & Schramm, 2008), that could simply be driven by the different cutoff years considered, where for later cutoff years, non-linearities are more likely to be significant enough. Cities that were negatively affected were on average smaller and in more remote places, which may indicate WWII accelerating the migration out of smaller cities.

This paper mainly contributes to the strand of empirical research on the lasting effect of a large, temporary shock on a system of cities (Davis & Weinstein, 2002; Brakman et al., 2004; Bosker, Brakman, Garretsen, & Schramm, 2008; Miguel & Roland, 2011), with the aim of providing a stronger empirical method to distinguish the large, temporary shock from any other parallel development happening. The results of this particular strand also has implications on the relevance of some competing fundamental theories on the location of economic activity. (Davis & Weinstein, 2002) defines those theories as locational fundamentals theory (e.g. Rappaport & Sachs (2003), also called 'first-nature geography' in the literature), random growth theory (e.g. Gabaix (1999b) and increasing returns theory (following Krugman (1991))).

This paper is also related to papers on multiple equilibria (Bosker et al. (2007), Davis & Weinstein (2008), Redding et al. (2011), Bleakley & Lin (2012)), although whether the German city system can be characterized by multiple equilibria or not is not the focus here.

To a lesser extent, this paper is related to the disaster literature (duPont IV & Noy, 2015; Siodla, 2015), with the difference that they look at natural disasters typically affecting a single city and not an entire city system.⁴

⁴This distinction could be important in that shocks on the city system may induce relatively little migration given the magnitude of the shock if most cities experienced some form of destruction. This gives room for the conjecture that as long as the large, temporary shock on the city system is even, it does not induce migration whereas a shock on a single city in a city does lead to migration. This notion could explain why we overall find a permanent impact in the disaster

In section 3.2, we will discuss the specific reasons of our approach in the light of the literature as well as possible drivers of non-linearities. Section 3.3 shows the construction of the synthetic control. In section 3.4, we will describe the data. Section 3.5 reports the descriptive statistics and shows that the population trajectory of our synthetic cities is inverse-U shaped. Section 3.6 describes our baseline empirical strategy, while section 3.7 reports the results for city size, city share. Section 3.8 will report some additional characteristics on a group of cities conditional on how the individual city was impacted. Section 3.9 will bring the evidence established in the previous sections together, where we will discuss the aforementioned competing fundamental theories on the location of economic activity in the light of the new evidence. Section 3.10 concludes.

3.2 Non-linear Mean Reversion in the 20th Century

When it comes to measuring whether there is a long-term impact of the WWII bombings on city population or whether the population level will recover after WWII, one worry are simultaneous adjustments that may influence the supposed population level towards which a city has to recover. This would make it difficult to disentangle the effect of the bombings from other developments.

For example, the decline of the coal industry in Europe during the 1960s may have led to an overall decline of cities close to or within coal fields, such as the Ruhr area in Germany or the Midland region in England. One may then assume a lower population level that coal cities are supposed to recover to in order to be considered as recovered from the WWII shock. On the other hand, the post-WWII increase in trade may had benefitted cities with river or sea-access, so one may then assume a higher population level that those cities are supposed to recover to.

A similar argument can be made for cities in recent history regardless of its type, as cities were first in the process of urbanization and then suburbanization with increasing motorization that especially coincided with post-WWII. The overall population decline after the baby boomer generation may be another factor that should be taken into account, which overall contributed to the slow down or even decline of the population of Western European cities.

That is, whether we find that a city is reverting to the mean or whether we find a permanent effect through finding a unit root, for example, also depends on what mean exactly it is supposed to be reverting to. Given these possible adjustments accompanying the post-WWII world for which an inverse-U shaped population growth trajectory is likely, city population may not conveniently be assumed as (linear-trend) stationary, at least throughout the 20th century.

literature concerning a single city, while mean reversion if the disasters affect the whole country.

Then, in the empirical evaluation, these nonlinear adjustments that have nothing to do with the WWII shock itself may lead to an overacceptance of the unit root null in models that considers a linear trend (Perron, 1989). Bosker, Brakman, Garretsen, & Schramm (2008) used a sharp structural break while detrending through linear trends i.e. allowing for a kink, which they found to be around WWII. Initially, they failed to reject the unit root hypothesis for all German cities with the population linear-trend stationary model, but once allowing for a structural break, they reject the unit root hypothesis for about 75 percent of cities.

However, introducing a structural break does not disentangle the impact of the WWII bombing shock from the aforementioned simultaneous nonlinear adjustments as this break would capture both the bombing shock effect as well as that of the adjustment. In one scenario, we may wrongly attribute a decline to the temporary large shock on cities, although such a decline may have taken place in any case without the WWII shock. Furthermore, even if we think that the temporary bombing impact is characterized by a sharp break with a new level and trend being immediate; simultaneous adjustments such as suburbanization and the transition out of coal may be a long process characterized by a smooth break, so that we may still overaccept the unit root null even in the case that a sharp break is detected. This relationship of non-linear mean reverting processes and overacceptance of the false null of a unit root has been shown by Taylor et al. (2001), which could have been the case in Bosker, Brakman, Garretsen, & Schramm (2008) for the linear-trend stationary model that accepted the unit root null for all cities.

In order to address these concerns for the bombed German cities we will attempt to detrend through a non-linear trend. One way to estimate this non-linear trend is through estimating counterfactual cities that were not bombed or were largely undamaged during the war.

We will use the synthetic control method (SCM) (Abadie et al., 2010), which we will describe briefly in section 3.3, to construct these counterfactual synthetic cities. In principle, by considering a counterfactual, we should be able to abstract away from possible non-linear simultaneous adjustments that affected both German and synthetic cities in the same way. As stated before, we choose the SCM as looking at the upper tail of the city size distribution, it is rarer to find equally sized single unit counterfactuals, which SCM can possibly extend. We also add objectivity to the counterfactual selection as this convex combination is selected from a data-driven procedure.

Also, we can make SCM prioritize to match with similar cities that can be described as specialized cities that in its development we think a priori deviate from other cities in the same region or country. For instance, an increase in world trade may lead to higher economic activity in cities that have access to a river or

the sea, or an overall increase in alternative energies such as nuclear or oil that overall lead to decline in the relatively more expensive coal. Ideally in this case, we would want to create the synthetic city out of similar, specialized cities. The SCM can consider these developments as well.

What the SCM cannot do is to abstract from developments that are unique to Germany in this case, or unique to the synthetic cities. For example, the division of Germany was a rather unique development worldwide that could have lead to an overall population movement away from the division border. However, we should also note that most of German cities were not close to the division border, with Braunschweig, Kassel and Lübeck being the only ones close to 50km of the border. Furthermore, the division border region is not highly population, with the center of population in West Germany being further away to the West around the Rhine river in West Germany, or Berlin and Saxony in East-Germany. We do not see an obvious way to adjust for this through introducing another variable in SCM. However we can see whether those cities close to the border were relatively more negatively affected, and vice versa, which could point towards this kind of adjustment being relevant.

Nevertheless, we more closely boil down the possible attributable causes to the WWII shock plus the shocks that we cannot account for that are unique to Germany. Thus, it makes interpretation of any permanent impact easier to attribute to the WWII shock, which in our case we may now conveniently think of as consisting of the bombing shock as well as the division shock. However, other causes unique to Germany that we do not immediately think of here and have nothing to do with the WWII shock may also play a role, which we cannot account for through our approach. Yet still, we can still adjust for two types of developments i.e. developments specific to (West European) cities, as well those specific to 'specialized' cities; thus arguably we approach the actual counterfactual better than not adjusting for these developments, which in the empirical evaluation we think can lead to a lower overacceptance of the unit root null.

3.3 Synthetic City Construction

We turn to the construction of the synthetic control city here. We will follow closely the notation of Abadie et al. (2015) in the following (see Abadie (2019) for a comprehensive guide and discussion on the use of the synthetic control method). That is, the synthetic German city, indexed by i , is represented by a convex combination of donor pool cities as indexed by j . Therefore, we have a sample of $I + J$ individual cities, with I number of 'treated unit' German cities, and J number of 'comparison units' donor cities in the 'donor pool'. Hence, the donor pool is

$j = I + 1, \dots, I + J + 1$.

In the following, we will leave out the individual city superscript i for each variable for simplicity. The synthetic control for the individual city is a convex combination of cities in the donor pool i.e. it can be represented by the $(J \times 1)$ vector of weights $W = (w_{I+1}, \dots, w_{I+J+1})$ for each city, where $0 \leq w_j \leq 1$ and $w_{I+1} + \dots + w_{I+J+1} = 1$.

Let X_1 be a $(K \times 1)$ vector which contains the pre-treatment covariate of a treated city and let X_0 be a $(K \times J)$ matrix which contains the same pre-treatment covariate but from the donor pool.

We want to minimize the discrepancy between X_1 and X_0W i.e. minimize

$$\|X_1 - X_0W\| = \sqrt{(X_1 - X_0W)'V(X_1 - X_0W)} \quad (3.1)$$

which is done in a two-step procedure: Conditional on V , the synthetic control $W^*(V) = (w_{I+1}^*, \dots, w_{I+J+1}^*)$ is selected to minimize (1) subject to $w_{I+1}^* + \dots + w_{I+J+1}^* = 1$ and $0 \leq w_j^* \leq 1 \forall j$. V is a $(k \times k)$ diagonal matrix with non-negative entries, where its values weights the relative importance of each pre-treatment covariate. The idea of this approach is that we want the covariates to have strong predictive power over the outcome of interest, so that relative strong predictors are given relatively larger weights. Equivalently, let X_{0jm} be the value for the m -th covariate for donor pool unit j and X_{1m} be the value for the treated unit covariate, where $m = 1, \dots, K$, so that we choose W^* which minimizes

$$\sum_{m=1}^K v_m \left(X_{1m} - \sum_{j=I+1}^{I+J+1} w_j X_{0jm} \right)^2 \quad (3.2)$$

where v_m are the diagonal entries of V , i.e. $V = \text{diag}(v_1, \dots, v_K)$.

One way to choose V , as done by Abadie & Gardeazabal (2003); Abadie et al. (2010, 2015) is based on minimizing the mean squared prediction error (MSPE) of the city population outcome variable S given the chosen W^* before:

$$\sum_{t=1}^{T_0} \left(S_{1t} - \sum_{j=I+1}^{I+J+1} w_j^*(V) S_{jt} \right)^2 \quad (3.3)$$

where T_0 is the treatment period. Summarizing, we first calculate $W^*(V)$ for any V in the first step, and then choose the V that gives us the $W^*(V)$ which minimizes the MSPE.

The literature follows the principle in the choice of the covariate in that it should be a good predictor of the variable of interest, so that typically, the pre-treatment lagged outcome variables are included.

In our case, we should not include all pre-treatment lags of the outcome as covariates, as SCM would choose V such that the synthetic and actual population in the pre-treatment years is as close as possible to another, resulting also in other covariates to be irrelevant in the matching process (Kaul et al., 2015). Adding all pre-treatment lags would hence ignore specialized cities that we try to match with the synthetic specialized counterpart through the sea, river and coal covariate. Second, we also want the covariates to the WWII treatment period, as covariates further away may have a worse prediction power on the post-WWII period, including the lagged population covariates.

3.4 Data

For the outcome of interest, city population, we extend the original German city dataset of Brakman et al. (2004) and Bosker, Brakman, Garretsen, & Schramm (2008) that considered cities with at least 50000 inhabitants at some point in time, where we have additionally included the years 1871 to 1925 where possible, so that we have data for 1871 to 2000. As for the donor pool cities, we overall want the pool of cities to be comparable to German cities, but to not be destroyed by WWII itself.

Within Europe, WWII neutral countries such as Switzerland, Portugal and Sweden seem to be the obvious choice in this case as no fighting took place. We have also included Denmark, as its cities came out unscathed during the 'Six Hour War'. We have not included Norway, Ireland, the Netherlands, Belgium and France as its censuses were not taken in an 10 year interval consistently throughout the 19th century. Especially in the year 1940 that is crucial for the synthetic control fitting, the censuses were often delayed until right after WWII. However, these neutral countries lack coal fields, so that they would not be able to reflect the 'Strukturwandel' decline of the coal industry and factories that are build close to these coal fields, and which we suspect to be one main driver of decline of cities around the Ruhr coal area in Germany.

Within Europe, Great Britain has similarly large coal field and its industrialized regions mainly around the Midland regions, so that we also add English and Welsh cities taken from GB Historical GIS / University of Portsmouth (2021).⁵

Note here that we also do not have a census for England and Wales in 1941 due to the war. However, there was a register during the war in 1939 to issue identity cards in order to coordinate rations. These do not include military personnel, so that we may especially undercount the population in cities with unusually large military presence, such as Portsmouth as a Navy port. To address this, we have

⁵We have excluded Scottish cities due to the lack of data in 1890 and 1900.

used the 1951 number of military personnel in a city and added them to the 1939 register with a factor 1.54, which reflects the ratio of the number of total military personnel (1.27 million, from Tucker (2003)) in September 1939 when the register took place and the personnel (0.8238 million, from Defence Analytical Services and Advice (2011)) in 1951.⁶ For the cities that are not naval bases, this adjustment would on average increase the population figures by 1.3 percent, so that this adjustment changes relatively little for most cities.⁷

Due to possible changes in municipal border over time, we have opted for the constant border population estimates for European cities by DG REGIO (2021), available for 1960 onwards. We have also looked up whether there were any municipal border changes before 1960 for each city in our dataset and adjusted these population numbers if possible to allow constant borders throughout.

Furthermore, these cities were also bombed during what is commonly known as the Blitz; but overall less severe and with the main focus on London and air fields in particular, as opposed to the widely used strategy of 'area bombing' employed by commander Arthur Harris during the war that targeted German city centers more severely without a particular focus on military importance (Brakman et al., 2004). The U.K. were able to accurately register deaths and injuries for each bombing raid sortie, which was documented twice a day for the Ministry of war security and were made available in the National Archives.⁸ This stands in contrast to German cities for which only estimates are available, such as from Friedrich (2002). With this data, we are able to construct deaths per capita in each city within Great Britain that were caused by the Blitz, and consider only those that can be considered as being unscathed during the war.

The lowest estimated deaths per 1939 capita of German cities for which we have casualty data are Heidelberg, Göttingen, and Regensburg with 0.0008, 0.0017, and 0.0021 respectively, and which are considered to have come out of the war unscathed, leaving the city largely undamaged (Lauer, 2020; Heinzelmann, 2003; Friedrich, 2002).⁹ We have in total census data for 75 cities in England and Wales, taken from GB Historical GIS / University of Portsmouth (2021) at the district or unitary authority level at a decennial interval from 1881 to 1991.

⁶Here, we should note that the high number of military personnel even in 1951 is partially due to the National Service Act of 1948 that kept the number of military personnel high for a few years after WWII. On the other hand, the relatively low number of personnel conscripted is reflected in September 1939 with 1.27 million, as compared to the 3.29 in April 1941 or 4.91 million in April 1945. The comparable personnel numbers for 1939 and 1951 helps us in avoiding overextrapolation through a low factor of 1.54.

⁷For the naval bases Plymouth and Portsmouth however, this increase is at 11.3 and 13.9 percent.

⁸The dataset was cleaned up and made available at www.warstateandsociety.com/Bombing-Britain from the original source National Archives UK (2021).

⁹Respectively, the total deaths were 72 (Lauer, 2020), 107 (Heinzelmann, 2003) and 203 (Friedrich, 2002)

Taking these figures as a point of orientation on which cities can be considered as largely undamaged out of the English and Welsh cities, we have 13 that are above a 'unscathed' threshold of 0.002, and 20 that are above 0.001 deaths per 1940 capita.

For comparison, at least for the cities for which we have estimated death figures by using the source of Friedrich (2002), the average rate for German cities is a magnitude higher at around 0.02 against the average of 0.0008 during the Blitz. Only 2 out of 42 German cities for which we have estimations on death figures caused by the WWII bombings would be below the 0.002 threshold, which stands in contrast to the cities of 75 Great Britain in our dataset with at least 50000 inhabitants at some point between 1880 to 1990, where 62 English and Welsh cities are below the 0.002 threshold, and 55 are below the 0.001 threshold. Hence, although most cities in Great Britain were raided, most can be considered as undamaged or only slightly damaged. From the 62 'unscathed' cities, we discard 12 due to inconsistent borders over time that we were not able to fix. This leaves us with 50 English and Welsh cities in total.

Additionally, with the requirement of at least 50000 inhabitants at some point in time, we have added 16 Swiss from the Bundesamt für Statistik (Bundesamt für Statistik, Schweiz, 2016), 38 Swedish cities from (Swärd, 2016), 30 Portuguese cities from (Instituto Nacional de Estatística, 2011), and 10 Danish cities (Danmarks Statistik, 1985) and (Danmarks Statistik, 2011).

As for the German cities, we usually do not have the population data for the surrounding cities which were annexed by the German cities in our sample. To adjust for that, we adapt the series in the same manner as Bosker, Brakman, Garretsen, & Schramm (2008), where we assume the same growth rate for annexing and annexed places and extrapolate it on the additional population; this time with the city boundary fixed around WWII. However, this adaptation is likely to be less accurate the farther away we extrapolate from the time of annexation/amalgamation. We can account for that in the SCM by excluding the covariates which are far away from the WWII treatment period. We have done the same adjustment for some English and Welsh cities when they switched the border definition for which the city population was recorded. Second, we had to exclude Berlin and Hamburg due to the lack of cities that are as large as them in our donor pool. However, we would have excluded Berlin in any case as it can be considered an island for most of its post-WWII history until the reunification. Third, there was a change in how the population was counted in 1987 throughout Germany, so that there may be a break which we cannot adjust for properly.¹⁰ Hence, our preferred specification

¹⁰The definition was changed from *Wohnbevölkerung* (inhabitants living there) to *Bevölkerung am Ort der Hauptwohnung* (inhabitants at the place of the main address)

would end right before or around 1987.

In the end due to year-data limitations of each individual country dataset, we will have a donor pool of 144 cities consisting of English, Welsh, Swiss, Swedish, Danish and Portuguese cities that came out unscathed from the war, and a treatment pool of 53 German cities, all ranging from the years 1880 until 1990 at a decennial interval.

As discussed above on specialized cities, first we will also consider a coal covariate. This inclusion is largely motivated by the decline of the coal industry in the Ruhr area and the UK, that started around the 1960s. For that, we have obtained the number of workers in the coal industry from Fischer (1989, 1995) on a Regierungsbezirk (administration area) level in Germany. Similarly, we have coal worker employed on a historic county level for both England and Wales employed in the coal industry, taken from commerce reports on the UK Dye (1921). There was no significant coal deposit in any other country, or large industrial areas for that matter. There is a large iron industry in Sweden, but it is located in the remote and not very populated place in Kiruna and the iron industry there did not decline over the course of the 20th century.

We will show in the next section that most synthetic cities exhibit an inverse U-shaped trajectory, a development that may be attributable to the post-war suburbanization movement. The job of the coal covariate should be to capture the 'extra' decline of the population due to that decline in coal, apart from the already occurring decline. Second, the main industrial areas are found within or nearby coal regions, so that a coal covariate should also cover the decline of manufacturing cities or areas and deindustrialization. Indeed, the industrial heartland between these two countries, the Ruhr and the Midland area, are close to or within major coal fields.

Third, the population trajectory may already capture the coal covariate, as industrializing cities around the beginning of the 20th century attracted notably more workers, that the lagged population covariates would capture as well. However, we may have a city in the donor pool that exhibits a very similar pre-war trajectory as these coal cities but are not subject to that decline. Thus, if the SCM would consider these cities to match coal cities, we may underestimate the counterfactual post-war decline. The idea of the coal covariate is that SCM will give 'preference', given equally well-fitting pre-war population trajectory, towards matching cities within coal regions with coal region synthetic cities specifically, and vice versa cities within non-coal regions with non-coal region synthetic cities. For that reason, we consider the coal covariate important to include as it may capture the 'extra' population decline, apart from the ones induced by suburbanization for example.

Lastly, for the same reason as including a coal covariate, we include a river and sea-access dummy as simple first-nature determinants that considers adjustments of international trade over time, where we consider a city a river or sea city if today's border is 1 kilometer away from either. These may capture the 'extra' population increase if trade became more important post-WWII.

3.5 Descriptive Statistics and Inverse U-shaped Growth

Table 3.1 presents the summary statistics of our 53 German sample cities and the cities of the donor countries. Next, we will show that the synthetic city population growth throughout time is found to be typically inverse U-shaped. An inverse U-shaped city population growth for German cities can be found by looking at the behavior of the synthetically estimated population growth of German cities over time. That is, we estimate the following simple quadratic model:

$$S_{i,t}^{synth} = \alpha_i + \beta_{1,i}T + \beta_{2,i}T^2 + \epsilon_{i,t} \quad (3.4)$$

with time trend T , and where $S_{i,t}^{synth}$ is the synthetically estimated city population for the corresponding German city i at time t .

To construct the synthetic control for now, we consider the covariates to be 1900 and 1920 to 1940 lagged population, 1920 number of coal workers, dummies for sea or river access, and lastly the simple 1940 urban potential measure for each individual country, thus:

$$UP_i = \sum_j^n \frac{S_{i,1940}}{\tau_{i,j}} \quad (3.5)$$

where $\tau_{i,j}$ is the kilometer distance between city i and j , and $S_{i,1940}$ is the population of city i in year 1940. For each individual synthetic city from 1900 to 1990, we find 42 out of 53 cities where $\beta_{1,i} > 0$ and $\beta_{2,i} < 0$, of which 36 out of 42 are significant for both at the 5 percent level. The average maximum of those 36 cities is around the year 1979.0.¹¹ Thus, it suggests that there are non-linear adjustments that have less to do with WWII at play here, with the trajectory being mostly described as an inverse U-shaped. These results indicate that there are underlying, systemic non-random factors which can influence the patterns of growth of the majority of cities which are not initiated by a large, temporary shock on cities itself. Furthermore, it provides some justification for our approach to accommodate for a possible non-linear counterfactual.

¹¹Similarly, if we consider 1920 to 1990 that is close to the time period that Bosker, Brakman, Garretsen, & Schramm (2008) considered, we would find 46 out of 53 cities where $\beta_{1,i} > 0$ and $\beta_{2,i} < 0$, out of which 32 significant at the 5 percent level. The average maximum of those 32 cities is around the year 1977.3.

Table 3.2 presents 3 arbitrarily chosen example cities, Aachen, Lübeck and Düsseldorf. For Aachen, the discrepancy of the lagged actual and synthetic population is relatively low, while the discrepancy for the coal worker covariate is relatively high. As discussed in the previous section, this may be due to county level coal workers between regions not being directly comparable or having cities that are far away from coal fields included within coal counties. Similarly for Düsseldorf, the synthetic control method does not match the city with any coal region city. Indeed, we would not think of Düsseldorf as a typical city that is characterized as a coal area city, as it is close but not within the Ruhr coal area. Coal workers assigned at a Regierungsbezirk level may then wrongly include cities such as Düsseldorf as a coal area. However, the fact that Düsseldorf is not matched with any coal area counterfactual cities is reassuring, as it indicates that the most weight in the matching process is given towards the lagged population covariates. This may be due to the different population trajectories of coal and non-coal cities, as the often rapidly industrializing coal cities had much higher growth rates pre-WWII.

Table 3.1: Descriptive Statistics

	mean (std. dev.)	min / max	mean (std. dev.)	min / max	mean (std. dev.)	min / max
Covariates	German cities		Synthetic cities		Donor pool	
population						
- (1900)	134480 (112831)	11704 / 534175	137582 (117275)	16198 / 533926	101383 (129882)	1138 / 842277
- (1920)	194007 (170816)	46990 / 699513	188816 (161679)	47003 / 698638	120747 (151217)	1628 / 943941
- (1930)	209416 (184136)	54500 / 759381	207006 (180159)	52197 / 767957	129505 (164122)	2716 / 1030956
- (1940)	220939 (193467)	58713 / 840188	222419 (195616)	57125 / 835473	135698 (168584)	3584 / 1046228
river	0.774 (0.423)	0 / 1			0.917 (0.277)	0 / 1
sea	0.019 (0.137)	0 / 1			0.319 (0.468)	0 / 1
coal worker						
- (1920)	62518 (82687)	0 / 219758			39346 (69607)	0 / 255000
urb. pot.	381356 (214026)	136102 / 972784			233538 (235975)	35326 / 1298276
- (1940)						
Non-covariates						
population						
- (1950)	211312 (175908)	54100 / 816924	248277 (219625)	62761 / 918915	149520 (188550)	7269 / 1148872
- (1960)	263290 (226693)	68400 / 1061359	255478 (212027)	70829 / 884986	159883 (197952)	13688 / 1249621
- (1970)	268313 (241572)	68700 / 1290071	251027 (190815)	78391 / 784843	166579 (201427)	26743 / 1485589
- (1980)	252265 (227949)	65154 / 1277230	240350 (176452)	79449 / 778944	168212 (219135)	40193 / 1961031
- (1990)	249724 (219536)	63525 / 1208484	236850 (171675)	80939 / 780176	168226 (219198)	40462 / 2010080
Sample size	53		53		143	

3.6 Empirical Strategy

To see whether we have had any permanent impact or not on a city's population after WWII, we test for mean-reversion of the city population by testing for the absence of a unit root. Consider the typical augmented Dickey-Fuller estimable equation given below:

Table 3.2: Individual City Examples

covariates	Treatment Aachen	Synthetic Aachen	Treatment Lübeck	Synthetic Lübeck	Treatment Düsseldorf	Synthetic Düsseldorf
population						
- (1900)	142427	142164.9	82098	82168.29	281492	293331.3
- (1920)	151731	151319.6	118709	118850.4	453026	435615.5
- (1930)	154682	154856	129842	130004.7	476300	482981.9
- (1940)	162164	161832.1	154811	155007.7	541410	533579.4
river	0	0.041	1	1	1	1
sea	0	0.252	1	1	0	1
coal worker	15426	43509.6	0	270.8	219758	0
- (1920)						
urb. pot.	265614.3	268480	245207	249135.5	741768	621672.5
- (1940)						
	Donor city	Weight	Donor city	Weight	Donor city	Weight
	Barrow, UK	2.9%	Aarhus, DK	32.1%	Kopenhagen, DK	57.9%
	Bradford, UK	10.9%	Blackpool, UK	0.2%	Lisboa, PT	1.7%
	Braga, PT	71.8%	Brighton, UK	28.3%	Southend, UK	40.5%
	Brighton, UK	1.2%	Kolding, DK	13.9%		
	Oldham, UK	5%	Malmö, SWE	7.3%		
	Sheffield, UK	8.1%	Southend, UK	18.3%		

$$\Delta S_{i,t} = c_i + \delta_i t + \rho_i S_{i,t} + \sum_{k=1}^p \beta_{i,k} \Delta S_{i,t-k} + \epsilon_{i,t} \quad (3.6)$$

For the city size $S_{i,t}$ for city i at time t . Finding ρ_i to not be significantly different from 0 would constitute a unit root here, such that city population trajectory is not linear trend mean reverting. On the other hand, estimating to be significantly $\rho_i < 0$, we would find a trend-stationary process here so that the city is linear trend mean reverting after a negative WWII shock.

Equation (3.6) had been estimated in Bosker, Brakman, Garretsen, & Schramm (2008), however by taking the natural log before. In the case of an inverse-U shaped counterfactual development, whether we take the log or not would not matter if we abstract away the trend in (7). However, we would recover approximate growth rates at the left side of the equation, which would make it the ideal test for Gibrat's law that states that urban growth rates are independent of its size. For now, we focus on whether cities are mean-reverting to its assumed post-war city size trajectory. As a preview to a robustness exercise, Table 3.10 reports the results of Table 3.3, however where we took the log of $S_{i,t}$ and $S_{i,t}^{synth}$ in (3.6) and (3.7). As compared to Table 3.3, we generally reject the unit root null more often.

We obtain the synthetically corrected (SC) form as the difference of the actual individual city population $S_{i,t}$ with its corresponding synthetic city population $S_{i,t}^{synth}$, thus 'detrending' it:

$$S_{i,t}^{SC} \equiv S_{i,t} - S_{i,t}^{synth} \quad (3.7)$$

Note that (3.7) is equivalent to the cointegration relation of $S_{i,t} = \alpha_i + \gamma_i S_{i,t}^{synth} + u_{i,t}$ by imposing $\alpha_i = 0$ and $\gamma_i = 1$, so that $S_{i,t}^{SC} = u_{i,t}$. In this case of known coefficients, the usual Dickey-Fuller critical values apply. However, we are only interested in whether the difference of the actual $S_{i,t}$ and synthetic series $S_{i,t}^{synth}$, $S_{i,t}^{SC}$, is mean-reverting, and not whether there is a α_i and γ_i such that two I(1) series are cointegrated. Hence, we are not interested in the order of integration as well.

Then, equivalent to 3.6, we write:

$$\Delta S_{i,t}^{SC} = \rho_i S_{i,t}^{SC} + \sum_{k=1}^p \beta_{i,k} \Delta S_{i,t-k}^{SC} + \epsilon_{i,t}^{SC} \quad (3.8)$$

That is, as we detrend through the corresponding synthetic city for each German city, we drop the linear-trend term δt . Also by construction, the values for $S_{i,t}^{SC}$ should be close to 0 before the war and are allowed to diverge from 0 after the war, so we also drop the constant c_i in (3.7), and consider the Dickey-Fuller critical values without a constant/drift in the following to be appropriate. The interpretation is similar to above, with the difference that we estimate whether cities return to its supposed mean as estimated by its synthetic control. Estimating ρ_i to be 0 would similarly constitute a unit root in the population trajectory after detrending with the synthetic city, while $\rho_i < 0$ would be interpreted as a reversion to the synthetic mean population trajectory after a negative divergence from it i.e. no permanent effect.

3.7 Results

3.7.1 WWII Bombings Impact on City Size

We keep the same synthetic city as in section 3.5, where the covariates are 1900 and 1920 to 1940 lagged population, 1920 number of coal workers, dummies for sea or river access, and the 1940 urban potential measure as above. Furthermore, the MSPE was set between 1880 and 1940.

Figure 3.1 shows the gap of the population with the synthetic population for the individual city. As expected, the gap is close to 0 for almost all cities between 1920-1940 by construction due to the choice of lagged covariates. Also notice that the orange dashed linear fit from 1900 to 1940 is close to 0, which reaffirms our exclusion of a trend term and a constant term in equation (3.8). Secondly, ignoring the WWII shock and similar to the cutoff year chosen by Davis & Weinstein (2002) and Brakman et al. (2004) in the 1960s, the yellow dashed line is the linear fit between 1960 and 1990. Again, we see a line close to 0 i.e. on average, there

Figure 3.1: Individual City Gap of Population and Synthetic Population (Donor Pool: <0.002 bombing deaths p.c.)

Note: The total number of cities is 53 here, and 143 in the donor pool which had below 0.002 deaths per capita during the Blitz. The figure shows the first specification with the 1900 and 1920 to 1940 lagged population, 1920 number of coal workers, dummies for sea or river access, and the 1940 urban potential measure. The MSPE was set between 1880 and 1940. The orange, dashed line shows the linear fit before 1940, while the yellow dashed line shows the linear fit after 1960 to exclude most of the WWII shock. The city that reaches a gap of over 500000 in 1980 is Munich.

seems to be no difference between the German city population and its synthetic counterpart decades after the war.

In Figure 3.2 we show the gap of the 56 cities in the donor pool which have a weight of at least 0.05 as a comparison city for some German city. The synthetic cities are the first specification with the 1900 and 1920 to 1940 lagged population, 1920 number of coal workers, dummies for sea or river access, and the 1940 urban potential measure. Here, we show the gap of these individual comparison cities with its synthetic control, where the donor pool is simply the same donor pool as before, with the exception of the comparison city itself. Here, we exclude Birmingham as there are no city comparable in size that we can use as a counterfactual. That the linear fit is a horizontal line around 0 is reassuring as this is what we should expect. That is, as the donor pool cities did not experience a large, temporary shock on its city system as compared to Germany, we should find on average no effect i.e. a horizontal line at 0.

Table 3.3a shows the individual city result of the Dickey-Fuller test on the

Table 3.3: Unit Root Test on City Size (Donor Pool: <0.002 bombing deaths p.c.)

A. Individual City (augmented) Dickey-Fuller test					
Period	1900-1990	1900-1980	1900-1970	1900-1960	1900-2000
a. Synthetic detrended ($c_i = 0$)					
Significance level	% Unit Root Rejected				
1%	7.5	11.3	11.3	28.3	
5%	24.5	26.4	34.0	54.7	
10%	32.1	35.8	49.1	66.0	
Out of which % Unit Root with Positive Impact ($S_{i,t} > S_{i,t}^{synth}$)					
1%	53.1	48.9	53.2	42.1	
5%	47.5	46.2	48.6	45.8	
10%	47.2	41.2	51.9	50.0	
b. Conventional Method with Linear Trend ($\delta_i t = 0$ or $\delta_i t \neq 0$)					
Significance level	% Unit Root Rejected				
1%	0	0	0	3.8	0
5%	5.7	5.7	7.5	9.4	3.8
10%	7.5	9.4	13.2	13.2	5.7
B. Panel Unit root test					
a. Synthetic detrended ($c_i = 0$)					
	t-stat (p-val)	t-stat (p-val)	t-stat (p-val)	t-stat (p-val)	t-stat (p-val)
Levin-Lin-Chu	-3.425 (0.000)	-4.992 (0.000)	-7.415 (0.000)	-11.807 (0.000)	
b. Conventional Method with Linear Trend (i.e. with trend term $\delta_i t$)					
Levin-Lin-Chu	-11.856 (0.000)	-13.078 (0.000)	-18.242 (0.000)	-18.530 (0.000)	-12.179 (0.000)
Im-Pesaran-Shin	-0.071 (0.472)	-0.600 (0.274)	-3.833 (0.000)	-4.811 (0.000)	-0.732 (0.232)

Note: The null hypothesis is a unit root in city size detrended by the synthetic city size. The total number of cities is 53 for all specifications, and 143 in the donor pool which had below 0.002 deaths per capita during the Blitz. The synthetic detrend Dickey-Fuller test in 3a, as in equation (8), is estimated through the first specification with the 1900 and 1920 to 1940 lagged population, 1920 number of coal workers, dummies for sea or river access, and the 1940 urban potential measure. The MSPE was set between 1880 and 1940. The linear detrended Dickey-Fuller test in 3b follows equation (6) with trend term $\delta_i t$. The individual city unit root number of lags are chosen following Ng & Perron (1995), where the maximum lag length started at 2 for this procedure. Incidentally, the lag is chosen to be zero for all cities. $c_i = 0$ indicates a suppressed constant in the regression. $\delta_i t = 0$ or $\delta_i t \neq 0$ indicates an added trend in the regression. Given the individual city optimal number of lags, we set the lag to 0 for the panel unit root tests. Panel statistics are calculated with adjusting for cross-sectional dependence.

Figure 3.2: ≥ 0.05 Weight Comparison City Gap

Note: The total number of cities is 56, which corresponds to the cities of the first specification (Table 3.3) in the donor pool that had at least a weight of 0.05 in the creation of any of the 53 synthetic cities of Figure 3.1. See the footnote of Figure 3.1 for the synthetic control creation specifications. We exclude Birmingham from this figure as there is no city in the donor pool that can match its population pre-WWII. The orange, dashed line shows the linear fit before 1940, while the yellow dashed line shows the linear fit after 1960.

population as in equation (3.8) detrended by the synthetic city, and in Table 3.3b as in equation (3.6) the conventional linear detrended with $\delta_i t$ Dickey-Fuller test. First, we reject the unit root null for around 24.5% at the 5% of the cities if we consider the cutoff year of 1990. We reject a similar amount of cities at the 5% for the cutoff years 1980 or 1970. Only around the cutoff year of 1960 do we reject the unit root null for 54.7%. The pattern seems clear that a earlier cutoff year leads to a higher rejection of the unit root hypothesis. One reason is that the synthetic counterfactual is less accurate the further away we move from the treatment period. Still, the rate of cities with rejected unit roots stays remarkably stable over time.

Even more remarkable is the percentage of the cities where we find a unit root and for which we find the population to be above the corresponding synthetic population. Throughout all cutoff years, we find this percentage to be around 50%, which may imply that the percentage of rejected unit roots is underestimated given that the counterfactual is less accurate over time. This stands in contrast to the mostly negative structural breaks around WWII that were found by Bosker,

Brakman, Garretsen, & Schramm (2008) within a very similar sample.¹²

They also found a unit root for all cities when not considering a structural break, which is consistent with the conventional linear detrended approach results in Table 3.3b, where we reject the unit root null for only 5.6% at the 5% level with a cutoff year of either 1990 or 1980. We overall consider this as a strong support for the notion of overacceptance of the unit root null by the conventional linear detrended approach in the light of simultaneous non-linear adjustments.¹³

For the sake of completeness to the previous literature that have considered a panel unit root test, we have also included some panel unit root tests similar to Bosker, Brakman, Garretsen, & Schramm (2008). The initial studies here have resorted to panel unit root tests, noting the low power of individual unit root tests with few observations. As noted in Perron (1991), it is rather a short time span that would be problematic, and not the frequency of the observations. First, we have included the Levin et al. (2002) test, with the null of all panels containing the unit root against the alternative of all panels being stationary with the same autoregressive parameters. For both, the synthetic and linear detrended approach, we reject the null of all panels containing a unit root.

A somewhat less restrictive test that allows for different autoregressive parameters between cities is Im et al. (2003) that tests the null of all panels having a unit root against the alternative of some panels being stationary and some having a unit root. This test however does not allow for suppressing the constant/drift term. Hence, we decided to not include to test the synthetically detrended approach with a constant as it would artificially inflate the critical values to overaccept the null hypothesis for the synthetically detrended approach, as the constant is by construction set to 0 in the pre-war period. For the linear detrended approach, we fail to reject the null of all panels having a unit root until 1980, which is more in line with the corresponding findings for individual cities. Hence, it seems that allowing for different autoregressive parameters matters here. However, the conclusions drawn seem to be sensitive to the cutoff year, as we would conclude the alternative for a cutoff year in 1970 despite the individual city unit root results not changing much throughout. In that regard, one advantage of the individual city unit roots results taken together is that it adds nuance to the conclusions drawn and is, taken together at least, less sensitive to these sudden changes in the conclusion.

As a robustness exercise, we may think of taking a log transformation of $S_{i,t}$ and $S_{i,t}^{synth}$ in (6) and (7), similar to the literature. As argued before, taking the log for synthetic controls would not be needed as the deterministic mean is already

¹²Out of 62 cities, they find 17 cities with a structural break around WWII, out of which 15 have a negative structural break. For the other cities, they find a unit root.

¹³We have also done the same exercise of Table 3.3b on the synthetic counterpart series instead of the actual population, and find a slightly lower rejection rate of the unit root null.

taken away. Instead, a log transformation would deemphasize the divergences of the gap of actual and synthetic population that happens after the treatment, as seen in Figure 3.1. In Table 3.10a, we report the same specification as in Table 3.3a, however where we took the log as mentioned just before. As compared to Table 3.3a, we generally reject the unit root null more often by around 15 percentage points for basically all specifications, consistent with the idea that such a transformation would decrease the fanning out after the treatment in Figure 3.1. Whether this is the right specification is not obvious, but it may reduce the likely inaccuracy of the synthetic control the further we move away from the treatment period.

Lastly, we reduce the threshold of deaths per capita during the Blitz from 0.002 as in Table 3.3, to 0.001 in Table 3.4, while Figure 3.3 is the corresponding gap of the population to the synthetic population graph to this table. This reduced threshold drops cities such as Birmingham from the donor pool, which stands at a rate of just above 0.0011. Overall, the donor pool will drop from 143 cities to 138. Hence, even though the original threshold of 0.002 corresponds to the very few German cities that were considered to be unharmed during the war, such as Göttingen or Regensburg, we reduce the threshold even further to the level of undamaged Heidelberg, in exchange for having less cities in the donor pool, that may result in trouble matching cities at the upper end of the population distribution. Overall, we now have a higher rate of cities where we find mean reversion. Second, the ratio of positive to negative impact unit root cities is also slightly but consistently lower for all cutoff years. We suspect this to be due to the donor pool being less covered at the upper tail, so that the SCM has fewer options to match larger cities.

Bosker, Brakman, Garretsen, & Schramm (2008) also looked at the evolution of the city size distribution. They found a more even city size distribution i.e. relatively more middle sized cities. However, they were only able to compare the post-war distribution to the pre-war distribution. We think that we are now in a better position for this, since we have an actual counterfactual in the synthetic control, so we can compare the actual city size distribution with the synthetic city size distribution over time. For that, we will perform the two-sample Kolmogorov-Smirnov test on the actual and the synthetic city size for each year. We report the results in Table 3.5.

For each year, we fail to reject the null that both distributions are the same. Note that we more strongly reject the null between 1920 and 1940, which is by construction due to the synthetic control lag covariates we included in those years. Also note that the p-value declines after the war over time, which may indicate that the effect of the bombings on the city size distribution did not fully dissipate yet and that we may reject the null if we extend the cutoff year; and this would

Table 3.4: Unit Root Test on City Size (Donor Pool: <0.001 bombing deaths p.c.)

A. Individual City (augmented) Dickey-Fuller test					
Period	1900-1990	1900-1980	1900-1970	1900-1960	1900-2000
a. Synthetic detrended ($c_i = 0$)					
Significance level	% Unit Root Rejected				
1%	11.3	17.0	17.0	33.4	
5%	26.4	34.0	45.3	54.7	
10%	37.7	41.5	52.8	67.9	
Out of which % Unit Root with Positive Impact ($S_{i,t} > S_{i,t}^{synth}$)					
1%	40.4	43.2	52.3	45.7	
5%	41.0	40.9	48.3	45.8	
10%	48.4	38.7	44.0	47.0	
b. Conventional Method with Linear Trend (i.e. with trend term $\delta_i t$)					
Significance level	% Unit Root Rejected				
1%	0	0	0	3.8	0
5%	5.7	5.7	7.5	9.4	3.8
10%	7.5	9.4	13.2	13.2	5.6
B. Panel Unit root test					
a. Synthetic detrended ($c_i = 0$)					
	t-stat (p-val)	t-stat (p-val)	t-stat (p-val)	t-stat (p-val)	t-stat (p-val)
Levin-Lin-Chu	-5.575 (0.000)	-7.596 (0.000)	-8.667 (0.000)	-11.911 (0.000)	
b. Conventional Method with Linear Trend (i.e. with trend term $\delta_i t$)					
Levin-Lin-Chu	-11.856 (0.000)	-13.078 (0.000)	-18.242 (0.000)	-18.530 (0.000)	-12.179 (0.000)
Im-Pesaran-Shin	-0.071 (0.472)	-0.600 (0.274)	-3.833 (0.000)	-4.811 (0.000)	-0.732 (0.232)

Note: The null hypothesis is a unit root in city size detrended by the synthetic city size. The total number of cities is 53 for all specifications, and 138 in the donor pool which had below 0.001 deaths per capita during the Blitz. The synthetic detrend Dickey-Fuller test in 3a, as in equation (8), is estimated through the first specification with the 1900 and 1920 to 1940 lagged population, 1920 number of coal workers, dummies for sea or river access, and the 1940 urban potential measure. The MSPE was set between 1880 and 1940. The linear detrended Dickey-Fuller test in 3b follows equation (6) with trend term $\delta_i t$. The individual city unit root number of lags are chosen following Ng & Perron (1995), where the maximum lag length started at 2 for this procedure. Incidentally, the lag is chosen to be zero for all cities. $c_i = 0$ indicates a suppressed constant in the regression. $\delta_i t = 0$ or $\delta_i t \neq 0$ indicates an added trend in the regression. Given the individual city optimal number of lags, we set the lag to 0 for the panel unit root tests. Panel statistics are calculated with adjusting for cross-sectional dependence.

Figure 3.3: Individual City Gap of Population and Synthetic Population (Donor Pool: <0.001 bombing deaths p.c.)

Note: The total number of cities is 53 here, and 138 in the donor pool which had below 0.001 deaths per capita during the Blitz. The figure shows the second specification (Table 3.4), with the reduced 0.001 threshold donor pool, with the 1900 and 1920 to 1940 lagged population, 1920 number of coal workers, dummies for sea or river access, and the 1940 urban potential measure. The orange, dashed line shows the linear fit before 1940, while the yellow dashed line shows the linear fit after 1960 to exclude most of the WWII shock. The city that reaches a gap of over 400000 in 1980 is Munich.

go against the very idea of having a cutoff year. Furthermore, given 33 cities out of 53 in our sample were above to the synthetic city population already in 1960, and heavily based on our estimate of this section, it would seem incredulous that it would take another 50 years for the effect of WWII to fully dissipate.

To conclude this section, despite the widespread bombings on German cities, the city size distribution did not diverge from its supposed counterfactual. Instead, based on the evidence in this section, it suggests the distribution change results of Bosker, Brakman, Garretsen, & Schramm (2008) to be driven by underlying factors which are not captured by their method and not necessarily by WWII, as well as an overacceptance of the unit root null. This however does not suggest that there is no permanent effect stemming from the WWII bombings, but that it did not induce a change in the distribution. Whether the WWII bombings were indeed not 'enough' to induce a change in the distribution i.e. that this change in the distribution would even be possible in the first place, cannot be answered here.

Table 3.5: Actual and synthetic city equality of distributions

Kolmogorov-Smirnov			
Year	p-value	Year	p-value
1880	0.273	1940	0.958
1890	0.529	1950	0.837
1900	0.905	1960	0.987
1910	0.744	1970	0.920
1920	0.784	1980	0.476
1930	1.000	1990	0.617

Note: The null hypothesis is the equality of both actual and synthetic distribution. The synthetic cities are the first specification (<0.002 bombing deaths p.c.) with the 1900 and 1920 to 1940 lagged population, 1920 number of coal workers, dummies for sea or river access, and the 1940 urban potential measure. The MSPE was set between 1880 and 1940. We report the exact p-values. The same exercise with the second specification (<0.001 bombing deaths p.c.) yields the same conclusion of no significant p-value for any time.

3.7.2 WWII Bombings Impact on the Relative City Size

Following Davis & Weinstein (2002), who looked at the impact of WWII bombings on the relative city size of Japanese cities on total population, other papers (Brakman et al. (2004), Bosker, Brakman, Garretsen, & Schramm (2008)) followed suit in looking at the impact of WWII on the relative city share of German cities among the total population. This allows us to specifically look at whether the population permanently changed its position within the city size distribution.

Here, Davis & Weinstein (2002) and Brakman et al. (2004) that set their cutoff years at 1963 and 1965 respectively, concluded mean reversion of the relative city size for both Japan and Germany, while Bosker, Brakman, Garretsen, & Schramm (2008) with a cutoff year of 1999 concluded no reversion to the mean for the relative city size. That the WWII shock did not matter two decades after WWII, but over the long-term of over 5 decades, is a peculiar result that is hard to explain. We can reconcile the literature by allowing for non-linear mean reversion especially as specialized cities may have developed differently to another that could have driven the relative city size irrespective of the WWII shock. The possible interpretation of these nonlinearities in the light of these theories will not be discussed here, but in section 9, we will briefly give interpretation of these non-linearities in the light of all theories together.

We follow Brakman et al. (2004) in the assumption that cities that were bombed were chosen at random, at least later in the war not depending on the size or military importance that may correlate with industrial capacity. The bombing targets were initially chosen to be of importance to the military, which proved to be unsuccessful in limiting war production. This led to area bombing as the new strategy that targeted not only industrialized areas, but also easier to hit historical city

centers, aiming to destroy the housing stock to disrupt overall production through the dislocation of the work force as well as to lower morale. Another criteria was the visibility from the air (weather conditions, landmarks), susceptibility to incendiary bombs (Friedrich, 2002), or distance to Britain as earlier aircraft could not reach more Southern areas. However, cities close Britain were sometimes not targeted as much i.e. Lübeck, Oldenburg as well.

With the total West German population¹⁴ $S_{i,t}^{total}$, the city share is $s_{i,t}$

$$s_{i,t} \equiv \frac{S_{i,t}}{S_{i,t}^{total}} \quad (3.9)$$

$$s_{i,t}^{synth} \equiv \frac{S_{i,t}^{synth}}{S_{i,t}^{total}} \quad (3.10)$$

We consider the synthetic city share to be the ratio of the same synthetic control as in the previous section, and the actual total West German population $s_{i,t}^{total}$. The synthetically detrended city share is

$$s_{i,t}^{SC} \equiv s_{i,t} - s_{i,t}^{synth} \quad (3.11)$$

Then, equivalent to (1), we write:

$$\Delta s_{i,t}^{SC} = \rho s_{i,t}^{SC} + \sum_{k=1}^p \beta_{i,k} \Delta s_{i,t-k}^{SC} + \nu_{i,t}^{SC} \quad (3.12)$$

Table 3.6 reports the results of the relative city size given the previously used synthetic control specification with a donor pool that considers 0.002 bombing deaths per capita. As compared to the absolute size effect, we overall find a higher rate of rejection of the unit root null for the relative share of the population of the total population. Second, we again find that the percentage of the cities where we find a unit root and for which we find the population to be above the corresponding synthetic population to be slightly below 50% throughout all cutoff years. Third, depending on the cutoff year, we find a higher to much higher rate of a rejection of the unit root null as compared to the conventional linear detrended approach. For example, we find mean-reversion for 71.7% at a 5% level in cities with a cutoff of 1960 and 69.8% for 1970, while it is only 13.2% and 15.1% for the corresponding linear detrended approach respectively; but this discrepancy drops with a later cutoff year.

For our preferred cutoff year 1980, which we can consider to be late enough to register any permanent effect after WWII, while the German population figures are

¹⁴We consider the territory of 1957-90 and keep it constant for the total population between 1880-1990.

Table 3.6: Unit Root Tests on Relative City Size (Donor Pool: <0.002 bombing deaths p.c.)

A. Individual City (augmented) Dickey-Fuller test					
Period	1900-1990	1900-1980	1900-1970	1900-1960	1900-2000
a. Synthetic detrended ($c_i = 0$)					
Significance level	% Unit Root Rejected				
1%	17.0	22.6	24.5	35.8	
5%	43.4	49.1	69.8	71.7	
10%	58.5	66.0	77.4	75.5	
Out of which % Unit Root with Positive Impact ($S_{i,t} > S_{i,t}^{synth}$)					
1%	50.0	46.3	45.0	38.2	
5%	46.7	37.0	29.4	46.7	
10%	36.3	22.2	33.3	46.2	
b. Conventional Method with Linear Trend (i.e. with trend term $\delta_i t$)					
Significance level	% Unit Root Rejected				
1%	9.4	5.7	5.7	3.8	13.2
5%	26.4	22.6	15.1	13.2	28.3
10%	37.7	22.6	30.2	32.1	45.3
B. Panel Unit root test					
a. Synthetic detrended ($c_i = 0$)					
	t-stat (p-val)	t-stat (p-val)	t-stat (p-val)	t-stat (p-val)	t-stat (p-val)
Levin-Lin-Chu	-10.713 (0.000)	-12.723 (0.000)	-15.242 (0.000)	-16.843 (0.000)	
b. Conventional Method with Linear Trend (i.e. with trend term $\delta_i t$)					
Levin-Lin-Chu	-15.619 (0.000)	-14.536 (0.000)	-14.717 (0.000)	-17.233 (0.000)	-16.074 (0.000)
Im-Pesaran-Shin	-4.800 (0.000)	-3.202 (0.001)	-3.012 (0.001)	-3.159 (0.001)	-6.068 (0.000)

Note: The null hypothesis is a unit root in the relative city size detrended by the synthetic relative city size. The total number of cities is 53 for all specifications, and 143 in the donor pool which had below 0.002 deaths per capita during the Blitz. The synthetic cities are the first specification with the 1900 and 1920 to 1940 lagged population, 1920 number of coal workers, dummies for sea or river access, and the 1940 urban potential measure. The MSPE was set between 1880 and 1940. The individual city unit root number of lags are chosen following Ng & Perron (1995), where the maximum lag length started at 2 for this procedure. Incidentally, the lag is chosen to be zero for all cities. $c_i = 0$ indicates a suppressed constant in the regression. $\delta_i t = 0$ or $\delta_i t \neq 0$ indicates an added trend in the regression. Given the individual city optimal number of lags, we set the lag to 0 for the panel unit root tests.

not too much affected by the amalgamation wave of the mid-1970s or the change in the city population definition of 1987.¹⁵ We still find at a 5% level that 49.1% of the cities are mean-reverting to the synthetic counterfactual as compared to 22.6% to the linear trend. Given the balanced positive to negative impact within unit root cities, and as the synthetic counterfactual is likely to be less accurate for later cutoff years, as well as that most German cities were heavily bombed, we are more inclined to support the notion that the city share reverted back to a non-linear mean for most cities.

Now in Table 3.7, similar to the previous section, we exclude the cities in the donor pool for which the bombing deaths per capita is above the rate of 0.001 instead of 0.002. Here, we overall find slightly fewer cities where we conclude mean reversion, as well as divergences from the 50:50 balanced positive to negative impacted unit root cities. As argued in the previous section however, this exclusion may be too restrictive given that the 0.002 threshold is in line with the few German cities that are considered largely undamaged going through WWII. However, the results are more or less robust given the lower number of donor cities.

What can we conclude here? We find mean-reversion for the majority of city shares throughout all cutoff years, which is very much in line with previous findings that considered a shorter cutoff year in the 1960s (Davis & Weinstein, 2002; Brakman et al., 2004). We can reconcile these findings with the results of a unit root/negative structural breaks concluded by Bosker, Brakman, Garretsen, & Schramm (2008) that considered a later cutoff year in 1999. With later cutoff years, it opens up more so the possibility that non-linearities could have driven the results that leads to overacceptance of the unit root null. Secondly, trying to estimate the deterministic mean through non-linear methods would make interpretation of the results difficult to attribute to either non-linearities or the WWII shock. Arguably, our approach can better distinguish between these non-linearities and the long-term impact following WWII, where we find mean reversion of the city share for slightly fewer cities with a cutoff year of 1990 than when we consider a cutoff year of 1960, which in itself could have simply been driven by the increasing inaccuracies of the synthetic control method for later cutoff years. This also reconciles the results of the literature.

3.8 City Characteristics Conditional on Impact

In Table 3.8, we show some of the the pre-war characteristics of cities for which we find either a stationary series (below a 5% significance level), or otherwise a

¹⁵The definition was changed from Wohnbevölkerung (inhabitants living there) to Bevölkerung am Ort der Hauptwohnung (inhabitants at the place of the main address).

Table 3.7: Unit Root Tests on Relative City Size (Donor Pool: <0.001 bombing deaths p.c.)

A. Individual City (augmented) Dickey-Fuller test					
Period	1900-1990	1900-1980	1900-1970	1900-1960	1900-2000
a. Synthetic detrended ($c_i = 0$)					
Significance level	% Unit Root Rejected				
1%	22.6	30.2	35.8	43.4	
5%	52.8	56.6	69.8	71.7	
10%	60.3	66.0	81.1	73.6	
	Out of which % Unit Root with Positive Impact ($S_{i,t} > S_{i,t}^{synth}$)				
1%	41.5	37.8	47.0	40.0	
5%	40.0	34.8	25.0	46.7	
10%	33.3	27.8	40.0	42.9	
b. Conventional Method with Linear Trend (i.e. with trend term $\delta_i t$)					
Significance level	% Unit Root Rejected				
1%	9.4	5.7	5.7	3.8	13.2
5%	26.4	22.6	15.1	13.2	28.3
10%	37.7	22.6	30.2	32.1	45.3
B. Panel Unit root test					
a. Synthetic detrended ($c_i = 0$)					
	t-stat (p-val)	t-stat (p-val)	t-stat (p-val)	t-stat (p-val)	t-stat (p-val)
Levin-Lin-Chu	-11.846 (0.000)	-14.139 (0.000)	-16.579 (0.000)	-18.216 (0.000)	
b. Conventional Method with Linear Trend (i.e. with trend term $\delta_i t$)					
Levin-Lin-Chu	-15.619 (0.000)	-14.536 (0.000)	-14.717 (0.000)	-17.233 (0.000)	-16.074 (0.000)
Im-Pesaran-Shin	-4.800 (0.000)	-3.202 (0.001)	-3.012 (0.001)	-3.1590 (0.001)	-6.068 (0.000)

Note: The null hypothesis is a unit root in the relative city size detrended by the synthetic relative city size. The total number of cities is 53 for all specifications, and 138 in the donor pool which had below 0.001 deaths per capita during the Blitz. The synthetic cities are the first specification with the 1900 and 1920 to 1940 lagged population, 1920 number of coal workers, dummies for sea or river access, and the 1940 urban potential measure. The individual city unit root number of lags are chosen following Ng & Perron (1995), where the maximum lag length started at 2 for this procedure. Incidentally, the lag is chosen to be zero for all cities. $c_i = 0$ indicates a suppressed constant in the regression. $\delta_i t = 0$ or $\delta_i t \neq 0$ indicates an added trend in the regression. Given the individual city optimal number of lags, we set the lag to 0 for the panel unit root tests.

positive or a negative impact unit root of the city share with a cutoff year of 1980. Additionally, Figure 3.4 shows the spatial distribution of these cities on the map.

On average, cities for which we find mean reversion or a positive impact in the city share were larger before WWII than those negative impact unit root cities. Second, these negative impact cities seem to have been more remote from other larger cities than mean reverting or positive impact cities, as judged by the urban potential measures in 1939. Figure 3.4 for instance shows that rather remote cities from other cities in our sample to have been negatively impacted. Interestingly, the Ruhr area industrial heartland of Germany (mid-western part of the map), which is also the largest agglomeration of cities, mostly recovered or were positively impacted in 1980, with the exception of the city of Herne. We give the same conclusion to the Rhine-Main/Neckar region (the region to the south-west around Frankfurt down to Mannheim), with the exception of Ludwigshafen. Of note here in this region is the city of Heidelberg, which is the only city in this region for which we find a positive impact, as well as Regensburg to the south-east. We identified earlier that they have suffered the least deaths per capita during the bombings, and were regarded as largely undamaged. Also of note is that the rather remote city of Lübeck to the North that is also close to the East-German border, which is the only city in our sample to have more than 32.4% refugees in its total population in 1960; and this was probably a factor in finding a positive impact here.

Furthermore, positive impact cities have experienced on average significantly less wartime destruction, as measured by the rate of housing lost of total housing, as well as the cubic meter rubble per 1939 capita.¹⁶ However, there is no significant difference in wartime destruction between stationary and negative impact cities; the difference here is that stationary cities had a higher rate of refugees in 1960 as a percentage of population as compared to negatively impacted cities. This may also reflect refugees preferring larger cities or agglomerations to settle in.¹⁷ There is also no significant differences between being located in a coal region when it comes to the different impact.

Overall, it seems that there is little difference in wartime destruction between cities that were stationary, and those for that we find a permanent negative im-

¹⁶Both data is the same as in Brakman et al. (2004), taken from Kästner (1949).

¹⁷Note that initially after WWII, the French occupied zone restricted the number of refugees. They only received 60,000 at the end of 1947, around 1% of the total population, as compared to the total 4.379 million in the Soviet zone (around 24.3% of the total at that time), 2.957 million in the US zone (around 17.7%) and 3.320 million in the British zone (around 14.5%) occupied zones (Volkman, 1995), which overall more than compensated for the total German deaths during WWII. Yet, even though refugees were hindered initially to settle in the French occupied zone, we still find that the average rate is already 11% in 1960 for the 6 French occupied cities in our sample, Freiburg, Koblenz, Kaiserslautern, Ludwigshafen, Mainz and Trier. In fact, the other city not making it below 12.5% are Freiburg and Kaiserslautern with 12.8% and 12.6%. It seems not likely that differences in impact is driven by refugees not being allowed in one zone, but in others at least in the long-term.

pact. The discrepancy in whether a city recovers or not seems to be driven by other factors, such as the larger population size and agglomerations. However, the positive impact cities were also experiencing less wartime destruction, which is in line with what we would expect.

We suggest two drivers of divergences here: First, we may have simply accelerated the process of migration from smaller cities and agglomeration towards relatively bigger ones after WWII. The other one is that relatively less destroyed cities experienced a relative increase in the population. If we combine both situations, then we have on the one hand relatively less destroyed and larger cities and agglomerations that have gained in city share. On the other hand those that were relatively more destroyed, or rather as destroyed as the majority of stationary cities; and were smaller cities or agglomerations, have lost in relative city size. However, more sophisticated analysis would be required here to give a more definitive answer, and we suggest future avenues of research in this line.

Table 3.8: City Characteristics Conditional on Impact (Cutoff year 1980)

	Impact			t-test		
	Unit Root (pos.) i.e. $S_{i,t} > S_{i,t}^{synth}$	Stationary	Unit Root (neg.) i.e. $S_{i,t} < S_{i,t}^{synth}$	pos.-stat.	pos.-neg.	stat.-neg.
mean Pop (1939)	302525	228327	162651	0.8948	1.6302	1.1894
std. dev.	(275038)	(188814)	(140687)			
mean Urb. Pot. (1939)	484187	402582	284059	0.9259	2.2326	1.9209
std. dev.	(268524)	(208565)	(163805)			
mean Urb Pot. - Pop (1939)	181661	174256	121408	0.1966	1.5820	1.8440
std. dev.	(95761)	(94402)	(82263)			
mean Coal Worker (1920)	60986	72740	42893	-0.3360	0.5402	1.1451
std. dev.	(88225)	(87829)	(69903)			
mean refugee rate (1960)	0.182	0.165	0.161	0.9555	1.0617	0.3742
std. dev.	(0.062)	(0.039)	(0.036)			
mean rate of housing lost	0.392	0.428	0.392	-0.5641	0.0053	0.6674
std. dev.	(0.214)	(0.141)	(0.219)			
mean rubble p.c.	11.913	12.657	14.180	-0.2591	-0.4739	-0.5961
std. dev.	(10.612)	(6.124)	(11.085)			
Number of Cities	8	30	15			

Note: We consider 'stat.' a city where we find stationarity at a 5% significance level and a unit root city if the significance level is above 5%, while 'pos.' and 'neg.' respectively denote a positive and negative impact on the city population with a cutoff year of 1980, conditional on finding a unit root at a 5% significance level. The specification is equivalent to the one in Table 3.7 i.e. relative city size with a donor pool that consists of bombing deaths per capita below 0.001. The synthetic control was constructed through the 1900 and 1920 to 1940 lagged population, 1920 number of coal workers, dummies for sea or river access, and the 1940 urban potential measure covariates. 'Rubble p.c.' denotes the post-war cubic volume of rubble per 1939 capita in the given city.

Figure 3.4: City Location by Relative City Size Impact (Cutoff 1980)

Note: The total number of cities is 53 here, and 143 in the donor pool which had below 0.002 deaths per capita during the Blitz. The figure shows the results of the second specification (Table 3.7) on the relative city size i.e. relative city size with a donor pool that consists of bombing deaths per capita below 0.001.

3.9 Link between the Evidence and Theory

Although this paper is meant to primarily be a methodological contribution, we can say some things about the links between the evidence and the fundamental theories. The original purpose of the seminal paper by Davis & Weinstein (2002) was to evaluate three principal theories on the location of economic activity, namely locational fundamentals theory that considers fundamentals of a location, such as river or sea access, as the main driver; increasing returns theory following Krugman (1991) and random growth theories that have some form of Gibrat's law in its core to generate a Zipf distribution. One crucial element in that evaluation is whether cities are mean-reverting or not. In the following, we will have another look at the link between the theories and the empirical evidence. Then, based on the evidence we have found here, we will give an evaluation of these different theories.

First, it is useful to present the set of stylized facts we have found so far: (1) In the absence of a large, temporary shock, city population between 1920-1990 behaves in an inverse U-shaped way. (2) We find reversion to the mean of the synthetic city share of total population for most cities after WWII. (3) The ratio of positive to negative impact unit root cities is balanced for the city size and share. (4) Negative impact cities were smaller and more remote areas in 1939 than mean reverting or positive impact cities. (5) There is no statistically significant difference between the synthetic and the actual distribution, before or after WWII.

Locational Fundamentals

To start, we will first describe the typical characteristics and predictions of each mentioned theory, beginning with locational fundamentals theory (also often referred as first-nature geography in the literature), which explain the location of economic activity through the spatial distribution of geographic features. This is most prominent if we look at the typical birthplace of ancient civilizations which are often found along rivers, such as the Indus Valley Civilization along the Indus river, ancient China along the Yellow river, Egypt along the Nile river, or the Mesopotamia between the Tigris and Euphrates rivers. Even today, these regions are heavily populated; and through that it is earning locational fundamentals its name. One key prediction of locational fundamentals theory is that any large shock dissipates in effect in the long-run i.e. reversion to the mean. As our contribution here is that we consider a potential new mean defined by the synthetic control, being consistent with locational fundamentals theory in our case is equivalent to the reversion to the synthetic mean even after a large, temporary shock.

In section 5, we have found evidence in favor of inverse U-shaped city growth (1) for most cities if we look at the synthetic German city between 1900 to 1990.

Furthermore, the estimated turning point is on average around the year 1978. It is unlikely that with long spans in time series, the structure of the city or economy would stay constant; instead a non-linear trend seems more likely.

Together, it indicates that there are underlying non-random factors at play here, regardless of a large, temporary shock. These patterns can be explained by locational fundamentals being mutable throughout time, an idea which was already entertained by Michaels & Rauch (2017) and Bleakley & Lin (2012). For instance, motorization would change how people interact with the existing road and street network that made living in the city center relatively less attractive by reducing commuting costs (see Baum-Snow (2007); Baum-Snow et al. (2017)), similar to how steamships changed how people interacted with rivers and the sea that made living close to ports more attractive. Even further, the increase in motorization may have even changed the long-lasting fundamentals of roads within a city itself in a pressure to accommodate to the increasing motorization e.g. through wider streets. Similarly, the decline of coal and industry may be another explanation for mutable fundamentals. However, the issue of allowing for mutability is the added flexibility in explaining every possible observation, which would picking out the relevant theories rather difficult.¹⁸

Overall, we think that no change in the city size distribution after the WWII shock (5) is consistent with locational fundamentals theory, but not the inverse-U shaped one (1), unless we consider locational fundamentals as mutable that can accommodate this pattern.

The mean reversion stylized fact (2) is also consistent with locational fundamentals at least for the majority of cities. The extension of (2), in that the ratio of positive to negative unit root cities is balanced can be attributed to the limitations of our counterfactual approach, which produces a less accurate counterfactual the farther away we are from the treatment period. If the likely increase in this error is random, then it supports the case for us undercounting the mean-reverting cities estimated here (2). However, we cannot give a straight-forward explanation here why the cities for which we find a permanent, negative impact were smaller (4). This points towards population agglomerating towards larger cities over time, or rather, population going away from smaller cities in this case. Of course, we could let the flexibility of mutable fundamentals also explain this, but we do not have an immediately obvious explanation here for the overall bias against smaller cities.

¹⁸Same idea applies to how an overly flexible functional form fitting a trend would make it difficult to distinguish a unit root.

Increasing Returns

Next, increasing returns theory (Krugman, 1991) explain agglomeration by the interaction and proximity of people. Theories in this manner leave the possibility for a permanent effect from a large, temporary shock, a shift in equilibria, and a change in the shape of the city size distribution open. This also means that if we find none of those possibilities, we are still consistent with increasing returns theory. In other words, due to its flexibility in interpretation, it is not possible to falsify it.

Stylized fact (1) is not explainable by increasing returns theory alone. Rather, the opposite is the case in that one other prediction of increasing returns theory is the increasing concentration of economic activity over time or throughout history. Of course, we may not find (1) if we were to consider a metropolitan area definition instead of a city definition. This idea may suggest that increasing returns theory would be more relevant for metropolitan areas, while with smaller unit definitions, we would be more likely to find the effects of suburbanization that would explain the inverse-U shaped trajectory for most cities.

Second, the mean reversion findings for a majority of cities is still consistent with increasing returns theory, partly because it is not possible to falsify increasing returns even if we find mean reversion for all cities, but also because we do find unit roots for the relative size for some cities. It could be that strong lock-in effects that held a worse equilibrium were dismantled after the war, which brought the opportunity for a better equilibrium city size distribution to the detriment of smaller cities or less populated areas as in (4). The mean reverting majority of cities could then be seen to have already found their place in this better city size distribution equilibrium. Given the importance of interactions between humans that is core to these increasing returns theories, they provide a good explanation for the patterns seen in (4) of a decline of primarily less agglomerated places. We would tend to consider increasing returns to be the most straight-forward explanation for this pattern. Hence the mean-reversion stylized facts (2) and (3) are at least consistent with increasing returns theory, but we do not have an explanation for the distribution staying the same in (5).

Random Growth

Lastly, random growth theory has some form of Gibrat's law in its core, that consider urban growth to be independent of its size. The advantage of Gibrat's law is that it generates a Zipf distribution when combined with a reflexive lower bound. A prediction with cities following the same expected mean and variance is that the shape of the city size distribution remains the same given any temporary

shock, but that the rank of cities within the distribution is allowed to change. That even a large shock does not in any way alter the shape of the city size distribution (5) is consistent with random growth theories if the synthetic and actual cities growth with the same expected mean and variance. Next, the inverse-U shaped development in (1) would suggest an overacceptance of the unit root null, and thus, of Gibrat's law if the expected mean is not allowed to change accordingly. There is no natural and obvious explanation in random growth theory that would explain this non-linearity (1). Gabaix & Ioannides (2004) argued for a weakened version of Gibrat's law that would allow for possible non-linearity in the mean reverting component that could better simulate (1). However, the usual version of testing of Gibrat's law on the German dataset with more cities and a higher frequency of observation has been done already by Bosker, Brakman, Garretsen, & Schramm (2008), so that we refrain from doing this exercise here again. They have found that evidence for independence of size and growth, but reject Gibrat's law based on finding a change in the city size distribution.

Evaluation of the Theories Together

Given the added interpretational flexibility of the weakened version of Gibrat's law as well as mutable fundamentals, this makes fitting the results to the theory not so straightforward as we would have hoped. Instead, we have stated our tendencies on which historical context better plausibly fit these theories. We may summarize these tendencies together in the following:

The inverse-U shaped development (1) that we see can be explained for example via motorization and suburbanization that happened in most cities post-WWII. Random growth theories even with added flexibility of a non-linear mean reverting component does not offer a natural way to explain suburbanization. As argued above, it may better fit with mutable locational fundamentals in that roads became more important over time or by considering the deindustrialization wave of cities close to coal sources. It may also be the case that increasing returns theory would be more applicable for metropolitan areas, while suburbanization would apply to a more narrow border definition of a city.

We also find mean reversion of the city share for the majority of cities in (2) and (3). This again supports mutable locational fundamentals being the most plausible here, with increasing returns being at least consistent with the findings.

That the cities for which we find a permanent negative impact were smaller and more remote on average in 1939 and vice versa, positively impacted cities were larger and less remote (4) can be explained through increasing returns, where WWII could have acted as an accelerator of agglomeration. However, this is not consistent with the independence of growth, while (mutable) locational funda-

mentals do not offer a natural explanation for this. Hence, we tend to consider increasing returns as the most plausible explanation.

Where does this leave us? Similar to Davis & Weinstein (2002), mutable locational fundamentals seem to explain all stylized fact except for the decline of smaller cities, for which we tend to think of increasing returns as the plausible driver.

3.10 Conclusion

We began by outlining the potential problem that could arise from not separating the WWII shock from other underlying factors, such as suburbanization, that could show up to be non-linear. Not accounting for these non-linearities may lead to an overacceptance of the unit root, as well as potential misattribution of an effect between the WWII shock and these underlying factors. By creating a synthetic control city for each of the 53 German cities in our sample, we obtain a benchmark to which we compare the corresponding German cities after WWII, thus potentially controlling for underlying factors. We find that the synthetic city population is typically characterized by an inverse U-shaped development, explained by underlying non-random factors at play which are not driven by a large, temporary, negative shock.

First, we find that the majority of German cities recover to their supposed, synthetic mean in size and city share. We also find that the synthetic city exhibits an inverse-U shaped population trajectory during the 20th century. Of those cities where we find a negative impact after WWII, we find that they are on average smaller and situated in less agglomerated areas, which hints at the better ability of larger cities to recover. This can be rationalized by WWII accelerating migration away from these remote places that may have happened in any case. However, more research would be needed within this line of thought.

Together, we can reconcile the different results found in the literature, with our results suggesting a permanent effect being also driven by non-linearities combined with a larger timespan considered, that are more likely to make these non-linearities significant enough.

The results also tend to show the relevance of (mutable) locational fundamentals, for instance through motorization that led to lower commuting costs and suburbanization that explain the non-linear counterfactual population trajectory in the 20th century, with increasing returns effects better explaining the migration out of smaller and remote cities after the war.

3.A Data Appendix

Table 3.9: List of all sample cities

Aachen	Essen	Kassel	Oberhausen	Würzburg
Augsburg	Frankfurt am Main	Koblenz	Offenbach am Main	
Bamberg	Freiburg im Breisgau	Köln	Osnabrück	
Bielefeld	Gelsenkirchen	Krefeld	Pforzheim	
Bochum	Gladbeck	Ludwigshafen am Rhein	Recklinghausen	
Bonn	Hagen	Lübeck	Regensburg	
Bottrop	Hamm	Mainz	Remscheid	
Braunschweig	Hannover	Mannheim	Solingen	
Bremen	Heidelberg	Mönchengladbach	Stuttgart	
Darmstadt	Herne	Mülheim an der Ruhr	Trier	
Dortmund	Hildesheim	München	Ulm	
Duisburg	Kaiserslautern	Münster	Wiesbaden	
Düsseldorf	Karlsruhe	Nürnberg	Witten	

Table 3.9 shows the West German cities in our sample. As compared to the 62 cities in the sample in Bosker, Brakman, Garretsen, & Schramm (2008), we have excluded 7 cities based on not having complete population data before 1920 for each 10 years until 1871. The excluded cities are: Flensburg, Fürth, Kiel, Oldenburg, Wanne-Eickel, Wattenscheid, Wuppertal. We have also excluded Berlin and Hamburg due to the lack of cities that are as big as them pre-war in the donor pool.

Furthermore, we have excluded the following 18 cities, which we have also included (as well as the previously mentioned excluded cities) for the urban potential measure as there is census data for 1939 for all the 80 cities: Bergisch Gladbach, Bremerhaven, Erlangen, Göttingen, Heilbronn, Ingolstadt, Leverkusen, Mers, Neumünster, Neuss, Paderborn, Reutlingen, Saarbrücken, Salzgitter, Siegen-Wittgenstein, Wilhelmshaven, Wolfsburg and Worms.

Typically, we find city amalgamations, such as Wuppertal from Barmen, Elberfeld, Ronsdorf, Cronenberg and Vohwinkel, mostly in the 1920s and to a lesser degree in the 1930s. The immediate pre-war amalgamations were also typically significant, and a significant distortion more frequent than the annexations and fusion of the Gemeindereform of the late 1960s and the 1970s. For instance in the post-war wave for the 62 cities dataset, we find that for 48 out of 62 cities, the added population is below 20%, and 10% for 32 out of 62 cities.

In contrast, the pre-war amalgamations were typically more significant, where the added population is below 20% for 17 out of 62 cities, and below 10% for 9 out of 62 cities.

Also, we know very well how many people were added after an amalgamation, as we typically have records available. If this was not the case, we simply took the difference of the year before the annexation or fusion happened with the year

after, which is likely a very good estimate if the population growth around the annexation or fusion year was not too far away from 0. Typically, the 1920s and 1930s, and the 1960s and 1970s have a low growth rate as compared to the pre-WWI periods, or the immediate post-WWII period, so that this simple estimation seems appropriate enough.

With this information, we can adjust for the amalgamations with the simple adjustment as done by Bosker, Brakman, Garretsen, & Schramm (2008). That is, we take the city boundaries during WWII as the reference point. That means that pre-WWII, we add to the given city the population of the added areas. On the other hand in the post-WWII period, we deduct the added population from the given city. Furthermore, we add the simplifying assumption that the population growth rate of the original city and the added areas are the same. Keeping the notation of Bosker, Brakman, Garretsen, & Schramm (2008), we can more formally write for the pre-WWII case:

$$\hat{S}_{iT} = S_{iT} \frac{S_{iT-k}}{S_{iT} - S_{inew}} \quad (3.13)$$

with S_{iT} being the population of city i with the added population at time T , S_{inew} is the (estimated) population of the new areas which we know, S_{iT-k} is the population in year $T - k$ before the change of the city's boundary. Finally $\hat{S}_{iT}$ is the adjusted population with the population of the newly added areas being extrapolated in the past.

As for the post-WWII case Gemeindereform happening at time T , we write:

$$\hat{S}_{iT} = S_{iT} - S_{inew} \quad (3.14)$$

so that we simply deduct the population from the new area post-WWII at time T . For each subsequent years, as we do not know the (estimated) population of the new areas, we will extrapolate according to:

$$\hat{S}_{iT+k} = \hat{S}_{iT} \frac{S_{iT+k}}{S_{iT}} \quad (3.15)$$

The assumption that the growth rate of the usual surrounding, newly added areas being the same as the core, original city is implemented in equations 3.13 to 3.15. Lastly, we set the decennial data of British/Welsh that starts for instance in 1911 equal with German cities that typically starts in 1910.

Table 3.10: Unit Root Test on Log City Size (Donor Pool: <0.002 bombing deaths p.c.)

A. Individual City (augmented) Dickey-Fuller test					
Period	1900-1990	1900-1980	1900-1970	1900-1960	1900-2000
a. Synthetic detrended ($c_i = 0$)					
Significance level	% Unit Root Rejected				
1%	22.6	26.4	30.2	47.2	
5%	41.5	50.9	69.8	75.5	
10%	54.7	67.9	79.2	75.5	
Out of which % Unit Root with Positive Impact ($S_{i,t} > S_{i,t}^{synth}$)					
1%	48.8	48.7	48.6	35.7	
5%	48.4	42.3	37.5	46.2	
10%	45.8	35.3	36.3	46.2	
b. Conventional Method with Linear Trend (i.e. with trend term $\delta_i t$)					
Significance level	% Unit Root Rejected				
1%	5.7	3.8	7.5	3.8	7.5
5%	13.2	17.0	18.9	11.3	20.8
10%	28.3	26.4	34.0	22.6	32.1
B. Panel Unit root test					
a. Synthetic detrended ($c_i = 0$)					
	t-stat (p-val)	t-stat (p-val)	t-stat (p-val)	t-stat (p-val)	t-stat (p-val)
Levin-Lin-Chu	-10.934 (0.000)	-13.474 (0.000)	-16.709 (0.000)	-18.563 (0.000)	
b. Conventional Method with Linear Trend (i.e. with trend term $\delta_i t$)					
Levin-Lin-Chu	-26.995 (0.000)	-28.216 (0.000)	-24.732 (0.000)	-24.112 (0.000)	-26.995 (0.000)
Im-Pesaran-Shin	-7.066 (0.000)	-5.990 (0.000)	-4.828 (0.000)	-5.632 (0.000)	-7.066 (0.000)

Note: The null hypothesis is a unit root in log city size. The total number of cities is 53 for all specifications, and 143 in the donor pool which had below 0.002 deaths per capita during the Blitz. The synthetic detrend Dickey-Fuller test in 3a, as in equation (8), is estimated through the first specification with the 1900 and 1920 to 1940 lagged population, 1920 number of coal workers, dummies for sea or river access, and the 1940 urban potential measure. The MSPE was set between 1880 and 1940. The linear detrended Dickey-Fuller test in 3b follows equation (6) with trend term $\delta_i t$. The individual city unit root number of lags are chosen following Ng & Perron (1995), where the maximum lag length started at 2 for this procedure. Incidentally, the lag is chosen to be zero for all cities. $c_i = 0$ indicates a suppressed constant in the regression. $\delta_i t = 0$ or $\delta_i t \neq 0$ indicates an added trend in the regression. Given the individual city optimal number of lags, we set the lag to 0 for the panel unit root tests. Panel statistics are calculated with adjusting for cross-sectional dependence.

Chapter 4

From Bulwarks and Bombs to Boulevards[†]

Abstract: The dense medieval German city centers experienced a second destruction around the 1960s, which led to a widespread restructuring of the intra-city transportation infrastructure. We estimate the effect of the change in the density of streets and boulevards within West German cities on suburbanization. To deal with endogeneity, we instrument with existing boulevards that coincided around pre-19th century defortified walls and bulwarks, as well as the WWII destruction. From 1966 to 1986, in line with suburbanization, we estimate that the German city population declined by around 6 to 8 percent due to additional boulevards in the city center, while we find no statistically significant impact from a change in the street density. After this time period, we find no further impact, suggesting that the change in the intra-city transportation infrastructure had a permanent effect, but that it was temporary development.

Keywords: transportation, infrastructure, suburbanization, urban form

[†]This chapter is co-authored with Steven Brakman, Harry Garretsen and Tristan Kohl.

4.1 Introduction

In recent years, there had been a great interest within the empirical urban economics literature to estimate the impact of the transportation infrastructure on various economic outcomes, such as city growth, trade or suburbanization. One possible reason for this focus is the increasing urbanization rates across the world, which among other issues of urbanization also brought the issue of congestion to light, particularly within the cities of the developing economies.

Some papers looked at subways or urban rail network (Mayer & Trevien, 2017; Gonzalez-Navarro & Turner, 2018; Heblich et al., 2020). However, these networks are quite rare around the world, typically restricted towards a few of the largest metropolitan areas. The typical, medium sized city instead relies on a street network, and existing studies focused on the highway expansion after the emergence of cars (Baum-Snow, 2007; Garcia-López et al., 2015; Baum-Snow et al., 2017; Baum-Snow, 2020), which were however also meant for inter-city connections. We suspect that this focus is due to highway projects being easiest to realise outside the dense city to avoid the high cost of convincing people to relocate elsewhere, as well as the cost of razing entire housing blocks. Thus, the dense medieval European city center may be stuck in a street network which was initially built for a few pedestrians and hand-carriages, but not with more contemporary needs in mind. This also means that there is little variation in within-city street infrastructure for researchers to exploit across time. For city planners in emerging economies that may decide on whether to raze housing blocks to modernize its within city infrastructure before further housing development may take place, having a measure of the causal impact of such modernization may help in this politically difficult decision.

German city planners saw the post-WWII period as a unique opportunity to reimagine the medieval city in light of the increasing popularity of cars. We will study the impact of this change in the street network within a city and its impact on city access and suburbanization. For this we have constructed a unique dataset of 74 German cities by digitizing topographic maps, which were purchased in 1933/34 by the Allied forces, and which we will use to compare the street network right before WWII with the network of today.

We diverge from the aforementioned papers which exploit the number of highway rays emanating from the city (Baum-Snow, 2007, 2010; Garcia-López et al., 2015; Baum-Snow et al., 2017; Baum-Snow, 2020) by measuring the length of added boulevards after WWII, as well as the change in the street density of a city. We consider this approach more suitable as a measure for within city access as there is likely to be little change in rays over time within the German city center.

As changes in streets and boulevards are both possibly endogeneous to future

urban growth, for example because transportation infrastructure investments are higher in cities that are expected to grow faster, we will instrument both: First, we follow the literature in exploiting historical data through a 'boulevardized walls' (BW) measure; boulevards were often created around defortified walls and bulwarks, as defortification provided early space within the dense medieval city for infrastructure, hence easing demands for such boulevards. Second and similar to the instrumental strategy approach by Davis & Weinstein (2002) and Brakman et al. (2004), we will use the WWII destruction after the Allied bombings on German cities as another instrument. Some cities never had walls, or defortified very early before the need for boulevards became apparent. The WWII destructions provided another opportunity for cities to that are stuck within the medieval layout to modernize. This culminated in a *Flächensanierung* or 'second destruction' of areas within the city, starting around the 1960s.

To our knowledge, this paper is the first which investigates the causal impact of a change in the street network within the city, while making a distinction between normal streets and wider boulevards. We find an overall impact of around 6 to 8% decline in the population of a city between 1966 to 1986 due to the boulevard increase that initially started in the 1960s, and that this effect is temporary. This development is however concentrated mostly around the mid 1970s. We do not find a significant effect of normal streets, which may suggest that the commuting cost reduction that led to this development is primarily driven by the boulevard increase.

A similar paper on cities before and after major transportation innovation is Burchfield et al. (2006), who looked at cross-city differences in car-unfriendliness via a street-car public transportation usage proxy showing a city center less friendly to cars. They found that cities that had a higher street-car usage tended to be more compact i.e. cities that were more car-friendly were more sprawled. Rauch & Brandily (2020) looked at cross-city differences in the street density of Sub-Saharan African cities, finding a higher growth rate for cities with higher street density. However, these cities are typically very young, as compared to much older European cities. We differ from these papers in looking not at the cross-city difference, but the change in the inner street network.

Apart from the aforementioned intra-city infrastructure literature, this paper is related to the wider transport cost literature, for instance on the effect of inter-city roads or highways on the growth of cities (Duranton & Turner, 2012), on the economic activity (Chandra & Thompson, 2000; Storeygard, 2016; Gibbons et al., 2019) or trade (Duranton et al., 2014). This paper is also related to the literature on the inter-city impact of railroads (Donaldson & Hornbeck, 2016; Jedwab et al., 2017; Donaldson, 2018).

In the following section 4.2, we will discuss the context during defortification and around WWII, which is particularly important for the validity of the instruments we employ. In particular, we outline the political process of defortification 4.2.1, and discuss how different cities may have different transportation needs over time in section 4.2.2. Section 4.2.3 discusses the time around the 'second destruction' and the major infrastructure projects. Section 4.3 discusses our empirical strategy as well as the data. Section 4.4 provides the baseline results, while section 4.5 provides some additional robustness exercises. Section 4.6 concludes.

4.2 Context

We are looking at the post-war *Flächensanierung* infrastructure change in Germany that renovated the old, often medieval infrastructure to accommodate modern needs, translating into lower commuting or transportation cost from and to the city center.

Here we can consider the standard land use model of the internal city location of workers and firms, where firms are assumed to be located in a single central business district (CBD), and workers can choose between living close to the CBD to reduce commuting while paying more for housing and vice versa.¹ The reduction in commuting costs due to the infrastructure modernization would lead a fraction of workers to live further away from the CBD. Hence one implication of the standard model is as we have a reduction in the commuting cost, whether it is through infrastructure or technology, the share of the population around the CBD falls, suggesting suburbanization (see Duranton & Puga (2014) for an overview of these type of models).

One issue empirically is that we typically have to deal with infrastructure projects to be endogenous to (expected) growth. In the following, we will suggest and discuss two sources of instruments, WWII destruction that made post-WWII infrastructure changes easier, as well as defortification of walls that led to initial boulevards in the dense medieval city center. Also given that such infrastructure changed little over time, these events coincide with where a significant change in the inner city infrastructure could have even taken place, allowing us to exploit this larger variation.

First, most medieval cities either had thin walls (so called *enceinte*) or more sophisticated bulwarks. After defortification, boulevards were often build on the area of those razed walls or bulwarks, as these areas are the closest buildable and

¹The Alonso-Mills-Muth framework (Alonso et al., 1964; Mills, 1967; Muth, 1969) is one such example.

free space to the city center.² Hence, one advantage of considering walls is that they are naturally situated around the historical European city center that is typically still the center of commerce today.

For most medieval European cities, the wall area offered a politically and economically feasible way to create boulevards, as the alternative would be to raze entire housing blocks throughout the inner city. As such, this alternative rarely occurred in the absence of outside forces or absolutist monarch in power.³

Instead, the boulevards of the typical walled city were ring-shaped around the medieval city center, known as 'der Stadtring' in many German cities today.⁴

We will use these boulevards that were built on defortified walls as an instrument that corresponds to less demand for an increase in boulevards after WWII, which we in turn suspect to be endogenous to expected urban growth. For one, this instrument is similar to the historical transportation network approach that instruments more modern networks (see Redding & Turner (2015) for an overview on the road network). Typically to establish exogeneity, these type of instruments rely on the time span between the creation of these networks, combined with whether fundamental economic changes took place during that time span. In our sample, most of the defortification took place before the foundation of the German Empire in 1871 (see Figure 4.6), but mostly around the beginning of the 19th century.⁵ Second, innovations in intra-city transport, such as the horse-drawn and electric omnibus, as well as widespread urbanization and industrialization after most of defortification took place could be regarded as fundamental enough to how the city functioned. Lastly, we may be worried that the choice to defortify itself is correlated with post-WWII urban economic outcomes. Although the time span between defortification and WWII is typically large, in section 4.2.1, we will in more detail outline the circumstances that surrounded the decision to defortify. In short, the decisions were more driven by city politics of the pre-German Empire and likely to not be correlated with urban economic outcomes today, which together with the already large time span further strengthens the case of exogeneity.

²Notice here that the French 'boulevard' was derived from the Dutch 'bulwarc', signifying the scarcity of such wide streets before widespread defortification took place. Similarly, esplanades were originally areas that were to be kept empty to provide a free field of shooting, and were also used to build boulevards on.

³The famous example is the Haussmann renovation of Paris in the 19th and 20th century, which saw the transformation of the dense Parisian streets towards the Paris of today with its star-shaped Boulevards.

⁴Baum-Snow et al. (2017) also considered 'ring roads', with the difference that these roads were built a few kilometers outside the city center where space is abundant. The ring road in European cities are typically within the 500m to 1000m radius of the townhall, with the rare situation of walls going past the 1500m radius often found for more modern fortifications e.g. Bundesfestung of Mainz and Ulm.

⁵Empire foundation dates were important in defortification history, as it meant that defense were now shifted from the individual city towards the border of the Empire. In this regard, countries deep inside the borders became relatively unimportant.

Second, the WWII destruction is used to instrument both the change in street and increase in boulevard density after WWII. For one, rebuilding of the cities happened right after WWII, but it took around the 1960s for initial boulevards to be completed; also due to city planners needing to convince property owners first to give up their land. Section 4.2.3 discusses the historical context of this second destruction and *Flächensanierung* ("area renovation"). We would think that it would be more difficult to build boulevards as compared to normal streets, as boulevards typically need to be longer, which would mean more property owners to convince in a way that would not lead to too much curve fitting. On the other hand, changes in the street density may be less difficult as it could be done more locally, such that changes could have been made before the 1960s. Unfortunately we can only speculate on when these changes happened, as we do not have any topographic maps right after WWII, and commercially available maps are not detailed and accurate enough. However, if we were to find an impact of boulevards, it should be around and after the 1960s as we will highlight in section 4.2.3.

The validity of the WWII destruction instruments rely on the same arguments as Brakman et al. (2004), with the destruction largely to not be driven by the economic importance of the city, which may have an effect on future growth rates. To summarize their arguments, 1940 to early 1942 targets were mostly industrial areas, with limited success in disrupting production. This led to area bombing to be employed as the new strategy, targeting not only industrialized areas, but also historical city centers. In particular, the goal was to destroy the housing stock to disrupt overall production through the dislocation of the work force, as well as to lower morale. Another criteria for the bombing choice was the visibility from the air (weather conditions, landmarks) or susceptibility to incendiary bombs (Friedrich, 2002). Distance to Britain was also a criteria as deep bombing runs were initially not possible for most aircrafts. That is, for most of the WWII bombings, the target itself was arbitrarily chosen.

Stylistically, we can say that these two instruments cover the significant street infrastructure changes within medieval European city: Defortified walls created the first opportunity in the dense medieval city to create boulevards for more modern transportation needs of the 19th century. If this was not done, maybe because the defortification happened too early, then the period after the WWII destruction offered another opportunity for street infrastructure modernization to take place.

Lastly, section 4.2.2 will discuss in which way the post-WWII transportation infrastructure adjustments run. In particular, we may think of some cities as having too few streets before WWII, and some very old cities as having too many streets, but too few boulevards. This may not be relevant to the estimation results on the impact of the transportation infrastructure, but it is relevant if we want to know

the different impact of the second destruction on different cities.

4.2.1 Decision to Defortify and Boulevardize

Beginning with the time around defortification, for our instrumental variable strategy to be valid, it requires that the decision whether to boulevardize the walls should not be driven by underlying economic factors that still play a role today and which influences post-WWII city growth. First, our instrument is similar to the identifying assumption by other historical road instrument strategy approaches in the literature (Redding & Turner, 2015), where going further back in time would lessen the likelihood of significant correlated unobservables.

In the following, from the perspective of historians we further show that the decision to defortify were primarily not driven by economic decisions that may be more likely to be correlated with modern city growth. Mintzker (2012) presents three classical explanation for defortification, where walls represented an obstacle for expansion due to the increasing urbanization, an increasing need for transportation between cities and countryside due to industrialization, and the obsolence of the usefulness of walls and bulwarks due to military innovations, such as improvements in artillery. Mintzker dismisses both the industrialization and expansion explanation as the primary cause for defortification. First, early industrialization took place after most of defortification of Western and Central European cities already happened. In our sample, this is also the case as most of defortification happened or started before very early industrialization in the 1830s (see Figure 4.6).

Second, Mintzker argued that expansion was not a primary driver of defortification in the 18th century, and contrary to our expectations, in the 19th century the biggest and most rapidly expanding central European cities retained until much later or even strengthened their walls.

The military explanation is too simplistic, as it ignores the other functions of walls, as well as the risks involved. For example, one reason for the defortification of the wall can primarily be found in the Napoleonic Wars, where razing a city meant the demolition of its defenses to turn it back into a village, and not the levelling of the city. Cities such as Bremen, Hamburg and Lübeck preemptively defortified, as even though their walls were very sophisticated bulwarks, useful even against the modern weapons at that time, they also attracted the much larger state armies of the Napoleonic wars, which could had resulted in long sieges and a direct bombardment of the city itself. Hence, a defortification wave occurred at the beginning of the 19th century, driven not by the need to expand the city border, but by political considerations. In Figure 4.6, we can see such a concentration

around 1800 with very different outcomes of whether they boulevardized the walls or used it for other means.

Similarly, in the 18th century before the Napoleonic wars, it was primarily the conflict between the republican city to retain its independence and the ability to defend itself, as symbolized by the walls, and the absolutist ruler to integrate the city and its people into their realm as their subjects. We would then find fortress cities at the border, and defortification of walls deep within the state border, as city walls did not serve the city anymore but the state to defend against the enemies outside the state. This is also pronounced in the German unification wars of the 1860s, where the regions to defend shifted towards the border, thus making the remaining 'internal' walls and bulwarks strategically unimportant, while making the 'border' fortresses relatively more important.

Apart from these political conflicts between the citizens of the city ('burgher') and the state, there were corporations (guild, town etc.) and the non-citizen ('suburbanites') with their own interest. In Munich for example, walls were used to collect specific taxes and tariffs from both citizens and non-citizens by corporations, and were still valuable against smaller and nocturnal threats, such as bandits. Non-citizens were also barred from working in the city, which enforced the control of guilds over manufacturing and labor markets within the city walls. The enforcement of these differences between the city and the suburbs, as well as within the city, would effectively be abandoned with the defortification of the walls. The analysis by Mintzker (2012) indicates that there seems to be common, single explanation which forced cities to defortify at one point in time, such as the expansion of cities, the early industrialization or military innovations rendering walls useless. Instead, defortification was a rather political process with many actors benefitting from or being disadvantaged by these walls.

Overall, it reassures us about our instrumental variable strategy: Even if later defortifying cities turn more of their walls into boulevards due to an increasing need for transportation; for the instrument to be valid, we would need the decision to defortify by the decision makers back then to not be driven by underlying factors that could still drive post-WWII central city growth rates. The large time span between defortification and WWII in that these 18th and early 19th century planners likely did not anticipate the emergence of cars, also work in favor of the validity of the boulevardization of walls instrument.

4.2.2 Ideal Infrastructure over Time

Rauch & Brandily (2020) expected that cities that were founded after the adoption of cars would better accommodate cars, which would represent itself in wider streets

and a higher street density. However, already before cars were common place, there was an effort to have right-angled and wider streets away from the organic street structure set in place during the Middle Ages. The motivation here however was to get away from the organically developed and dense housing blocks where fire can easily spread. For example, after the fires in 1512, 1763 and finally a complete destruction in 1817, the city of Regensburg through King Max I of Bavaria was completely redesigned in 1820 with wider, angular streets. It became known as the Modellstadt Regensburg ('model city'), as cities that were similarly destroyed were to follow the example of Regensburg in its reconstruction. Similarly, the Great Fire of London in 1666 saw reconstruction with wider streets as well, driven by fire prevention as well as the use of carriages:

"Almost inevitably, the fire was seen as providing an opportunity to replan as well as rebuild the City so that it would be 'finer than it ever was' [...] The necessity of broadening at least some of the principal streets was also unquestioned. Graunt believed that the increased use of coaches, for which the narrow streets of the city were unsuited, was one reason for the drift of business westwards." (Porter (1998) pp 94-95)

We would think that wider streets, but contrary to Rauch & Brandily (2020), a lower housing density and along with it, a lower street density were also desirable traits. There may be two opposing factors now: First, the transportation improvement effect of a higher street density, and the modernization effect of a lower housing and street density.

To see which cities in terms of its street network may be stuck in the past, we could look at the different foundation years of a city as a proxy for the age of the street network. Here, instead of the age of a city, we think that the year a city received its city rights can better suggest whether a city is stuck in the past or not in terms of street organisation. This is because with the exception of Karlsruhe and Ludwigshafen, the cities in our sample were founded about a millenia ago, while many more cities were issued city rights relatively recently.⁶ However, as places typically need to be of a certain size to afford city rights, those cities that received its rights during the 18th-20th century can be characterized as cities that were small, but developed quickly relatively recently. This recently laid down street organization would hence be closer in line with the more modern needs, where the street and boulevard density for these cities can be implied as optimal for its time, and arguably closer to optimal today than the street organization that was laid down during the Middle Ages. As such, we can consider them as 'young' cities, in contrast to the 'old' cities that received the rights during the Middle Ages.

⁶Karlsruhe was first mentioned in 1715, and Ludwigshafen in 1824.

The data for the following graphs on boulevard and street density will be described in section 4.3.1. Figure 4.1 shows the relationship of the 1934 street density and city rights. In general, further away from the town hall, it is more likely that there are areas that are not yet occupied, which may be an issue particularly for smaller, fast developing cities i.e. cities that received the city rights later. Hence, the most relevant range is the within 500m range of the town center, as it typically encloses the medieval city center as well, thus making cross-city comparison the most appropriate for this narrow range. We find a negative relationship between year of city rights received and street density, in particular for the within 500m radius, suggesting a lower street density as more desirable for young cities.

As in the previous section, we may think that an increasing need for boulevards would drive the decision to defortify in order to free up the space for boulevards. However, we find a weak relationship in Figure 4.2 for 1934 boulevard density and city rights year. The little difference in the boulevard density between young and old cities indicates that around the time of the expansion of the young city, boulevards did not become more important. In other words, old cities may not be stuck yet with too little boulevards.

However, around WWII, cars became more important as the mode of transportation. Figure 4.3 shows the 1934 street density and the 1934 to 2020 *change* in street density, where we find street densities converging across cities: Cities with an initially low street density level experienced an increase in street density between 1934 and today, while those with a high street density level experienced a decrease. Figure 4.4 shows the 1934 boulevard density and the 1934 to 2020 *increase* in boulevard density. Cities with an initially lower boulevard density experienced a larger increase in boulevards.

Hence, it suggests an ideal towards a lower street density from the Middle Age to approximately the end of the Late Modern Period (around 1900), while we do not see such a clear pattern for boulevards. As for the time after WWII, the patterns suggest some convergence in the street density.

4.2.3 Second Destruction and *Flächensanierung*

The construction of additional boulevards did not happen immediately after the end of WWII, as there was a shortage of housing so that the general focus was on the quick supply of shelter and basic services within the city, and not on further tearing down the city. Apart from that, it would also take some time for city planners to convince enough owners of property to sacrifice their property for the greater good. We cannot determine a precise year on when this second destruction and the subsequent reimagination of the infrastructure initially took place, as there

Figure 4.1: Ideal Street Density: Middle Age to Late Modern Period

Figure 4.2: Ideal Boulevard Density: Middle Age to Late Modern Period

Figure 4.3: Ideal Street Density: around WWII

was no country-wide initiative or policy that could serve as a starting point which would apply for every city in our sample. However the urban planning history literature can provide us some general trends. First, Hanover at the end of 1959 was stylized as the prototype car-friendly city, with its post-war development being called 'the wonder of Hanover' (Spiegel, 1959). Bernhardt (2017) posits that the exchange program during the 1950s of continental European planners to the United States established the automotive city as the guiding principle for spatial planning for the years to come. Following Zhang (2011) the leitmotif of adapting the city to the increasing levels of motorization continued until the beginning of the 1970s, with the 1970s to the beginning of the 1980s seeing a shift towards preservation. This development was symbolized by the increase of preservation programs among renewal projects from 5% to 30-50% in West-Berlin (Tessin et al. (1983), pp. 23), the foundation of the Deutsches Nationalkomitee für Denkmalschutz (cultural heritage preservation) in 1973, the European cultural heritage preservation year in 1975, and also driven out of necessity due to the oil crisis of the 1970s, such that preservation policies were established in most cities at the beginning of the 1980s (Zhang (2011) pp. 108).⁷

We therefore consider the 1960s as the beginning of the second destruction and *Flächensanierung*, and the 1970s as the beginning of the end of it. This assumption is crucial as we will compare the topographic maps of 1934 to the maps of today, which are the only available post-war maps that we consider to be comparable to the pre-war maps in terms of accuracy. That the street network has not changed much during the 1970s and the beginning of the 1980s justifies the use of today's maps, where we can then narrow down the changes in the street network, as seen between the 1934 topographic maps and the maps of today, to this narrower period around the 1960s to 1970s.

We should then expect to find any impact of a change in the street network only during and after this period, but not before.⁸

4.3 Empirical Strategy

As noted by Redding & Turner (2015), the literature that empirically adopts land use models on the infrastructure impact on suburbanization, whether it is about inter or intra-city infrastructure, follows more or less a very similar reduced form estimation. It typically starts with a simple equation of the outcome of interest regressed on the measure for the transportation infrastructure. Equation (1) follows

⁷Another symbol was the motto of the Deutsche Städtetag (the German association of cities) in 1971, 'Rettet unsere Städte jetzt' (rescue our cities now).

⁸In terms of the period length, it is similar to the 20 year period considered by (Baum-Snow et al., 2017).

this ubiquitous reduced-form estimating equation, where we take the the long-difference similar to the literature (Baum-Snow, 2007; Garcia-López et al., 2015; Baum-Snow et al., 2017), that is

$$\log population_{i,t} = \beta_0 + \beta_1 boulevard_{i,t} + \beta_2 street_{i,t} + \beta_3 \log population_{S,t} + X\zeta + \mu + \sigma t + \epsilon_{i,t} \quad (4.1)$$

where $\log population_{i,t}$ is the log population of city i at time t , $\log population_{S,t}$ the log population of state S , the boulevard density $boulevard_{i,t}$ and the street density $street_{i,t}$. X contains the vector of controls and μ is the time-invariant unobservable. As we compare the pre-WWII city map to the post-WWII city map of today, and therefore have no annual time variation i.e. do not observe the evolution of the change in the street infrastructure. We will consider the long-difference estimator typical to this literature that deals with historical data. $\#T$ long differencing equation (2) leads to:

$$\Delta_T \log population_i = \sigma T + \beta_1 \Delta boulevard_i + \beta_2 \Delta street_i + \beta_3 \Delta_T \log population_S + X \Delta \zeta + \eta_i \quad (4.2)$$

which drops the time-invariant unobservable μ . Here, $\eta_i = \epsilon_{i,t} - \epsilon_{i,t-T}$ and Δ_T are the $\#T$ years long difference operator. As we only measure the 1934 and today difference of streets and boulevards, and other measures derived from the street length, we drop the T in the difference operator for these variables for the sake of clarity. Lastly, for ease of interpretation in order to display the annual growth rates, we will divide both city and state population by $T - 1$. We consider both street and boulevard measures to be endogeneous. For instance, growth may be endogenous with an increase in the demand for commuting. Hence, we will instrument through the WWII destruction rate of a city i , as well as the 'boulevardized walls' (BW) density measure. We expect the BW measure to be a relatively better negative predictor of post-WWII changes in boulevards, and less so of changes in street density, while the opposite would be the case for street density. The reason is that boulevards are typically much longer than individual streets, so that a longer stretch of housing blocks need to be deconstructed. Even if the houses were destroyed, the city planner still needed to convince owners to give up this stretch of land, unless there are enough boulevards already.

As infrastructure improvements may both have an effect on suburbanization of the population as well as draw overall population in due to its attractiveness, we control with the change $\Delta_T \log population_{S,t}$, and as it itself may be endogeneous to the outcome, we control this with its level $\log population_{S,t}$. This approach is similar to the MSA controls of Garcia-López et al. (2015); Baum-Snow et al.

(2017) in that the estimated coefficients of $street_{i,t}$ and $boulevard_{i,t}$ can then be attributed to suburbanization.

Second, given section 4.2 and 4.2.1 that the WWII bombings as well as de-fortification is not concentrated by states, the aggregate impact of this change in the inner city infrastructure on the attractiveness of the whole state itself should as well be averaged out. Nevertheless, in a robustness exercise in Table 4.6, and similar to Baum-Snow et al. (2017), we opt to exclude $logpopulation_{S,t}$ to give a bounds on the true value coefficients of $street_{i,t}$ and $boulevard_{i,t}$. As it turns out, the estimates are almost identical whether we include it or not.

Third, X are also controls that may be correlated with both our instruments and growth at the individual city level. We consider geographical characteristics that may hinder suburbanization e.g. (great) rivers or physical possible extent as measured today.

Given the above model in that transportation improvements reduces central city population density, we would hence expect to estimate $\beta_1, \beta_2 < 0$. However, note that this is the indirect way to find evidence for suburbanization due to infrastructure changes, while for the stronger case for suburbanization, we should also expect to find the opposite signs for the unobserved suburbs as well. That is, we may replace the dependent variable of equation (3) with suburban population, for which there is no sufficient data available. Lastly, given that wider streets are probably better for commuting, we can also expect that $\beta_1 < \beta_2$ i.e. that the street density impact on population is smaller than the boulevard density impact.

4.3.1 Data

We use the so-called 'Meßtischblatt' topographic maps for the pre-WWII road network data, which were purchased by the Allied forces around 1933/34 and which were used during WWII.⁹ We overall restrict our analysis to the 1500 radius of the town hall, as for most of the cities in our sample, the extent of the city can be roughly found around this radius, except for the largest cities. Distance to the town hall itself also seems to be a good proxy for the proximity to the central business district (Ahlfeldt et al., 2020).

The Meßtischblatt maps are in general very accurate if we compare them to the maps of today, which allows us to see whether segments of the boulevards that exists today were also present back when these maps were created. We consider

⁹In the following, whenever we mention the year 1934, we refer to the these maps as if these maps were done in 1934 out of convenience. However, the maps were merely purchased around 1933/34, so that the maps record the street network slightly before that date. We think that this is not an issue, as little seems to have changed within the city before WWII, as we have cross-checked them with similar topographic maps from the year 1877.

streets as boulevards if they classify as at least 'highways' in the ESRI World Street Map.

We use depictions of walls that were then georeferenced on modern maps. As these depictions come from a variety of historical maps, they vary more in quality as compared to the *Meßtischblatt* topographic maps. Figure 4.8 compares the street network of 2020 to the one in 1934 Ulm. Figure 4.9 shows the 2020 boulevard network in 2020 compared to the one in 1934, as well as the part of the boulevards in 1934 that coincide with the historical walls. Figure 4.10 depicts the enceinte (inner ring walls) and bulwarks (outer ring, thicker walls) of the 'Bundesfestung' Ulm.

We have digitized 2020 boulevards as defined above, and digitized 1934 boulevards by comparing the width of individual boulevard segments between them. If one segment of the 1934 boulevard is significantly less wide compared to the same 2020 boulevard segment, then we do not assign the 1934 boulevards to this segment.¹⁰ The appendix A presents the digitization process of our boulevard measures in more detail. As a consequence of this approach, the total length of 2020 boulevards is greater or equal to the 1934 boulevard length. Also this digitization approach does take into account the possibility that boulevards which existed in 1934 would not exist anymore today.

We consider the same German municipal population data as Bosker, Brakman, Garretsen, & Schramm (2008), that had a population of at least 50,000 between 1939 and 2000.¹¹ Of those, we have excluded Munich and Heilbronn from the sample due to the *Meßtischblatt* topographic maps not being retrievable. We have excluded Salzgitter, as it was formed as a collection of disjoint villages, and Wolfsburg, which was formed in 1938 to house Volkswagen workers. Lastly, we have excluded West-Berlin as it can be considered an island for most of the period of interest, so there was little room for suburbanization. Thus in total, this leads to 74 West German cities within our sample.

Note that here was a change in how the population in a city was counted in 1987, which is enough to detect a significant break in 1987, and which we cannot adjust for properly.¹²

Also, the population data is measured on the basis of city-municipality, defined by the area around WWII. A wave of amalgamations of municipalities occurred in the 70s, which however was negligible in terms of added population, as compared to

¹⁰We have defined significance if the 2020 boulevard segment is at least 10% wider than the same segment in the *Messtischblatt*. Taking a typical width of the boulevard between 20 to 40 meters, a significant change would be at least 2 to 4 meters respectively.

¹¹There were some typos that were fixed, so the dataset is very similar but not exactly the same.

¹²The definition was changed from *Wohnbevölkerung* (inhabitants living there) to *Bevölkerung am Ort der Hauptwohnung* (inhabitants at the place of the main address).

the major wave that occurred in the 20s and 30s. We correct for it through the same method as Bosker, Brakman, Garretsen, & Schramm (2008) by assuming that the growth rate of these newly added municipalities have the same growth rate as the one it was added to. We describe this procedure in more detail in the appendix B on amalgamations. Note here that this correction is most appropriate right after the amalgamation, and less accurate the further away from this date. 'Bundesländer' state level population data are taken from various sources as compiled by the Histat historical statistics of the Leibniz Institute (Sensch, 2007), which provide data for 1946 to 1999. We will then restrict our analysis to 1946 onwards and consider the years before 1987 to be the least problematic. Individual city destruction data during WWII are the same as in Brakman et al. (2004), taken from Kästner (1949).

We include a river dummy as simple first-nature determinants, where we consider a city a 'river' city if today's municipal border is at most 1 kilometer away from the nearest river. We will also consider a 'great' river measure of rivers that run through the city within 1500m of the town hall and is over 100m wide in this city (Rhine, Danube, Weser, Elbe, Main).¹³ As medieval city center are typically build close to the river, the extent of the city will be disrupted by these great rivers, which manifests itself in the other side of the river being typically less populated and which may limit the growth of a city. We can see these patterns with the more populated left side of the Rhine river, which was first settled by the Romans.

We will consider a proxy for the maximum extent of a city, which will be derived as a by-product of the unevenness measure, which we will describe below. In essence, it measures the number of points in a city with a maximum distance to the nearest street today. As streets cannot be build on sea, or mountains, it serves as a proxy for the maximum buildable area within a 1500m radius.¹⁴ Lastly, we consider cities with a port close to the sea (Hamburg, Kiel, Flensburg, Lübeck) as well as control for whether the city is a 'Bundesländer' state capital to control for divergent growth rates from other cities. Given that younger cities tend to have infrastructure that are more accommodating to modern needs, we have city rights data to determine the effective age of the city, as these rights take into account that the city was already a city at the time the city rights were awarded, and were not merely founded.

¹³The Neckar river may be considered a great river around Heidelberg with a wideness of 100m, but not in Reutlingen with a wideness of around 20m.

¹⁴Similar to Harari (2020) who consider the impact of the shape of cities, we consider the shape of the inner city within a specific radius. We may hence think of the shape of cities as 'outside' shape, while our measure considers the 'inside' shape.

4.3.2 Other Controls: Unevenness and Orientation

We may think about other forms of measures that benefits transportation within the city center. Following Donaldson & Hornbeck (2016) and Rauch & Brandily (2020), we measure the unevenness of the city grid, that is, we measure the distance from a random point to the nearest street and take the average of it.

Here, we consider points in a squared grid formation as random. For each of our circle range radii 500m, 1000m and 1500m, we then have at most 16, 80 and 180 points respectively. The intuition is that more spatially evenly distributed streets within a circle decreases this average.

Note here that we do not consider points that are too far away from the nearest street, which is often the case for cities with a wide river, at the sea or surrounded by mountains.¹⁵ Otherwise, it would artificially increase the unevenness measure as we would otherwise consider points that are outside a city's natural geographic border. That is, with this simple measure, a highly uneven but dense street network may be considered as highly even. To give an example: Typically the medieval city center exhibits the most dense street network, and the denser it is, everything else equal, the less uneven we would measure the overall city even if 'inequality' in street density between the the center and outside of it would be higher. In effect, we may have instead measured the inverse of a street density proxy, and not the uneven access to the nearest streets across the whole city.

That is, we may think about unevenness as a function that is negatively affected by street density. In particular, for newly created neighborhoods, changes in unevenness and changes in street density would move in the opposite direction. To deal with potential multicollinearity, we consider the standard deviation of this measure instead of the average, as the deviation is closer to an actual measure of unevenness. From this measure of random points, we can also control for the change in the extent of a city by counting the number of these points, before and after WWII. Most cities in our sample are close to 180 points, with few notable exceptions such as Heidelberg, which is surrounded by a mountain and a great river. We can then measure the city extension of the city boundary within an overall range, and the change of the extension over time, as well as the above described maximum buildable area that is restricted by geography.

Lastly, following the indexation by Rauch & Brandily (2020) of the orientation measure by (Boeing, 2019), we measure the orientation or alignment of the grid through the degree orientation of streets in the grid, and take this to create an orientation Herfindahl index. Overall, we consider 18 bins, each representing 10 out of 180 degrees of street orientations. Considering

¹⁵We have considered too far away as the distance above $200 + 1$ meters from one point to the closest other point.

$$Orientation_{i,t} = \sum_{bin}^{18} street_length_{i,bin} \quad (4.3)$$

where $street_length_{bin}$ measures the total length of streets of city i that are orientated in the direction of one of 18 directions, indexed by bin . Figure 4.11 shows two examples for 1934: Remscheid with a chaotic street network at the top left with a Herfindahl index of 0.058, and Wilhelmshaven with a less chaotic one at the bottom left with an index of 0.15. Note that in a perfectly rectangular street grid system, which represents itself as having an equal share between two bins, we would find a Herfindahl index of 0.5. The closest to that measure in 1934 in our sample is the 500m range inner city of Mannheim with a Herfindahl index of 0.43 and Krefeld with 0.39. This is not surprising, as these cities were initially planned to have a rectangular city grid. On the other hand, Karlsruhe with an index of 0.09 was also planned as a radial grid in which the grid was initially created with the palace of an absolutist ruler in the center, and roads emanating out of this palace, with perfectly circular ring roads surrounding the palace.¹⁶ Here, an unplanned, chaotic grid type of random orientation would be close to a Herfindahl index of $1/18$. Hence, the limitation of this measure is that although the city center of Karlsruhe is a radial grid and highly symmetrical, it is not reflected in our measure. To illustrate this issue, we could imagine a city that consists of a single, perfectly circular ring road and nothing else, for which we would find a Herfindahl index of $1/18$ with 18 different orientation bins.

Another issue may be called the 'opposing neighborhood problem', which can be illustrated by two perfectly rectangular grid neighborhoods of the same size on the opposite side of the city center. Even though we would find an index of 0.5 for both neighborhoods individually, if one is tilted 45 degrees, the measure for both neighborhoods would be 0.25. In our sample, as neighborhoods were initially build along existing roads emanating from the city center, they would be orientated as such towards the city center. This is indeed what we find as we increase the range from 500 to 1500m, the index of the average 1934 city decreases from 0.112 to 0.088 to 0.076. This measure is then best defined as a measure for how close it is to a squared grid, and not necessarily how less chaotic a grid is, especially if we increase the range. As these are 'level' issues, we can abstract away from these issues by considering the post-war changes of these measure. Finally, Table 4.1 reports the descriptive statistics for the aforementioned variables.

¹⁶Apparently, this grid form was beneficial for par force hunting.

Table 4.1: Descriptive Statistics

Variable	Range (meter)	mean	std. dev.	min	max
log population 1939		11.7727	0.8490	10.0123	14.3531
log population 1945		11.5474	0.8083	9.7410	14.1158
log population 1960		11.9820	0.7849	10.6115	14.4196
log population 1980		11.9657	0.7317	10.8824	14.3133
log population 2000		11.9706	0.7143	10.9231	14.3552
rubble m^3 per capita		11.2973	8.3658	0.3	31.3
river		0.8378	0.3711	0	1
great river		0.3514	0.4807	0	1
port		0.0811	0.2748	0	1
log city rights		7.1675	0.4656	3.9120	7.5632
log first record city		6.6980	0.7544	3.7842	7.5332
wall length density	0 to 500	0.0012	0.0013	0	0.00501
	0 to 1000	0.0008	0.0007	0	0.0025
	0 to 1500	0.0005	0.0005	0	0.0020
density of boulevardized walls	0 to 500	0.0004	0.0006	0	0.0020
	0 to 1000	0.0004	0.0005	0	0.0025
	0 to 1500	0.0002	0.0003	0	0.0014
boulevard density, 1934	0 to 500	0.0014	0.0009	0	0.0040
	0 to 1000	0.0016	0.0007	0.0005	0.0048
	0 to 1500	0.0015	0.0005	0.0007	0.0041
boulevard density, 2020	0 to 500	0.0024	0.0012	0	0.0059
	0 to 1000	0.0025	0.0008	0.0010	0.0058
	0 to 1500	0.0023	0.0006	0.0011	0.0048
street density, 1934	0 to 500	0.0170	0.0038	0.0087	0.0261
	0 to 1000	0.0131	0.0029	0.0078	0.0203
	0 to 1500	0.0110	0.0025	0.0051	0.0164
street density, 2020	0 to 500	0.0169	0.0033	0.0098	0.0238
	0 to 1000	0.0134	0.0025	0.0086	0.0211
	0 to 1500	0.0117	0.0022	0.0063	0.0167
orientation Herfindahl, 1934	0 to 500	0.1115	0.0643	0.0609	0.4300
	0 to 1000	0.0876	0.0342	0.0583	0.2286
	0 to 1500	0.0763	0.0195	0.0578	0.1541
orientation Herfindahl, 2020	0 to 500	0.1117	0.0640	0.0633	0.4459
	0 to 1000	0.0869	0.0332	0.0581	0.2271
	0 to 1500	0.0758	0.0181	0.0573	0.1464
unevenness, 1934	0 to 500	23.8905	7.6575	12.9985	47.4585
	0 to 1000	34.5178	9.2840	17.2916	57.3928
	0 to 1500	40.7408	8.4493	24.1	59.7519
unevenness, 2020	0 to 500	23.6320	7.4430	11.7994	48.5814
	0 to 1000	32.2791	8.6298	16.7992	60.2597
	0 to 1500	36.9641	7.1712	23.4921	60.8603
unevenness st. dev., 1934	0 to 500	21.2601	9.8169	9.3834	56.8236
	0 to 1000	33.5220	9.4694	14.1247	55.6911
	0 to 1500	38.6327	7.0302	21.8243	54.4088
unevenness st. dev., 2020	0 to 500	21.3378	10.8128	7.8951	54.5181
	0 to 1000	31.3643	9.4307	14.9121	54.9911
	0 to 1500	36.2839	6.8023	22.8110	54.8428
extent 1934	0 to 1500	171.9189	13.8067	102	180
extent 2020	0 to 1500	173.6351	11.3301	119	180

Note: This table reports the summary statistics for 74 German cities, which had a population of 50,000 at any point in time between 1939 and 2000, with the exception of Berlin, Munich, Salzgitter, Wolfsburg, Heilbronn, as described in section 4.3.1. The unevenness and orientation Herfindahl measures are described in section 4.3.2. The population data is adjusted for amalgamations of municipalities over time, as described in the appendix B. The river dummy indicates a river within 1km of the municipal border of today. The great river dummy indicates a river that is at least 100m wide and that flows through the city, within 1500 meters away from the town hall of 1945. These rivers are the Rhine, Danube, Elbe, Weser, the Main river and the Neckar river only at Heidelberg. Boulevardized walls measure the boulevardized parts of today, that existed in its width as seen in the Meßtischblätter topographic maps of 1933/34, and were found around the historical walls. This includes the area on the defortified walls, as well as the immediate area in front that coincides with the course of these walls. We describe this measure in detail in section 4.3.3.

4.3.3 Boulevardization of Walls Instrument

With a radius of 1.5km around the town hall, Figure 4.5 shows the positive relationship between the length density of walls and that of boulevards in 1934 as described before, that are however situated on or close to the walls. That is, these specific boulevard parts that existed in 1934, was not widened after that, and stem from the defortification of the walls. Hence, we call this measure 'boulevardized walls'. These parts may coincide with the areas directly on top of the defortified walls, but we also consider the area right in front of the walls, as long as the course of the walls and the course of the boulevard of 1934 coincide, reflecting that some of the area in front of the walls were left empty in order to have a free field of fire.

In Figure 4.5, there are a few notable outliers: Frankfurt am Main is above the 90 degree line, suggesting that the length of boulevardized walls exceeded the length of the walls itself. In the case of Frankfurt, they created boulevards right behind the walls, as well as another boulevard on or along the ditches 'outside' the walls, which we also consider part of the walls.

Also notable are Mainz and Ulm for its very long walls. These cities were known as 'Bundesfestung' (Fortress of the German Confederation), which were designated as strategically important against France. These state of the art bulwarks were paid by the French after the Peace of Paris in 1815. Figure 4.10 shows the extent of the Fortress of Ulm. The city of Ulm itself was also directly surrounded by 'normal' walls (and not depicted, broad fortified ditches), thus leading to an unusually large area occupied by walls. Similarly, Nuernberg had both the simple 'enceinte' medieval walls, which were build at a time when the citizens did not yet worry about modern cannons, as well as an actual bulwark with its typical outward bastions. Notable here is that Nuernberg is the only city with an almost fully intact enceinte, while the more space occupying bulwarks were razed.¹⁷

Lastly, Figure 4.6 shows that cities that decided to defortify later also boulevardized more along the walls, which is in line with the notion that there is more need for boulevards throughout time, whether it is due to industrialization, urbanization, or the emergence of horse-drawn carriages and omnibusses. It also suggests that cities which defortified early may be more likely to be stuck with a lower boulevard density that aligns better with the needs of the medieval city.

For instance, the city of Pforzheim lost its walls in 1645 during the Thirty Years' War in a fire, and none of the wall area were used to create boulevards as seen today. Instead, that area was used for housing. Similarly in Figure 4.7, we show the map of Kassel in 1934 and today, which began defortification in 1767 right after the Seven Years' War, as do a few other cities during that time period, where

¹⁷The other outlier in our sample is Duisburg, with its enceinte's surviving even today.

Figure 4.5: Boulevardization of Walls

We show the length of boulevards that were built on or around the former walls and its relationship with the total length of these walls. The upper dotted line shows the 90 degree line i.e. a 1 to 1 conversion. Frankfurt is the only city above that line as boulevards were build right behind and in front of its defortified bulwark. The dashed line is the linear fit.

again the area was primarily used for new housing. Vice versa, we can consider cities as lucky if they defortified at a later date, giving them the opportunity to create boulevards that align closer to more modern infrastructure needs.

4.4 Results

Table 4.2 reports the naive OLS regression estimates, where we have included all the aforementioned control variables, some of which may already deal with bias arising from omitted variables. We progressively expand the range (from the town hall) that can be considered as the inner city, starting with a 500 meter radius in (1) and (2) to a 1500 meter radius in (5) and (6). We further split the time period investigated between 1946-66, and 1966-86 corresponding to the 1960s and 1970s during the 'second destruction' i.e. around the time when the bulk of post-WWII boulevards were built. Specification (1), (3) and (5) reports the estimates for 1946-66, while (2), (4) and (6) reports the corresponding estimates for 1966-86. As described in the data section 4.3.1, we chose 1946 due to data limitations of the Bundesländer state level population data, which covers the years of 1946 to 1999 in a time consistent manner. We also choose 1986 due to a change in the way of how

Figure 4.6: Boulevardized Walls Rate and Approx. Earliest Defortification Year

We show the Boulevardization rate of walls, and its relationship with the approximate defortification date. Later defortification corresponds to a higher boulevardization rate of walls. Note that very low rates of boulevardization close to 0 may be due to 1934 boulevards intersecting with a single point around former wall areas. We also exclude Solingen due to a lack of data. Sources in the line of 'beginning of the 19th century' are assigned as the year 1800. The dashed line is the linear fit.

the population is counted in 1987, which led to a significant break in some cities.¹⁸ Lastly, as we have a non-randomized sample of the largest cities in Germany, and as the variation of interest is at the city level, we follow the guideline by Abadie et al. (2017) and cluster standard errors at the city level, equivalent in this case to the robust standard errors with the long estimation. Furthermore, clustering standard errors at the 'Bundesländer' state level generally provided lower standard errors, but may be less reliable due to the low number clusters of 11 in our sample.

As stated before, we should expect no significant impact of the increase in boulevard density between 1946 to 1966, that is right after the 'first destruction' of WWII took place, as the expansion in boulevards took place during the 'second destruction'. We however report a significant and negative coefficient throughout all specifications, suggesting suburbanization before a significant amount of boulevards were built, as well as after. As discussed before, these results seem implausible especially for boulevards.

¹⁸The definition was changed from Wohnbevölkerung (inhabitants living there) to Bevölkerung am Ort der Hauptwohnung (inhabitants at the place of the main address).

We also report a positive effect of an increase in the street density of a city, but only between 1966-86.

We have also used the change in the population between 1939 and 1945 as a proxy for a population disruptions caused by WWII, for which we find a strong and significant effect between 1946 and 1966. We interpret this as a strong recovery effect of population growth after the war. This effect disappears in the 20 year time period from 1966 to 1986.

We report the first-stage regression in Table 4.3. The signs for the instruments are as expected in panel A, with a higher level of boulevardized walls BW within the 1500m range corresponding to a lower post-war increase in boulevards $\Delta boulevard_i$. Second, we find that the destruction, as measured by the cubic meter rubble in 1945 per 1939 capita, is not significant for the within 500m range, but for the within 1000m and within 1500m range, with a positive sign on additional boulevards. One explanation is that city planners may not have wanted to build boulevards directly through the within 500m core city.

In panel B, we find a higher level of destruction corresponding to a lower post-war change of street density $\Delta street_i$, but no impact of previous levels of boulevardized walls BW on a change in street density. The conditional first-stage F for the $\Delta boulevard_i$ first-stage are above the 10 and 15 percent critical value for two endogenous variables and two instruments (Stock & Yogo, 2005) of 7.03 and 4.58, indicating that the instruments for $\Delta boulevard_i$ are not weak.

Lastly, Table 4.4 reports the second-stage results, directly corresponding to columns of the first-stage of Table 4.3. The specifications in Table 4.4 also correspond directly to the uninstrumented naive OLS specification of Table 4.2 to retain comparability. We do not estimate any statistically significant impact between 1946 to 1966 of additional boulevards anymore, as compared to the OLS estimates of Table 4.2. As these results seem more plausible, we are more certain to have isolated the impact of boulevards. Second, our estimated coefficients for the period of 1966 to 1986 are now significantly larger as compared to the OLS estimates. The ratio of the 2SLS to OLS coefficient is about 2, which is similar to the one found in the US suburbanization study by Baum-Snow (2007), but smaller than the ratio estimated for Spain (García-López et al., 2015) and China (Baum-Snow et al., 2017).

Third, we find no impact of a change in street density on population growth anymore for any time period. Lastly, the impact of the WWII recovery on population growth seem to not be there anymore from 1966 onwards, similar to the OLS results.

Taking our preferred specification of a 1000m radius, we estimate that the post-WWII increase from the average boulevard length density of 0.0016 to 0.0025 led to

an overall decline of around 8.22% between 1966 to 1986. Similarly for the 500m range, we estimate an overall decline of 7.00% and 8.41% for the 1500m range, which are larger than the estimated effect through OLS. The effect of boulevards are roughly in line with previous estimates on changes in the highway infrastructure on the central city population growth of around 10% to 11% (Baum-Snow, 2007), 8-9% (Garcia-López et al., 2015) and 4% (Baum-Snow et al., 2017). Secondly, we find no evidence that changes in the stock of narrower streets within the city has any impact.¹⁹

Table 4.2: OLS estimates: Effect of a Change in the Street Infrastructure

specification	(1)	(2)	(3)	(4)	(5)	(6)
range radius (meter)	0 to 500	0 to 500	0 to 1000	0 to 1000	0 to 1500	0 to 1500
time period	1946-1966	1966-1986	1946-1966	1966-1986	1946-1966	1966-1986
Dependent variable: $\Delta_T \log population_i$						
$\Delta boulevard_i$	-2.2809** (0.9976)	-1.4528** (0.5645)	-3.0244** (1.3727)	-2.2247*** (0.7179)	-5.4462*** (2.0428)	-3.8678*** (1.1435)
$\Delta street_i$	0.1715 (0.7639)	0.9376** (0.3604)	-0.0483 (1.3780)	2.2621*** (0.6851)	-0.6659 (1.8931)	1.9824** (0.8497)
$\Delta \log population_S$ 1946 – 1966	0.2989		0.3653*		0.4220*	
$\Delta \log population_S$ 1966 – 1986		0.4097*		0.3410		0.2488
$\log population_S$ 1946	0.0024		0.0024		0.0028	
$\log population_S$ 1966		0.0000		-0.0001		0.0002
$\Delta \log population$ 1939 – 1945	-0.0151**	0.0015	-0.0148**	0.0003	-0.0146**	0.0007
log city rights year	-0.0031	-0.0025	-0.0037*	-0.0035*	-0.0040**	-0.0034*
state capital indicator	-0.0018	-0.0016	-0.0021	-0.0019	-0.0019	-0.0012
port	-0.0074**	-0.0033	-0.0060	-0.0019	-0.0043	-0.0019
river	0.0022	0.0007	0.0028	0.0021	0.0034	0.0021
within 1500m great river	-0.0002	-0.0027*	0.0003	-0.0027*	0.0002	-0.0021
within 1500m max. extent 2020	-0.0001	-0.0001	-0.0002	-0.0001**	-0.0002	-0.0001**
within 1500m $\Delta extent$	-0.0004	-0.0001	-0.0004	-0.0001	-0.0002	-0.0000
$\Delta Herfindahl Orientation$	-0.0875	0.0161	0.0767	0.2471*	0.5731*	0.2513**
$\Delta Unevenness st.dev.$	-0.0002*	-0.0000	-0.0003	0.0000	-0.0003	-0.0000
constant	0.0172	0.0302	0.0328	0.0440	0.0307	0.0332
R^2	0.4814	0.3172	0.4512	0.3909	0.4749	0.4428

Note: Each regression has 74 observations. *, **, *** denotes significance at the 10%, 5%, 1%. Robust standard errors are in parentheses. Note that the difference operator Δ_T measures the $\#T$ long difference of the log population variable for the indicated time period. The restricted range measured from the town hall for each regression of the variables $\Delta boulevard_i$, $\Delta street_i$, $\Delta Herfindahl Orientation$ and $\Delta Unevenness st.dev.$ are indicated by the range radius.

4.5 Robustness

4.5.1 Finer Patterns over Time

We have already provided different definitions of what can be considered as a city center i.e. 500m to 1500m. In this section, we show some additional time patterns.

¹⁹This is not consistent with the findings by Rauch & Brandily (2020) that Sub-Saharan African cities that are denser in roads grew faster. This may already hint at a heterogeneous effect of street density depending on the level of development or urbanisation.

Table 4.3: First Stage: Effect of a Change in the Street Infrastructure

specification	(1)	(2)	(3)	(4)	(5)	(6)
range radius (meter)	0 to 500	0 to 500	0 to 1000	0 to 1000	0 to 1500	0 to 1500
time period	1946-1966	1966-1986	1946-1966	1966-1986	1946-1966	1966-1986
Panel A: Dependent variable: $\Delta \text{boulevard}_i$						
Boulevardized Walls (1500m)	-1.0522** (0.4264)	-1.0342** (0.4391)	-0.7903*** (0.2634)	-0.7701*** (0.2818)	-0.5328*** (0.1593)	-0.5104*** (0.1716)
Rubble per capita	0.00003* (0.00001)	0.00003* (0.00001)	0.00003** (0.00001)	0.00003** (0.00001)	0.00003*** (0.00000)	0.00002*** (0.00000)
$\Delta \log \text{population}_S$ 1946 – 1966	0.0252		0.0246*		0.0165*	
$\Delta \log \text{population}_S$ 1966 – 1986		-0.0495		-0.0429		-0.0350
$\log \text{population}_S$ 1946	0.0001		0.0000		0.0001	
$\log \text{population}_S$ 1966		0.0001		0.0001		0.0001*
$\Delta \log \text{population}$ 1939 – 1945	-0.0005*	-0.0006**	-0.0002	-0.0003	-0.0000	-0.0001
log city rights year	-0.0001	-0.0001	-0.0004*	-0.0003*	-0.0002	-0.0002
state capital indicator	-0.0003	-0.0004	-0.0002	-0.0002	0.0000	-0.0000
port	-0.0002	-0.0004	0.0000	-0.0002	0.0001	-0.0000
river	-0.0002	-0.0002	0.0000	0.0000	0.0001	0.0001
within 1500m great river	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001
within 1500m max. extent 2020	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
within 1500m Δextent	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	0.0000
$\Delta \text{Herfindahl Orientation}$	-0.0062	-0.0051	0.0218	0.0239	0.0197	0.0229
$\Delta \text{Unevenness st.dev.}$	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000**	-0.0000**
constant	0.0022	0.0016	0.0043	0.0037	0.0006	0.0002
Panel B: Dependent variable: Δstreet_i						
Boulevardized Walls (1500m)	-0.8209 (0.8426)	-0.6755 (0.8223)	-0.4062 (0.3309)	-0.3664 (0.3169)	-0.0332 (0.2778)	-0.0072 (0.2723)
Rubble per capita	-0.00007*** (0.00002)	-0.00008*** (0.00002)	-0.00005*** (0.00001)	-0.00005*** (0.00001)	-0.00005*** (0.00001)	-0.00005*** (0.00001)
$\Delta \log \text{population}_S$ 1946 – 1966	-0.0047		0.0199		0.0204	
$\Delta \log \text{population}_S$ 1966 – 1986		-0.0952*		-0.0456		-0.0393
$\log \text{population}_S$ 1946	0.0002		0.0002		0.0002*	
$\log \text{population}_S$ 1966		0.0002		0.0002		0.0003***
$\Delta \log \text{population}$ 1939 – 1945	-0.0009	-0.0008*	-0.0001	-0.0001	-0.0003	-0.0003
log city rights year	-0.0005	-0.0005	-0.0003	-0.0003	-0.0001	-0.0001
state capital indicator	-0.0003	-0.0003	0.0000	-0.0000	-0.0001	-0.0001
port	-0.0002	-0.0001	-0.0002	-0.0003	0.0000	-0.0001
river	-0.0006	-0.0004	-0.0006*	-0.0005*	-0.0004*	-0.0004*
within 1500m great river	0.0009**	0.0008**	0.0004**	0.0004**	0.0003	0.0003
within 1500m extent 2020	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
within 1500m Δextent	0.0001	0.0001*	0.0000	0.0000	0.0000	0.0000
$\Delta \text{Herfindahl Orientation}$	-0.0303***	-0.0280**	-0.0291	-0.0257	0.0025	0.0061
$\Delta \text{Unevenness st.dev.}$	-0.0000**	-0.0000**	-0.0001***	-0.0001***	-0.0001***	-0.0001***
constant	-0.0030	-0.0009	-0.0018	-0.0020	-0.0035	-0.0039
SW F Stat $\Delta_T \text{boulevard}_i$	10.14***	9.55***	15.21***	12.91***	8.50***	6.97**
SW F Stat $\Delta_T \text{street}_i$	12.31***	13.18***	15.58***	15.03***	9.36***	9.07***
SW Chi $\Delta_T \text{boulevard}_i$	12.71***	11.98***	19.08***	16.20***	10.66***	8.74***
SW Chi $\Delta_T \text{street}_i$	15.43***	16.53***	17.04***	18.85***	11.74***	11.37***

Note: Each regression has 74 observations. *, **, *** denotes significance at the 10%, 5%, 1%. Robust standard errors are in parentheses. Note that the difference operator Δ_T measures the $\#T$ long difference of the log population variable for the indicated time period. The restricted range measured from the town hall for each regression of the variables $\Delta \text{boulevard}_i$, Δstreet_i , $\Delta \text{Herfindahl Orientation}$ and $\Delta \text{Unevenness st.dev.}$ are indicated by the range radius. SW F Stat and SW Chi respectively are the Sanderson & Windmeijer (2016) test for weak- and underidentification respectively. The Stock & Yogo (2005) 10 and 15 percent critical values for two endogeneous variables and instruments are 7.03 and 4.58.

Table 4.4: Second Stage: Effect of a Change in the Street Infrastructure

specification	(1)	(2)	(3)	(4)	(5)	(6)
range radius (meter)	0 to 500	0 to 500	0 to 1000	0 to 1000	0 to 1500	0 to 1500
time period	1946-1966	1966-1986	1946-1966	1966-1986	1946-1966	1966-1986
Dependent variable: $\Delta_T \log population_i$						
$\Delta boulevard_i$	-3.7443 (3.5947)	-3.4738* (1.8414)	-4.5664 (4.3106)	-4.6763** (2.1043)	-11.4856 (7.7666)	-5.8402* (3.4576)
$\Delta street_i$	-2.8238 (2.1594)	1.1153 (0.9339)	-5.0610 (3.5476)	1.3374 (1.5233)	-8.6316 (5.2185)	0.3540 (2.3022)
$\Delta \log population_S$ 1946 – 1966	0.3182		0.4958		0.6822**	
$\Delta \log population_S$ 1966 – 1986		0.2758		0.1362		0.1001
$\log population_S$ 1946	0.0030		0.0090		0.0049	
$\log population_S$ 1966		0.0005		0.0005		0.0009
$\Delta \log population$ 1939 – 1945	-0.0156***	-0.0004	-0.0125**	-0.0007	-0.0136***	0.0005
log city rights year	-0.0029	-0.0024	-0.0039*	-0.0041***	-0.0043**	-0.0035**
state capital indicator	-0.0020	-0.0025	-0.0018	-0.0026	-0.0021	-0.0015
port	-0.0095*	-0.0037**	-0.0081*	-0.0026	-0.0046	-0.0023
river	0.0001	0.0004	-0.0001	0.0016	0.0006	0.0016
within 1500m great river	0.0014	-0.0031*	0.0012	-0.0029*	0.0019	-0.0018
within 1500m max. extent 2020	-0.0001	-0.0001	-0.0002	-0.0002**	-0.0002	-0.0001*
within 1500m $\Delta extent$	-0.0002	-0.0001	-0.0003	-0.0000	0.0001	0.0000
$\Delta Herfindahl Orientation$	-0.2003	0.0120	-0.0560	0.3078**	0.7541	0.3178*
$\Delta Unevenness st.dev.$	-0.0004*	-0.0001	-0.0006*	-0.0000	-0.0012	-0.0002
constant	0.0035	0.02938	0.0253	0.0487*	0.0020	0.0258

Note: Each regression has 74 observations. *, **, *** denotes significance at the 10%, 5%, 1%. Robust standard errors are in parentheses. Note that the difference operator Δ_T measures the $\#T$ long difference of the log population variable for the indicated time period. The restricted range measured from the town hall for each regression of the variables $\Delta boulevard_i$, $\Delta street_i$, $\Delta Herfindahl Orientation$ and $\Delta Unevenness st.dev.$ are indicated by the range radius.

Table 4.5 reports the same specification patterns as Table 4.4, with the difference that we now consider 10 year intervals instead of 20 year intervals in order to learn about the patterns over time in more detail. We consistently find for all range specifications that the decentralization effect of an increase in boulevards were only found between the 1976 to 1986 period. The estimated coefficients are about twice the size of those estimated in Table 4.4, suggesting overall that the significant results of Table 4.4 were entirely driven during this period. That is, over this time period, we find an overall, non-annualized average effect of 5.73%, 6.78%, 7.90% decline for the 500, 1000 and 1500 meter range respectively. Lastly, we do not find any effect during the 1986 to 1996 period, suggesting that the additional boulevards segments had a permanent effect on displacing population, but that the development itself was of temporary nature.

Also similar to (Baum-Snow et al., 2017), it takes about a decade after initial changes in the 1960s to find an effect. This may simply reflect the time it takes to develop the land and to construct housing. We only find a negative, but weakly statistically significant effect for the immediate post-war rebuilding period between 1946-56 for both the 1000m and 1500m range. We also find a positive, but weakly

statistically significant impact of streets for the 500m and 1000m range between 1966-76. Since this coincides with the Flächensanierung period, it may reflect that new streets are build faster than larger boulevards.

Lastly, for all range specifications, the impact of WWII could only be measured during the period of 1946 to 1956, many years before we find the effect of the increase in boulevards. We consider this gap as reassurance that the exclusion restriction assumption holds.

4.5.2 Robustness: Street Infrastructure and Hinterland Attractiveness

We consider an urban potential measure as another measure for the changing attractiveness of a region due to the changes in the street infrastructure in a city. That is, we define a simple linear urban potential measure:

$$UP_{i,t} = \sum_j^n \frac{S_{i,t}}{\tau_{i,j}} \quad (4.4)$$

where $\tau_{i,j}$ is the kilometer distance between city i and j , and $S_{i,1940}$ is the population of city i in year t . We consider the other cities in our sample for this measure, also including the already excluded Heilbronn, München, Wolfsburg and West-Berlin here, for a total of 79 cities of a population above 50000.²⁰ However, as the growth of this measure is very highly correlated with our variable of interest city population growth, we subtract the population of the city itself from here. As such, we measure the surrounding economic activity of larger cities. For instance, commuting could take place between cities, which is more in line with this measure. It also takes into account the loss in attractiveness of cities at the East-West German border due to the division of Germany as we do not count the urban potential effects of East German cities here (Redding & Sturm, 2008). For instance, the division could have led to a lower expected growth in cities along the border, and lower infrastructure investments in these cities.

In Table 4.6, we consider the same specification as previous subsection for the years 1976 to 1986 of Table 4.5, but add urban potential growth between 1976 to 1986 in column 1 to 3. As compared to Table 4.5, the estimated coefficients drop slightly from -6.00 to -5.60 (500m range), -8.15 to -7.79 (1000m) and -11.58 to -10.25 (1500m).

Additionally, following Baum-Snow et al. (2017), we can obtain a bound on the value for the coefficients of both streets and boulevards by excluding both state level population measures as well as the urban potential measure. In column 4

²⁰Note that we did not consider Salzgitter a city.

Table 4.5: IV Estimates: Finer Patterns over Time

specification time period	(1) 1946-1956	(2) 1956-1966	(3) 1966-1976	(4) 1976-1986	(5) 1986-1996
Dependent variable: $\Delta_T \log population_i$					
Panel A: 0 to 500 meter					
$\Delta boulevard_i$	-7.7909 (6.6990)	-0.0351 (4.2056)	-1.5333 (2.3061)	-6.0084** (2.3202)	-1.7782 (2.6082)
$\Delta street_i$	-5.1931 (3.5817)	-0.7628 (1.9120)	2.0736* (1.2013)	0.3331 (1.2779)	-1.2195 (1.1134)
$\Delta \log population$ 1939 – 1945	-0.0283*** (0.0105)	-0.0047 (0.0055)	0.0016 (0.0033)	-0.0027 (0.0026)	-0.0001 (0.0033)
SW F Stat $\Delta_T boulevard_i$	9.82***	11.90***	10.26***	8.67***	10.94***
SW F Stat $\Delta_T street_i$	12.28***	12.47***	13.08***	13.28***	13.39***
SW Chi $\Delta_T boulevard_i$	12.32***	14.92***	12.86***	10.87***	13.72***
SW Chi $\Delta_T street_i$	15.40***	15.64***	16.41***	16.66***	16.80***
Panel B: 0 to 1000 meter					
$\Delta boulevard_i$	-9.4061 (7.4941)	-0.4441 (5.0518)	-1.9507 (2.6222)	-8.1467*** (2.7893)	-1.9117 (3.5324)
$\Delta street_i$	-10.0349* (5.7603)	-0.6181 (3.4085)	3.4778* (1.9720)	-0.5938 (2.3489)	-2.4754 (1.9631)
$\Delta \log population$ 1939 – 1945	-0.0224** (0.0089)	-0.0042 (0.0049)	-0.0004 (0.0032)	-0.0020 (0.0023)	0.0012 (0.0033)
SW F Stat $\Delta_T boulevard_i$	14.25***	18.39***	13.48***	12.35***	15.88***
SW F Stat $\Delta_T street_i$	13.60***	13.36***	15.28***	14.64***	13.58***
SW Chi $\Delta_T boulevard_i$	17.87***	23.07***	16.91***	15.49***	19.91***
SW Chi $\Delta_T street_i$	17.06***	16.75***	19.16***	18.36***	17.03***
Panel C: 0 to 1500 meter					
$\Delta boulevard_i$	-22.5191 (14.1079)	-0.7801 (7.7387)	-0.9007 (4.3887)	-11.5827** (5.3644)	-5.2556 (7.0524)
$\Delta street_i$	-15.6698* (9.2141)	-2.2836 (4.8720)	3.8624 (3.0796)	-3.0127 (3.5569)	-3.6908 (3.9412)
$\Delta \log population$ 1945 – 1939	-0.0250*** (0.0067)	-0.0044 (0.0045)	0.0021 (0.0033)	-0.0012 (0.0023)	0.0000 (0.0038)
SW F Stat $\Delta_T boulevard_i$	8.20***	8.75***	6.78**	7.28***	7.28***
SW F Stat $\Delta_T street_i$	9.39***	8.58***	9.04***	9.04***	7.56***
SW Chi $\Delta_T boulevard_i$	10.29***	10.98***	8.50***	9.13***	9.13***
SW Chi $\Delta_T street_i$	11.78***	10.76***	11.34***	11.34***	9.48***

Note: Each regression has 74 observations. The specifications are the same as in Table 4.4, with the exception that the state population in a 10 year instead of 20 year interval i.e. for the first specification 1946-1956 we consider, $\Delta \log population_S$ 1946 – 1956, as well as $\log population_S$ 1946. Equivalently, for the second specification 1956-1966, we consider $\Delta \log population_S$ 1956 – 1966 and $\log population_S$ 1956 etc. *, **, *** denotes significance at the 10%, 5%, 1%. Robust standard errors are in parentheses. Note that the difference operator Δ_T measures the $\#T$ long difference of the log population variable for the indicated time period. SW F Stat and SW Chi respectively are the Sanderson & Windmeijer (2016) test for weak- and underidentification respectively. The Stock & Yogo (2005) 10 and 15 percent critical values for two endogeneous variables and instruments are 7.03 and 4.58.

to 6 of Table 4.6, the coefficients barely change, indicating that the attractiveness hypothesis does not seem to be very relevant in this case.

Table 4.6: IV Estimates: Street Infrastructure and Hinterland Attractiveness

specification	(1)	(2)	(3)	(4)	(5)	(6)
range radius (meter)	0 to 500	0 to 1000	0 to 1500	0 to 500	0 to 1000	0 to 1500
time period	1976-1986	1976-1986	1976-1986	1976-1986	1976-1986	1976-1986
Dependent variable: $\Delta_T \log population_i$						
$\Delta boulevard_i$	-5.6053** (2.5539)	-7.7885*** (3.0004)	-10.2456** (5.0760)	-5.6276*** (1.7793)	-7.2684*** (1.9724)	-9.8127** (4.0172)
$\Delta effective street_i$	0.7182 (1.4570)	-0.1284 (2.8428)	-1.9853 (3.5095)	0.1481 (1.2906)	-0.8978 (2.4416)	-3.3418 (3.7222)
Base controls	yes	yes	yes	yes	yes	yes
Urban potential control	yes	yes	yes	no	no	no
Exclude state pop. controls	no	no	no	yes	yes	yes
SW F Stat $\Delta_T boulevard_i$	8.63***	10.28***	6.23**	13.85***	22.52***	9.40***
SW F Stat $\Delta_T street_i$	10.46***	11.37***	8.06***	12.97***	12.01***	7.94***
SW Chi $\Delta_T boulevard_i$	11.01***	13.12***	7.95***	16.80***	27.32***	11.40***
SW Chi $\Delta_T street_i$	13.35***	14.50***	10.28***	15.74***	14.58***	9.63***

Note: Each regression has 74 observations. Second stage dependent variable: $\Delta_T \log population_i$. The specification is the same as in table 5 (base controls). The urban potential controls consists of the difference between the 1986 and 1976 log urban potential measure equation (4) minus the population of the given city i in 1976 i.e. the urban potential growth of 1976 to 1986. *, **, *** denotes significance at the 10%, 5%, 1%. Robust standard errors are in parentheses. Note that the difference operator Δ_T measures the $\#T$ long difference of the log population variable for the indicated time period. SW F Stat and SW Chi respectively are the Sanderson & Windmeijer (2016) test for weak- and underidentification respectively. The Stock & Yogo (2005) 10 and 15 percent critical values for two endogeneous variables and instruments are 7.03 and 4.58.

4.5.3 Effective Street Density

In Table 4.1, we see that the orientation Herfindahl did not change much over time, especially within the 500m range, and even decreased slightly, suggesting that the city was not rebuild so that street would follow the overall city alignment of street orientations to another. However, more bombed cities were subject to an overall increase in the Herfindahl index, although the correlation is rather weak with 0.1338, 0.1579, 0.0518 for the 500m to 1500m range respectively. Still, city planners after WWII may have also looked to increase the alignment of streets along with the change of the street network, at least as a secondary objective. For example, modern cities may have wanted to avoid zig-zag streets, often found along the more sophisticated and modern bulwarks, that are not made for a smooth and straight traffic of cars. This would mean that an increase in the alignment may not be a consequence of a change in the density of streets, and as such taken as random itself, but may jointly be the subject of change similar to the street density.

We consider an effective street density measure *effective street_i* as the interaction term of both the street density *street_i* and the Herfindahl orientation

measure, and treat the change of this measure from 1934 to 2020 as endogenous. The first two columns of Table 4.7 reports the specification as in Table 4.4 with this effective street density measure. For one, we still see no impact of the effective street density measure, while the boulevard impact is robust for the 1000m and 1500m range as compared the previously estimated coefficients of Table 4.4. Second, the Sanderson-Windmeijer (2016) conditional F-statistics for the first stage suggest that we cannot reject that the instruments are weak, particularly for the within 500m estimates. Within this range, the WWII destruction did not influence this effective street density measure, as compared to the usual street density measure seen in previous specifications. This however is less the case for the 1000m and 1500m range, overall suggesting a heterogeneous impact of the instruments depending on the range from the immediate city center. We suspect that this has something to do with the 'old city' Altstadt i.e. the historical city, often being very distinct in its street network from the parts outside the historical city. In particular, this part is often more dense in streets, where the streets are often less aligned to each other, contrary what the Table 4.1 descriptive statistics would tell us. The discussion in section 4.3.2 outlines the shortcomings of the Herfindahl orientation measure itself, which may be effective in measuring differences in the alignment of the streets between cities, but not between the different ranges within a city.²¹

One way to instrument for that is to consider the first record of when this city was first mentioned, which serves as a proxy for age.²² The third and fourth column of Table 4.7 includes this instrument. We find statistical significance of this instrument at the 5% level for the panel A 500m range, but it disappears for the specifications including the 1000m and 1500m range in panel B and C. This suggests that the age of the inner city matters the most for whether the effective street density changes after WWII, but very little for the areas outside the often historical city center. For the 1000m and 1500m specification, it appears to be a weaker instrument as expected.

As we suspect a heterogeneous impact of the instruments depending on the range considered, we will exclude the WWII destruction instrument in column 5 and 6 for the 500m range, and the first record instrument for the 1000m and 1500m range, while also restricting the analysis to the relevant 1966-76 and 1976-86 interval as in section 4.5.1. With these specifications, the Sanderson-Windmeijer (2016) conditional F-statistics for the first stage are above the (Stock & Yogo, 2005)

²¹From the discussion in section 4.3.2 about the orientation Herfindahl, as we increase the range, the Herfindahl index may decrease even if the streets are more aligned, due to the 'opposing neighborhood problem' and the issues surrounding the measurement of this index in the case of a radial grid.

²²City rights may also serve as a proxy, but we consider that as a proxy for a more matured city, whereas the first record may be considered as a more appropriate instrument for the older historical city center.

Table 4.7: IV Estimates: Effective Street Density

specification time period	(1) 1946-1966	(2) 1966-1986	(3) 1946-1966	(4) 1966-1986	(5) 1966-1976	(6) 1976-1986
Panel A: Second stage $\Delta_T \log population_i$ (0 to 500 meter)						
$\Delta boulevard_i$	-2.9530 (4.1251)	-3.7900 (2.3414)	-4.1125 (3.1629)	-3.3655* (1.7206)	-0.5613 (2.3705)	-5.5130** (2.1458)
$\Delta effective street_i$	-46.2120 (44.8111)	20.2166 (19.7956)	-12.6594 (17.3195)	6.9263 (10.2029)	14.4733 (13.2687)	-1.6571 (14.8329)
$\Delta \log population$ 1939 – 1945	-0.0175**	-0.0002	-0.0171**	0.0003	0.0037	-0.0021
First stage: $\Delta boulevard_i$						
Boulevardized Walls (1500m)	-1.0660**	-1.0414**	-1.0575**	-1.0155**	-0.9779**	-0.9268**
Rubble per capita	0.00003*	0.00003*	0.00002	0.00002		
log first record			0.0002	0.0003	0.0004*	0.0004*
First stage: $\Delta effective street_i$						
Boulevardized Walls (1500m)	-0.0628	-0.0548	-0.0669	-0.0648	-0.0682	-0.0704
Rubble per capita	-0.00000	-0.00000	-0.00000	-0.00000		
log first record			-0.0001**	-0.0001**	-0.0001**	-0.0001**
SW F Stat $\Delta_T boulevard_i$	7.99***	9.05***	5.85***	5.31***	7.48***	7.62***
SW F Stat $\Delta_T street_i$	2.64	2.83*	3.91**	4.11**	7.71***	7.94***
SW Chi $\Delta_T boulevard_i$	9.85***	11.16***	14.68***	13.33***	9.22***	9.39***
SW Chi $\Delta_T street_i$	3.26*	3.49*	9.81***	10.30***	9.51***	9.80***
Panel B: Second stage $\Delta_T \log population_i$ (0 to 1000 meter)						
$\Delta boulevard_i$	-2.3470 (4.0177)	-4.5310*** (1.6676)	-2.3177 (3.9182)	4.3081*** (1.5090)	-2.9424 (2.0629)	-6.6808*** (2.2703)
$\Delta effective street_i$	-74.1399 (56.3629)	17.3631 (21.4616)	-74.2179 (56.1096)	16.6057 (20.8705)	47.8822* (27.8721)	-11.3665 (29.5632)
$\Delta \log population$ 1939 – 1945	-0.0117*	-0.0005	-0.0117*	-0.0003	-0.0001	-0.0012
First stage: $\Delta boulevard_i$						
Boulevardized Walls (1500m)	-0.8256***	-0.8173***	-0.8146***	-0.7872***	-0.8289***	-0.8026***
Rubble per capita	0.00003**	0.00003**	0.00003**	0.00003*	0.00003***	0.00003**
log first record			0.0002**	0.0003**		
First stage: $\Delta effective street_i$						
Boulevardized Walls (1500m)	-0.0504*	-0.0524**	-0.0506*	-0.0529**	-0.0519*	-0.0532**
Rubble per capita	-0.00000**	-0.00000**	-0.00000**	-0.00000*	-0.00000**	-0.00000**
log first record			-0.0000	-0.0000		
SW F Stat $\Delta_T boulevard_i$	13.77***	13.28***	11.04***	11.75***	13.99***	12.28***
SW F Stat $\Delta_T street_i$	9.01***	9.10***	4.71**	4.92**	9.14***	9.04***
SW Chi $\Delta_T boulevard_i$	16.99***	16.38***	27.68***	29.46***	17.26***	15.14***
SW Chi $\Delta_T street_i$	11.11***	11.23***	11.82***	12.34***	11.27***	11.15***
Panel C: Second stage $\Delta_T \log population_i$ (0 to 1500 meter)						
$\Delta boulevard_i$	-9.4986 (7.1698)	-5.3793* (2.8560)	-7.3617 (6.5762)	-5.4878** (2.4774)	-1.2432 (3.5507)	-10.1636** (4.3085)
$\Delta effective street_i$	-97.3635 (61.5903)	8.7807 (25.2064)	-81.1188 (60.5918)	8.0227 (23.9067)	49.9333 (32.9172)	-31.2715 (37.0675)
$\Delta \log population$ 1939 – 1945	-0.0138***	-0.0006	-0.0134**	-0.0006	0.0023	-0.0010
First stage: $\Delta boulevard_i$						
Boulevardized Walls (1500m)	-0.5492***	-0.5345***	-0.5433***	-0.5175***	-0.5417***	-0.5259***
Rubble per capita	0.00003***	0.00003***	0.00002**	0.00002*	0.00002***	0.00002***
log first record			0.0001**	0.0002***		
First stage: $\Delta effective street_i$						
Boulevardized Walls (1500m)	-0.0079	-0.0106	-0.0076	-0.0101	-0.0102	-0.0112
Rubble per capita	-0.00000***	-0.00000***	-0.00000***	-0.00000***	-0.00000***	-0.00000***
log first record			0.0000	0.0002		
SW F Stat $\Delta_T boulevard_i$	11.89***	10.93***	8.93***	9.06***	10.77***	11.16***
SW F Stat $\Delta_T street_i$	15.75***	15.09***	8.90***	8.68***	15.13***	15.00***
SW Chi $\Delta_T boulevard_i$	14.67***	13.48***	22.41***	22.73***	13.28***	13.76***
SW Chi $\Delta_T street_i$	19.43***	18.61***	22.32***	21.77***	18.66***	18.50***

Note: Each regression has 74 observations. Second stage dependent variable: $\Delta_T \log population_i$. The specification of column 1 to 4 are equivalent to the Table 4.4, while the specifications of column 5 to 6 are equivalent to Table 4.5, column 3 and 4 respectively, except that we have not included the $\Delta Herfindahl Orientation$ for any specifications in this table and that we have replaced the $\Delta street_i$ with $\Delta effective street_i$. *, **, *** denotes significance at the 10%, 5%, 1%. Robust standard errors are in parentheses. Note that the difference operator Δ_T measures the $\#T$ long difference of the log population variable for the indicated time period. SW F Stat and SW Chi respectively are the Sanderson & Windmeijer (2016) test for weak- and underidentification respectively. The Stock & Yogo (2005) 10 and 15 percent critical values for two endogeneous variables and instruments are 7.03 and 4.58. The 10, 15, 20 and 25 percent critical values for two endogeneous variables and 3 instruments are 13.43, 8.18, 6.40 and 5.45. The overidentification p-values for the specifications of column 3 for the 500m, 1000m and 1500m are 0.2594, 0.9841 and 0.5266 respectively, while the specifications of column 4 for the 500m, 1000m and 1500m are 0.2985, 0.7124, and 0.9333 respectively.

10 percent critical value for two endogeneous and instruments of 7.03, suggesting that we can reject a weak instrument. We still however find, at most, weak evidence of an impact of an effective street density for the 1966 to 1976 period for the 1000m range. Lastly, we find a robust, but slightly lower estimate for the impact of boulevards between 1976 to the 1986 period, as compared to the estimates of Table 4.5.

4.6 Conclusion

The goal of this paper was to estimate the causal impact of changes in the street network *within* the city on population decentralization. By using topographic maps from before WWII, we were also able to distinguish between normal streets and wider boulevards. Our analysis suggests that the increase in the boulevard density following the second destruction around the 1960s led to large decentralization of the population within German cities of around 6 to 8% within a short period of time. These estimates are similar in magnitude to previous findings within the literature on suburbanization that use highway rays as a measure instead. We also find this development to be temporary but the impact to be permanent, starting around the 1970s until the 1980s, with no detectable impact anymore after the mid 1980s.

We do not find any strong evidence that a change in the street density within a city influences urban growth in any period, which is contrary to some of the evidence by Rauch & Brandily (2020) that Subsaharan African cities with a higher street density grew faster. This may hint at a heterogeneous impact of the street network depending on the level of development and urbanisation. Clearly, more research specifically on developing countries would be welcome here.

Lastly, even though our results show evidence of suburbanization due to these improvements in the within city infrastructure, this attribution was largely driven by the predictions made by the common urban economic models. Stronger empirical evidence that the infrastructure effects are clearly attributable to suburbanization requires that with the lower estimated growth rates within the central city, we should at the same time also find a corresponding higher population growth rates in the suburbs, for which population data is not systematically available. One possible way to get around this issue is to consider the built area as a proxy for population. Again, historical maps during our period of interest could prove to be a valuable source of these historical informations, possibly combined with the use of detailed light density data. A recent, promising initiative by the MapHis: Mapping History project, which uses image analysis to digitizes vast amounts of historical maps, could provide exactly this. However, these would need to mature

into ad hoc processes first to be considered reliable, so that manual digitization may still be the way to go, especially for relatively smaller datasets.

4.A Appendix: Data Digitization

Figure 4.7 shows Kassel, which heavily changed after WWII. We measured every segment of the boulevard in 2020 and compared them to the Meßtischblätter topographic maps of 1934 by measuring if the width are the same in 1934 as today. Hence, we can think of orange lines as the segment of a boulevard that existed already in 1934 i.e. the segments are comparable in width to the 2020 boulevards, while the green lines are the segments that were added after WWII. The added boulevard segments that run across entire housing blocks are easy to spot, notably in the first 500m radius circle. Harder to immediately spot are the added segments which do not run through housing blocks as seen in 1934. For these segments, we have measured the width and assigned it as an existing boulevard segment if this segment is at most 10% less wide than the same segment in 2020. Here, the points of reference to measure are the individual houses across from each other.

As for the street network, as seen in Figure 4.8 for Ulm, we did not rely on pre-existing shapefiles that have already digitized the street network, but assigned the streets by hand ourselves. This was to ensure consistency in assignment between the Meßtischblatt maps and the maps of today. For example, the OSM maps are overly detailed, also considering walks or paths inside housing blocks that we may consider as a part of a backyard. In Lübeck, the 'Gänge' are very narrow walks without access for cars, and resembles someones garden more than a typical street. Those are not depicted within the Meßtischblatt maps, but they are within the shapefiles of the OSM maps. We have also made use of satellite images to determine whether the street as seen within the OSM map depiction of today can be considered a street, or a backyard. However these cases were very rare, and for the vast majority of streets, the assignment was straightforward and obvious.

4.B Appendix: Amalgamation

One issue of the population data are 'Gemeindereform' amalgamations of municipalities in the mid 1970s i.e. the fusion of two or more into one municipality. We know quite well how many people were added as those are often recorded. If not, then we simply take the difference of the population between the year before amalgamation happened with the year after, which seem appropriate as the growth rate between individual years in the 1970s are not far from 0. Similar to Bosker,

Figure 4.7: Kassel Boulevards

The 1934 boulevards are shown in **orange** i.e. those segments of the boulevards that already existed before 1934, while the 2020 boulevard segments that are added are shown in **green**. The inner circle shows the immediate city center. All **green** lines show a change after WWII, where the inner 500m radius circle show that houses were razed and boulevards were run across the city center. The bottom left part show an area where houses were not razed, but houses were widened along the **green** lines. Note that the bottom figure shows the OSM map that show the position of individual houses very accurately, and which were used to compare the width of streets or boulevards between the Meßtischblatt topographic maps of 1934, and the width of the boulevards today. The classification of a boulevard in 2020 were done through the ESRI World Street Map, which also considers the white colored streets Kurfürstenstraße, Rudolf-Schwander-Straße and Ständeplatz as a boulevard, which we would not have considered as such if we were to use the OSM maps.

Figure 4.8: Ulm Street Grid

The 2020 street grid is shown in orange, while the 1934 street grid is shown dark green

Brakman, Garretsen, & Schramm (2008), in that we assume the growth rate of the added municipality to be the same as the one it was added to, which is the city of interest of our sample. The adjusted, estimated population of the added municipality will then be used to subtract it from the official, reported population figures. For amalgamations happening at time T , we write:

$$\hat{pop}_{iT} = pop_{iT} - pop_{inew} \quad (4.5)$$

with pop_{iT} being the population of city i with the added population at time T , pop_{inew} is the (estimated) population of the new areas which we know, pop_{iT-k} is the population in year $T - k$ before the change of the city's boundary. Finally $\hat{pop}_{iT}$ is the adjusted population with the population of the newly added areas being extrapolated in the past. We simply deduct the population from the new area post-WWII at time T . For each subsequent years, as we do not know the

This is a detailed topographic map of Bonn, Germany. The map shows the city's layout, including the Rhine river (Rhein) flowing through it. Key landmarks and features include:

- Geographical Features:** The Rhine river is prominent, with several islands and bends. Contour lines indicate elevations of 500m and 1,000m.
- Urban Layout:** The city is densely packed with buildings and streets. The map shows the city's expansion and the surrounding areas.
- Landmarks:** Labels include 'Bonn', 'Riedhöfe', 'Königsplatz', and 'Ludwigstr.'.
- Infrastructure:** The map shows the railway network (indicated by black lines with cross-ticks) and the road network (indicated by red and orange lines).
- Scale and Orientation:** A scale bar at the bottom right indicates a distance of 2,000m. The map is oriented with North at the top.

(estimated) population of the new areas after amalgamation at $T + k$, we will extrapolate this new area according to:

This procedure should be appropriate for short time periods after, but are less likely to be close to the true population figure the further away in time we adjust the population figures. Also, note that the amalgamation decisions of the 1970s are likely to be uncorrelated with the WWII destruction across cities, as it happened on a voluntary basis between municipalities, and were likely to not be driven by the expansion of the city borders that would require the assimilation of areas. That is, after WWII, there were very few amalgamations until 1974 and it was also a period of very low aggregate urban population growth rates,

Figure 4.10: Bundesfestung Ulm

1934 boulevards derived from walls or bulwarks are shown in purple, the 1934 boulevards are shown in green the new boulevards that exist in 2020 are shown in orange.

which signals that these decisions were mostly driven by politics. Indeed, the main driver was to increase efficiency of the municipal government. This stands in contrast to the amalgamations before WWII, which were larger and likely to be pressured by the relatively large population growth rates especially around the WWI period.²³ Then, adjusting for amalgamation through this procedure, although it would change the aggregate growth rates across cities downwards due to the overall trend of suburbanization, would not bias the estimates of a change in the street network on the growth rates itself as it can be assumed to be a random measurement error in the dependent variable.

²³One indicator for this is that pre-WWII amalgamations were scattered in time from around 1890 to 1940, while the post-WWII amalgamations were more or less around 1975 with little time variation.

Figure 4.11: Orientation Histogram

The top-right figure shows the Remscheid street network in 1934, as well as the histogram of the orientation of streets shown in the top-left figure, which gives us a Herfindahl index of 0.0578 for 18 bins, close to the Herfindahl of random orientations of 0.0555. Respectively, the bottom figures shows the Wilhelmshaven street network in 1934, with a Herfindahl index of 0.1541.

As a robustness exercise, we consider the alternative 2005 constant border population recalculated by DG REGIO (2021), which are only available at a decennial basis from 1960 onwards. Table 4.8 compares the estimates for the population measure in this paper with the DG Regio for 1960 to 2000 at a decennial basis by using the specifications of Table 4.5. The estimated coefficients for the variables which are economically and/or statistically significant are very similar, where we would draw the same conclusion for each period and range regardless of the population measure used.

Table 4.8: IV Estimates: Comparison of Different City Population Measures

time period	1960-1970	1970-1980	1980-1990	1990-2000
Panel A: 0 to 500 meter				
	Dependent variable: $\Delta_T \log population_i$			
$\Delta boulevard_i$	-1.2249 (3.0170)	-4.5180* (2.3856)	-3.7664 (2.4904)	-3.0515* (1.7636)
$\Delta street_i$	0.3743 (1.4678)	1.6541 (1.2140)	-1.0166 (1.2034)	0.2260 (1.2057)
	Dependent variable: $\Delta_T \log DG Regio population_i$			
$\Delta boulevard_i$	-2.5453 (3.2633)	-4.8192* (2.4961)	-3.7100 (2.4810)	-3.2762* (1.8332)
$\Delta street_i$	1.4126 (1.3821)	1.5142 (1.1992)	-0.9158 (1.1843)	0.4679 (1.2100)
SW F Stat $\Delta_T boulevard_i$	11.08***	10.46***	8.78***	14.24***
SW F Stat $\Delta_T street_i$	13.38***	12.81***	13.84***	8.08***
SW Chi $\Delta_T boulevard_i$	13.90***	13.12***	11.01***	18.20***
SW Chi $\Delta_T street_i$	16.78***	16.07***	17.35***	10.32***
Panel B: 0 to 1000 meter				
	Dependent variable: $\Delta_T \log population_i$			
$\Delta boulevard_i$	-0.7160 (3.8829)	-5.9926** (2.4517)	-4.2922 (3.2236)	-4.3065* (2.4491)
$\Delta street_i$	1.0270 (2.7481)	2.0043 (1.8916)	-2.5401 (2.3336)	-0.9508 (2.2615)
	Dependent variable: $\Delta_T \log DG Regio population_i$			
$\Delta boulevard_i$	-2.4438 (3.8046)	-6.3447** (2.8105)	-4.2307 (3.1979)	-4.6248* (2.5156)
$\Delta street_i$	2.5999 (2.365)	1.7412 (2.0037)	-2.3547 (2.2802)	-0.6594 (2.3060)
SW F Stat $\Delta_T boulevard_i$	15.46***	14.08***	13.09***	14.93***
SW F Stat $\Delta_T street_i$	14.23***	15.13***	13.58***	11.84***
SW Chi $\Delta_T boulevard_i$	19.39***	17.66***	16.42***	19.08***
SW Chi $\Delta_T street_i$	17.84***	18.98***	17.03***	15.13***
Panel C: 0 to 1500 meter				
	Dependent variable: $\Delta_T \log population_i$			
$\Delta boulevard_i$	-2.2790 (5.7048)	-7.0905* (3.8546)	-8.4408 (7.1152)	-6.7737 (4.4242)
$\Delta street_i$	-0.0659 (4.1486)	0.9767 (2.6416)	-4.5738 (4.1998)	-1.8281 (3.3380)
	Dependent variable: $\Delta_T \log DG Regio population_i$			
$\Delta boulevard_i$	-3.7269 (5.3749)	-7.9493* (4.4140)	-8.2454 (6.9615)	-7.1277 (4.4903)
$\Delta street_i$	1.6134 (3.6066)	0.2189 (2.9724)	-4.3152 (4.0983)	-1.7552 (3.4064)
SW F Stat $\Delta_T boulevard_i$	7.38***	7.30***	7.51***	6.08**
SW F Stat $\Delta_T street_i$	8.47***	9.74***	7.87***	5.75**
SW Chi $\Delta_T boulevard_i$	9.25***	9.15***	9.41***	7.76***
SW Chi $\Delta_T street_i$	10.63***	12.22***	9.87***	7.35***

Note: Each regression has 74 observations. The specifications are the same as in Table 4.5, with the corresponding year for the state population. *, **, *** denotes significance at the 10%, 5%, 1%. Robust standard errors are in parentheses. Note that the difference operator Δ_T measures the $\#T$ long difference of the log population variable for the indicated time period. SW F Stat and SW Chi respectively are the Sanderson & Windmeijer (2016) test for weak- and underidentification respectively. The Stock & Yogo (2005) 10 and 15 percent critical values for two endogeneous variables and instruments are 7.03 and 4.58.

4.C Appendix: Individual Cities

Table 4.9 presents the cities in our sample that we consider, as well as its population in 1966. Furthermore, we show the the WWII rubble per 1939 capita in each city, the first year the city was mentioned in any record that serves as a proxy for its foundation date, the city rights year where the city can be considered as a city and is actually allowed to even build fortification, as well as hold markets, courts etc. Lastly, we show the earliest year where defortification was mentioned. However, note that when the only source mentions a non definitive time such as 'beginning of 19th century', we take this as the year 1800 here. Furthermore, the year itself may correspond to either the date when defortification started or ended. Here we show the earliest of the two dates unless the starting date was not mentioned anywhere; then we take the latter of the two dates.

Table 4.9: Sample Cities

City	Pop. 1966	rubble p.c.	first rec.	city rights	de- fort.	City	Pop. 1966	rubble p.c.	first rec.	city rights	de- fort.
Aachen	177900	21.2	769	1166	1804	Köln	860200	31.2	50	50	1880
Augsburg	212200	6.3	962	1156	1867	Krefeld	223200	16.1	1097	1373	1819
Bamberg	72200	3.3	902	1306	1803	Leverkusen	106100	2.4	1107	1921	.
B.-Gladbach	47600	0.9	1018	1856	.	Ludwigshafen	177500	14.5	1824	1859	.
Bielefeld	170600	12.8	826	1214	1732	Lübeck	242616	4.5	1050	1160	1806
Bochum	355500	12.1	1041	1321	.	Mainz	149387	13.3	44	297	1921
Bonn	140500	9.4	105	1243	1713	Mannheim	329900	15.1	766	1607	1798
Bottrop	112300	3.7	1092	1919	.	M'gladbach	154200	17.8	1085	1366	1809
Braunschweig	233000	11.5	1031	1227	1802	Moers	49500	1.6	850	1300	1763
Bremen	599000	17.6	786	1186	1803	Mülheim	190878	5.8	1093	1808	.
Bremerhaven	146900	7.3	1830	1851	.	Münster	199300	17.7	819	1170	1764
Darmstadt	139300	26	1100	1330	1800	Neumünster	74200	3.2	1136	1869	.
Dortmund	657100	30.9	890	1152	1810	Neuss	113200	9.1	107	1190	1823
Düsseldorf	698400	18.5	1135	1288	1801	Nürnberg	472300	25.3	1022	1200	.
Duisburg	484000	12.9	883	1279	.	Oberhausen	257900	6.8	1314	1874	.
Erlangen	80300	0.3	1002	1374	.	Offenbach	114605	12.1	977	1698	.
Essen	721200	22.4	870	1244	1800	Oldenburg	128600	1.4	1108	1345	1764
Flensburg	96500	0.3	1196	1284	1796	Osnabrück	129915	17.1	803	1171	1794
Frankfurt	684800	21.1	794	1297	1806	Paderborn	60100	13.7	777	1000	1881
Freiburg	156628	9.4	1091	1120	1800	Pforzheim	87700	24.3	1067	1200	1645
Fürth	96200	2.6	907	1818	.	Recklinghausen	127900	1.3	1017	1236	1834
Gelsenkirchen	363015	12.2	1150	1875	.	Regensburg	125300	0.9	772	1080	1868
Gladbeck	82700	4	900	1919	.	Remscheid	134400	19.7	1168	1808	.
Göttingen	81948	1.2	953	1230	1763	Reutlingen	74900	3.1	1089	1180	1817
Hagen	203000	7.2	1150	1746	.	Saarbrücken	134500	15.0	999	1321	.
Hamburg	1851300	20.9	831	1189	1820	Solingen	175100	10.3	965	1374	.
Hamm	72500	20.3	1188	1213	1763	Stuttgart	559347	8.5	1160	1219	1806
Hannover	547800	17.8	1150	1160	1763	Trier	85500	10.2	50	1190	1888
Heidelberg	124400	0.3	1196	1200	1830	Ulm	91700	12.8	813	1181	1802
Herne	107100	0.7	880	1897	.	Wanne-Eickel	106600	9	850	1926	.
Hildesheim	98500	7.3	864	1249	1761	Wattenscheid	81300	1.5	880	1876	.
Ingolstadt	59601	0.9	806	1250	1937	Wiesbaden	200905	3.1	829	1232	1804
Kaiserslautern	86700	3.6	830	1276	1703	Wilhelmshaven	101400	12.7	1869	1873	.
Karlsruhe	255000	7.4	1715	1722	.	Witten	98398	4.5	1214	1825	.
Kassel	214100	26.7	913	1180	1767	Worms	63700	13.7	496	1184	1850
Kiel	266294	15.1	1232	1242	1700	Würzburg	122200	31.3	704	1069	1868
Koblenz	102600	10.4	150	1214	1899	Wuppertal	422900	18.9	1180	1808	.

Note: Rubble p.c. considers the cubic meter rubble post WWII per 1939 capita, first rec. is the first record of the city name in any record, city rights is the first year a city right was recorded, and de-fort. is the earliest known date of the defortification of the last walls.

Chapter 5

Closing Remarks

This thesis was initially driven by some methodological issues within the literature surrounding Davis & Weinstein (2002) and the central idea: What if locational fundamentals are allowed to change and how did their importance change throughout time? In chapter 2, the main question is which geographical characteristics were exactly fundamental, and which ones still positively impact population today even though these fundamentals have long become obsolete.

The impact of changing fundamentals on urban development may also be seen in a short-term period relative to the very long historical view of chapter 2. One potential source of this short-term change especially during the 20th century may be industrialisation and deindustrialisation, as well as urbanisation and suburbanisation. The worry here is that non-linear developments induced by changing fundamentals may lead to wrong conclusions drawn from impact studies on the population trajectory; an idea which was the main driver of chapter 3, where we took another look at the question whether WWII had any permanent impact on German cities. More concretely, we have looked at a shock on the fundamentals itself in the form of post-war changes in the German inner city street infrastructure during the 1960s and 70s, and whether these changes led to further suburbanization effects. The following will summarize each chapter and discuss ideas for new avenues of research.

In chapter 2, we have expanded the notion of first-nature geography and specifically emphasized man-made features as well as the 'first' part in first-nature geography to get closer to the initial geographical conditions of a location. We have tackled the measurement issue by considering place names (toponyms) that convey this kind of first-nature geography. The source of these place names also provide the year it was mentioned, and together with the year its city rights were awarded gives us a simple proxy for a location's quality of land, or geographical characteristics if you will. With this unique dataset, we contribute to several strands of

the literature: First, as we have vastly expanded the variety of geographical characteristics that can be considered through toponyms, we show that the frequency of geographical characteristics as suggested by its place name is Zipf distributed. Second, the simple proxy for land quality also follows a Zipf distribution, which we take as evidence in favor of a Zipf's law for geographical characteristics that was proposed by Krugman (1996); and which offers an alternative to the generation of a Zipf distribution through some form of Gibrat's law. Lastly, we show that characteristics that are associated with water-based trade were not as important very early in the development phase of a city, only later. However, land-based trade and castles were important at the beginning, with places that had castles very early on still being larger even today although they are not relevant anymore, which suggest strong path dependence of castle places.

In chapter 3, we have had another look at Davis & Weinstein (2002), as well as Brakman et al. (2004) and Bosker, Brakman, Garretsen, & Schramm (2008) on the second-nature geography impact of the WWII bombings. The initial two papers found no permanent impact of WWII on the relative city size in Japan and West Germany respectively, while the latter does for West Germany. We have conjectured that this discrepancy may be due the longer time span considered in the latter paper, where simultaneous developments such as the decline of coal and the rise of automotive commuting led to a decline of the city population from the 1970s onwards; a timespan only the latter of those papers covers. If this is the case, then the current method that does not explicitly attempt to disentangle these possible simultaneous developments from the WWII bombing effect may wrongly attribute the impact of these simultaneous developments to the WWII bombings.

In order to explain the discrepancy, and motivated by this conjecture, we have constructed counterfactual (synthetic) German cities following Abadie et al. (2010), directly aimed at disentangling the WWII long-term impact from all other developments. Then, we find no permanent impact i.e. reversion to the synthetic mean, for the majority of German cities even when we consider the time span from the 1970s onwards, which can reconcile the discrepancy we have initially found. However, we also find that smaller and remote cities typically did not recover, whereas we find a positive and permanent population impact in bigger and less remote cities. We interpret this as the WWII impact accelerating the move towards more agglomeration, indicating the importance of increasing returns and second-nature geography.

In chapter 4, we have looked at the impact of WWII on the first-nature geography of the inner city street infrastructure, after a few other papers following Davis & Weinstein (2002) looked at whether there is a disrupting and permanent impact on second-nature geography. We have distinguished between normal streets

and wider boulevard infrastructure, where we measured the post-WWII suburbanization impact of a change of the street network infrastructure towards a more car-friendly inner city. We contribute to the literature in considering the street network as inner city transportation infrastructure instead of highways that are found outside the city, or subways that only apply for a few very large agglomerations so that the estimates may be relevant for many cities around the world that are too small for their own subway system. We have also instrumented infrastructure changes with WWII destruction and walls converted to boulevards. Overall, our estimates suggest a suburbanization effect of the additional boulevards of around 6 to 8 percent, depending on the range definition of the inner city, while we find no effect of streets of normal width.

We can also tell a few more things about the 'historical roots of cities'. For one, rivers were not important for the quick development towards full-fledged Medieval cities. This suggests that they are not the deepest root nor the most important first-nature geography at least initially. Instead, land-based trade and security were more important characteristics in this initial phase of settlement than water-based trade, which only became important later. Hence, the focus on trade and the disregard for other first-nature characteristics may have been misguided by the literature on geography and the historical location of economic activity. We also find a different behavior of second-nature effects after the WWII shock not by a population movement being uprooted towards medium-sized cities (Bosker, Brakman, Garretsen, & Schramm, 2008), but in line with the increasing returns interpretation by Davis & Weinstein (2002) towards even more agglomeration.

However, apart from the novel findings within the three chapters, there may be some other contributions and implications that can be found here that do not fit within the individual chapters, and that were accumulated throughout the years.

For one, the idea of using place names was born out of the emphasis to get as close to the initial conditions of the early settlement as possible, while considering both natural and man-made features. Especially initial man-made features were not considered, thus providing an alternative or complementary way of measuring these geographical initial conditions that may be useful to the larger literature.

One issue of urban growth models is that they rely on some form of Gibrat's law to generate a Zipf distribution, where Gibrat's law itself has no economic explanation unless one is imposed, such as random technology shocks. One critique here is that this in itself adds too much freedom of choice in the economic mechanism that governs Gibrat's law, where anything goes as long as it somehow generates Gibrat's law without being too far-fetched in its mechanism. That a measure of the time it takes from first record of a place's name until its city right awarded itself is Zipf distributed adds some consistency to Zipf's law on cities, and is importantly

not ex-post chosen to fit Gibrat's law in order to generate this distribution, adding restrictions on the mechanism. Our findings support alternative ways to model a Zipf distribution e.g. Lee & Li (2013).

Further, the generation of a Zipf distribution by assuming locational fundamentals to be Zipf distributed as well instead of through some form of Gibrat's law may not change the dynamics within a model if the generation of a Zipf distribution is the sole purpose of introducing a Gibrat's law mechanism. However, it may change the policy implications due to the very different nature of their underlying mechanisms. Future research may be more careful in selecting their modelling assumptions going forward. Furthermore, information on both city rights and first record of place name are easily available for almost all European cities, thus making it possible to scale this dataset even further. So far for example, current datasets on early city economic activity were restricted on already developed, larger cities e.g. Bairoch et al. (1988) during the late Middle Ages, which stands in contrast to the 'foundation date'/first record of most settlements around half a millenia earlier. A rule of thumb is that historical data is sparser the further back in time we would like to measure, that for measures on economic activity would require plausible proxies of such data (see Nagy (2021) for an overview).

Data on cities during the antiquity that for example consider the number of archeological sites or stone tablet data as information sources for local economic activity (Davis & Weinstein, 2002; Bakker et al., 2019; Barjamovic et al., 2019) on the other hand may be too far away in time as especially most Middle-European cities were only founded during the Middle Ages, with very few that were already founded during antiquity.

This would make our type of data that primarily considers the Middle Ages, in this light, the most relevant to elicit on the origins of most contemporary cities. Especially in recent years, the urban and geographical economics literature relied on Bairoch et al. (1988) type of datasets on the historical development of population of cities Hanlon & Heblich (2020), so the hope is that chapter 2 offers the literature an alternative or complementary way to measure economic activity, that also covers the very early days in the development of cities that were previously thought to be inaccessible.

The limitation here is that it is not straight-forward how the relationship of this geographical advantage measure is with the more straight-forward measures of city population. One may need to impose some structure to recover relative population estimates, similar to Barjamovic et al. (2019), thus giving us time-consistent estimates of relative population over very long time spans. With these estimates, we may be able to provide a richer picture on the evolution of importance of different geographical circumstances throughout history, as well as the interplay

of technology and geography and its changing impact on local economic activity in the long-run. Nevertheless, due to the sparsity of information going backwards in time, we would rely on stronger assumptions than simple population estimates for cities over time which will, going forward, need to gain convincing as well.

On a similar note, chapter 4 uses historical maps to gather information on the street network just before WWII. This was done by hand here, but there have been recent developments that use machine learning methods to digitize these maps (Combes et al., 2020), with a focus however for very old maps. The next evolution of studies on the local economic effects of urban transportation infrastructure projects may make use of these new machine learning methods in digitizing vast quantities of maps to register more subtle changes throughout time.

We may also want to combine the recently popular light density data (Henderson et al., 2012) with these other means of measuring local economic activity, for example through aerial photographs that were done quite frequently reaching back to the time after WWII, as well as historical maps that reaches further back and that we have used in chapter 3. The goal is then to make these different types of top-view snapshots consistent with each other, which is possible if these snapshots overlap at some point in time. Then using the high serial correlation that exists for short time spans between these snapshots, which is less so the case for larger time spans, we can again paint a picture on how local economic activity changed throughout the 19th to 20th century. This would allow us to not only register changes in the infrastructure between the topographic maps of 1933 and today, but also in between these dates.

For Germany in chapter 4 specifically, for which annual suburban population data is missing, we could use this to estimate the effect of urban infrastructure on suburban development and land-use changes at a higher frequency. This would give us more precise treatment dates for each individual city, and not as we have done by giving an estimate between the 1960s and 1970s as a treatment data for all cities. This may allow us to better understand the dynamics of such changes in the urban infrastructure on local economic activity, as well as how far reaching they are on suburban development and land-use changes.

Another channel of interest on the impact of inner city infrastructure is through inequality. Although boulevards can still be crossed quite easily, this is not so much the case for highways that run through a city, leading to segregation similar to railroads Ananat (2011) or the Berlin wall (Ahlfeldt et al., 2015). Of interest here may be the endogeneous provision of amenities that may then relatively favor one side over another. The location of these amenities may be determined by the aforementioned newly acquired data through machine learning processes.

As for some wider implications of chapter 4, it is difficult to compare the impact

of different types of infrastructure projects, especially as current estimates come from very different regions in the world. Also, the scale of such projects may differ case by case, although they may act as a guiding light for policy makers in this case. The natural evolution from here is to compare the inner city infrastructure impact to the impact of highways outside the city for the same region, or that of other infrastructure projects that aim to reduce commuting costs; another is to estimate the cost-effectiveness of each type of infrastructure. Hence, in isolation, the results of chapter 4 may not be as useful in informing most policy makers right now, except that it suggests suburbanization effects. However, cities in developing countries that in any case would modernize the housing stock or commercial properties within the inner cities may want to coordinate this to allow for wider streets. The cost of razing to make room for transportation infrastructure would then reduce to only the cost of coordination.

One implication of the findings of chapter 3 is that we can reconcile the literature in whether a large shock on the city system will have a permanent effect, which is of interest to policymakers as a change in the relative city size equilibrium may also reinforce existing regional disparities. Of the few cities for which we find a negative impact on the relative city size, those are mostly remote and smaller cities. This may not point towards a permanent impact of WWII per se, but that the bombing shock may have accelerated developments that would have happened in any case i.e. a move of the population towards agglomerating areas. However, this still remains in the realm of speculation and it would be another interesting area of research.

Finally, the bombing shock was widespread, affecting many regions and cities. However, there had been estimates of a permanent impact of disasters that were more local, only affecting a single region, city or neighborhood (duPont IV & Noy, 2015; Siodla, 2015). Furthermore, following Krugman (1991), the idea of historical accidents and the idea of path dependence is still pervasive, at least at a local level. It could be that how widespread a catastrophe is may play a role in whether we would expect to find a permanent or a temporary impact. For example, if all cities are equally affected, there is no reason for firms or people to move to other cities. On the other hand, if only one city is shocked, then we would expect movement away to other cities. In this regard, whether a shock is widespread or local would matter here, which may explain the tendency of finding a permanent impact for local shocks, and a temporary impact for more widespread shocks. Whether this simple explanation can reconcile these different findings is not clear here, and it would be another interesting avenue of research that is worthy of investigation.

Overall, we aimed to contribute towards understanding the spatial distribution of activity within single urban areas or entire regions, past or present, where the

focus was also on tackling some important methodological issues on measurement or the disentanglement of effects. The hope is that the new ideas, approaches, measures, and instruments of this thesis will prove to be useful for subsequent studies, with the different angle taken here to also provide a novel perspective to the established literature.

Bibliography

- Abadie, A. (2019). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*.
- Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017). *When should you adjust standard errors for clustering?* (Working Paper No. 24003). National Bureau of Economic Research.
- Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, 105(490), 493–505.
- Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2), 495–510.
- Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. *American Economic Review*, 93(1), 113–132.
- Acemoglu, D., Johnson, S., & Robinson, J. (2005). The rise of Europe: Atlantic trade, institutional change, and economic growth. *American Economic Review*, 95(3), 546–579.
- Ahlfeldt, G. M., Albers, T., & Behrens, K. (2020). Prime locations. *CEPR Discussion Paper No. DP15470*.
- Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., & Wolf, N. (2015). The economics of density: Evidence from the berlin wall. *Econometrica*, 83(6), 2127–2189.
- Allen, T., & Donaldson, D. (2020). *Persistence and path dependence in the spatial economy* (Working Paper No. 28059). National Bureau of Economic Research.
- Alonso, W., et al. (1964). Location and land use. toward a general theory of land rent. *Cambridge, Mass.: Harvard Univ. Pr.*

- Alsan, M. (2015). The effect of the tsetse fly on African development. *American Economic Review*, 105(1), 382–410.
- Ananat, E. O. (2011). The wrong side (s) of the tracks: The causal effects of racial segregation on urban poverty and inequality. *American Economic Journal: Applied Economics*, 3(2), 34–66.
- Armenter, R., Koren, M., & Nagy, D. K. (2014). bridges. *Unpublished Manuscript*.
- Bairoch, P., Batou, J., & Chevre, P. (1988). La population des villes Europeennes, 800-1850. *Librairie Droz, Geneve*.
- Bakker, J., Maurer, S., Pischke, J., & Rauch, F. (2019). Of mice and merchants: connectedness and the location of economic activity in the Iron Age. *Review of Economics and Statistics*.
- Barjamovic, G., Chaney, T., Coşar, K., & Hortaçsu, A. (2019). Trade, merchants, and the lost cities of the bronze age. *The Quarterly Journal of Economics*, 134(3), 1455–1503.
- Baum-Snow, N. (2007). Did highways cause suburbanization? *The Quarterly Journal of Economics*, 122(2), 775–805.
- Baum-Snow, N. (2010). Changes in transportation infrastructure and commuting patterns in US metropolitan areas, 1960-2000. *American Economic Review*, 100(2), 378–82.
- Baum-Snow, N. (2020). Urban transport expansions and changes in the spatial structure of us cities: Implications for productivity and welfare. *Review of Economics and Statistics*, 102(5), 929–945.
- Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., & Zhang, Q. (2017). Roads, railroads, and decentralization of Chinese cities. *Review of Economics and Statistics*, 99(3), 435–448.
- Berliant, M., & Konishi, H. (2000). The endogenous formation of a city: population agglomeration and marketplaces in a location-specific production economy. *Regional Science and Urban Economics*, 30(3), 289–324.
- Bernhardt, C. (2017). Längst beerdigt und doch quicklebendig: Zur widersprüchlichen Geschichte der 'autogerechten Stadt'. *Zeithistorische Forschungen/Studies in Contemporary History*, 14(3), 526–540.
- Bleakley, H., & Lin, J. (2012). Portage and path dependence. *The Quarterly Journal of Economics*, 127(2), 587–644.

- Boeing, G. (2019). Urban spatial order: Street network orientation, configuration, and entropy. *Applied Network Science*, 4(1), 1–19.
- Bosker, M., Brakman, S., Garretsen, H., De Jong, H., & Schramm, M. (2008). Ports, plagues and politics: explaining Italian city growth 1300–1861. *European Review of Economic History*, 12(1), 97–131.
- Bosker, M., Brakman, S., Garretsen, H., & Schramm, M. (2007). Looking for multiple equilibria when geography matters: German city growth and the WWII shock. *Journal of Urban Economics*, 61(1), 152–169.
- Bosker, M., Brakman, S., Garretsen, H., & Schramm, M. (2008). A century of shocks: the evolution of the German city size distribution 1925–1999. *Regional Science and Urban Economics*, 38(4), 330–347.
- Bosker, M., & Buringh, E. (2017). City seeds: Geography and the origins of the European city system. *Journal of Urban Economics*, 98, 139–157.
- Bosker, M., Buringh, E., & Van Zanden, J. L. (2013). From Baghdad to London: Unraveling urban development in Europe, the Middle East, and North Africa, 800–1800. *Review of Economics and Statistics*, 95(4), 1418–1437.
- Brakman, S., Garretsen, H., & Schramm, M. (2004). The strategic bombing of German cities during World War II and its impact on city growth. *Journal of Economic Geography*, 4(2), 201–218.
- Brakman, S., Garretsen, H., Van Marrewijk, C., & Van Den Berg, M. (1999). The return of Zipf: Towards a further understanding of the rank-size distribution. *Journal of Regional Science*, 39(1), 183–213.
- Brendler, S. (2008). Sprachliche Orientierung im Raum: Zur Notwendigkeit der Ortsnamen. *Linguence Hamburg*.
- Brueckner, J. K., Thisse, J.-F., & Zenou, Y. (1999). Why is central Paris rich and downtown Detroit poor?: An amenity-based theory. *European Economic Review*, 43(1), 91–107.
- Bundesamt für Statistik, Schweiz. (2016). Bevölkerungsentwicklung der Gemeinden 1850–2000.
- Burchfield, M., Overman, H. G., Puga, D., & Turner, M. A. (2006). Causes of sprawl: A portrait from space. *The Quarterly Journal of Economics*, 121(2), 587–633.

- Chandra, A., & Thompson, E. (2000). Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system. *Regional Science and Urban Economics*, 30(4), 457–490.
- Chaney, E. (2013). Revolt on the Nile: Economic shocks, religion, and political power. *Econometrica*, 81(5), 2033–2053.
- Chaney, E. (2020). Modern library holdings and historic city growth. *CEPR Discussion Paper No. DP14686*.
- Combes, P.-P., Duranton, G., Gobillon, L., & Roux, S. (2010). Estimating agglomeration economies with history, geology, and worker effects. In *Agglomeration economics* (pp. 15–66). University of Chicago Press.
- Combes, P.-P., Gobillon, L., & Zylberberg, Y. (2020). Urban economics in a historical perspective: Recovering data with machine learning. *CEPR Discussion Paper No. DP15308*.
- Cronon, W. (1991). *Nature's metropolis : Chicago and the Great West*. New York :W. W. Norton.
- Danmarks Statistik. (1985). Danske byers folketal 1801-1981.
- Danmarks Statistik. (2011). Statbank Denmark folketal.
- Davis, D. R., & Weinstein, D. E. (2002). Bones, bombs, and break points: The geography of economic activity. *American Economic Review*, 92(5), 1269–1289.
- Davis, D. R., & Weinstein, D. E. (2008). A search for multiple equilibria in urban industrial structure. *Journal of Regional Science*, 48(1), 29–65.
- Defence Analytical Services and Advice. (2011). Total strength, UK service personnel. *Institute for fiscal studies*.
- DG REGIO. (2021). Population data collection for European local administrative units from 1960 onward. *Directorate-General for Regional and Urban Policy, European Commission*.
- Dittmar, J. (2011). Cities, markets, and growth: the emergence of Zipf's law. *Unpublished Manuscript*.
- Donaldson, D. (2018). Railroads of the Raj: Estimating the impact of transportation infrastructure. *American Economic Review*, 108(4-5), 899–934.
- Donaldson, D., & Hornbeck, R. (2016). Railroads and American economic growth: A “market access” approach. *The Quarterly Journal of Economics*, 131(2), 799–858.

- duPont IV, W., & Noy, I. (2015). What happened to Kobe? a reassessment of the impact of the 1995 earthquake in Japan. *Economic Development and Cultural Change*, 63(4), 777–812.
- Duranton, G. (2006). Some foundations for zipf’s law: Product proliferation and local spillovers. *Regional science and urban Economics*, 36(4), 542–563.
- Duranton, G., Morrow, P. M., & Turner, M. A. (2014). Roads and Trade: Evidence from the US. *The Review of Economic Studies*, 81(2), 681.
- Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In *Handbook of regional and urban economics* (Vol. 4, pp. 2063–2117). Elsevier.
- Duranton, G., & Puga, D. (2014). The growth of cities. In *Handbook of economic growth* (Vol. 2, pp. 781–853). Elsevier.
- Duranton, G., & Turner, M. A. (2012). Urban growth and transportation. *Review of Economic Studies*, 79(4), 1407–1440.
- Dye, A. V. (1921). Coal production in the United Kingdom. *US Commerce Reports, Volume 1, Issue 3*.
- Eeckhout, J. (2004). Gibrat’s law for (all) cities. *American Economic Review*, 94(5), 1429–1451.
- Fischer, W. (1989). Statistik der Bergbauproduktion Deutschlands 1850 – 1914. *GESIS Datenarchiv, Köln. histat. Studiennummer 8448, Datenfile Version 1.0.0*.
- Fischer, W. (1995). Statistik der Montanproduktion Deutschlands 1915 – 1985. *GESIS Datenarchiv, Köln. histat. Studiennummer 8400, Datenfile Version 1.0.0*.
- Flückiger, M., Hornung, E., Larch, M., Ludwig, M., & Mees, A. (2019). Roman transport network connectivity and economic integration. *CEPR Discussion Paper No. DP13838*.
- Friedrich, J. (2002). Der Brand; Deutschland im Bombenkrieg 1940-1945. *Munich, Propylaen*.
- Gabaix, X. (1999a). Zipf’s law and the growth of cities. *American Economic Review*, 89(2), 129–132.
- Gabaix, X. (1999b). Zipf’s law for cities: an explanation. *The Quarterly Journal of Economics*, 114(3), 739–767.

- Gabaix, X., & Ioannides, Y. M. (2004). The evolution of city size distributions. In *Handbook of regional and urban economics* (Vol. 4, pp. 2341–2378). Elsevier.
- Garcia-López, M.-À., Holl, A., & Viladecans-Marsal, E. (2015). Suburbanization and highways in Spain when the Romans and the Bourbons still shape its cities. *Journal of Urban Economics*, 85, 52–67.
- GB Historical GIS / University of Portsmouth. (2021). Population Statistics — Total Population, A Vision of Britain through Time.
- Gibbons, S., Lyytikäinen, T., Overman, H. G., & Sanchis-Guarner, R. (2019). New road infrastructure: the effects on firms. *Journal of Urban Economics*, 110, 35–50.
- Giesen, K., & Südekum, J. (2011). Zipf’s law for cities in the regions and the country. *Journal of Economic Geography*, 11(4), 667–686.
- Giesen, K., & Suedekum, J. (2014). City age and city size. *European Economic Review*, 71, 193–208.
- Glaeser, E. L. (2005). Reinventing Boston: 1630–2003. *Journal of Economic Geography*, 5(2), 119–153.
- Glaeser, E. L., & Kahn, M. E. (2004). Sprawl and urban growth. In *Handbook of regional and urban economics* (Vol. 4, pp. 2481–2527). Elsevier.
- Gonzalez-Navarro, M., & Turner, M. A. (2018). Subways and urban growth: Evidence from earth. *Journal of Urban Economics*, 108, 85–106.
- Hanlon, W. W., & Heblich, S. (2020). History and Urban Economics. *National Bureau of Economic Research Working Paper Series*(w27850).
- Harari, M. (2020). Cities in Bad Shape: Urban Geometry in India. *American Economic Review*, 110(8), 2377–2421.
- Heblich, S., Redding, S. J., & Sturm, D. M. (2020). The making of the modern metropolis: evidence from London. *The Quarterly Journal of Economics*, 135(4), 2059–2133.
- Heinzelmann, M. (2003). Göttingen im Luftkrieg. *Die Werkstatt, Göttingen*.
- Henderson, J. V., Squires, T., Storeygard, A., & Weil, D. (2017). The global distribution of economic activity: Nature, history, and the role of trade. *The Quarterly Journal of Economics*, 133(1), 357–406.

- Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *American economic review*, 102(2), 994–1028.
- Hornbeck, R. (2012). The enduring impact of the American Dust Bowl: Short- and long-run adjustments to environmental catastrophe. *American Economic Review*, 102(4), 1477–1507.
- Im, K. S., Pesaran, M. H., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, 115(1), 53–74.
- Instituto Nacional de Estadística. (2011). Censos 1878-2011.
- Ioannides, Y. M., & Overman, H. G. (2003). Zipf’s law for cities: an empirical examination. *Regional Science and Urban Economics*, 33(2), 127–137.
- Ioannides, Y. M., & Overman, H. G. (2004). Spatial evolution of the US urban system. *Journal of Economic Geography*, 4(2), 131–156.
- Jedwab, R., Johnson, N. D., & Koyama, M. (2020). Medieval cities through the lens of urban economics. *Regional Science and Urban Economics*, 103598.
- Jedwab, R., Kerby, E., & Moradi, A. (2017). History, path dependence and development: Evidence from colonial railways, settlers and cities in Kenya. *The Economic Journal*, 127(603), 1467–1494.
- Kaul, A., Klößner, S., Pfeifer, G., & Schieler, M. (2015). Synthetic control methods: Never use all pre-intervention outcomes together with covariates. *MPRA Paper 83790, University Library of Munich, Germany.*
- Konishi, H. (2000). Formation of hub cities: Transportation cost advantage and population agglomeration. *Journal of Urban Economics*, 48(1), 1–28.
- Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, 99(3), 483–499.
- Krugman, P. (1996). Confronting the mystery of urban hierarchy. *Journal of the Japanese and International economies*, 10(4), 399–418.
- Küntzel, T. (2004). Stadt und Grenze-die Landwehr der Stadtwüstung Nienover im südniedersächsischen Kontext. *Archaeologia historica.*, 29(1), 167–191.
- Kästner, F. (1949). Kriegsschäden (Trümmermengen, Wohnungsverluste, Grundsteuerausfall und Vermögensteuerausfall. *Statistisches Jahrbuch Deutscher Gemeinden*, 37: 361-391.

- Lauer, J. (2020). Das "Bombenloch" blieb noch Jahrzehnte. *Rhein-Neckar Zeitung*.
- Lee, S., & Li, Q. (2013). Uneven landscapes and city size distributions. *Journal of Urban Economics*, 78, 19–29.
- Lee, S., & Lin, J. (2018). Natural amenities, neighbourhood dynamics, and persistence in the spatial distribution of income. *The Review of Economic Studies*, 85(1), 663–694.
- Levin, A., Lin, C.-F., & Chu, C.-S. J. (2002). Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics*, 108(1), 1–24.
- Mayer, T., & Trevien, C. (2017). The impact of urban public transportation evidence from the Paris region. *Journal of Urban Economics*, 102, 1–21.
- Michaels, G., & Rauch, F. (2017). Resetting the urban network: 117–2012. *The Economic Journal*, 128(608), 378–412.
- Miguel, E., & Roland, G. (2011). The long-run impact of bombing Vietnam. *Journal of Development Economics*, 96(1), 1–15.
- Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *The American Economic Review, Papers and Proceedings*, 57(2), 197–210.
- Mintzker, Y. (2012). *The Defortification of the German City, 1689-1866*. Cambridge University Press.
- Muth, R. F. (1969). Cities and housing; the spatial pattern of urban residential land use. *Third Series: Studies in Business and Society, The University of Chicago Press*.
- Nagy, D. K. (2021). Quantitative economic geography meets history: Questions, answers and challenges. *Regional Science and Urban Economics*, 103675.
- National Archives UK. (2021). Ministry of Home Security: Daily Intelligence Reports.
- Ng, S., & Perron, P. (1995). Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association*, 90(429), 268–281.
- Niemeyer, M. (2012). Deutsches Ortsnamenbuch. *De Gruyter*.
- Nitsch, V. (2005). Zipf zipped. *Journal of Urban Economics*, 57(1), 86–100.

- Nunn, N., & Puga, D. (2012). Ruggedness: The blessing of bad geography in Africa. *Review of Economics and Statistics*, 94(1), 20–36.
- Oto-Peralías, D. (2017). What do street names tell us? the ‘city-text’ as socio-cultural data. *Journal of Economic Geography*, 18(1), 187–211.
- Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica: Journal of the Econometric Society*, 1361–1401.
- Perron, P. (1991). Test consistency with varying sampling frequency. *Econometric Theory*, 7(3), 341–368.
- Porter, S. (1998). *The Great Fire of London*. Sutton.
- Rappaport, J., & Sachs, J. D. (2003). The United States as a coastal nation. *Journal of Economic Growth*, 8(1), 5–46.
- Rauch, F., & Brandily, P. (2020). Within-city roads and urban growth. *Unpublished Manuscript*.
- Redding, S. J. (2010). The empirics of new economic geography. *Journal of Regional Science*, 50(1), 297–311.
- Redding, S. J., & Sturm, D. M. (2008). The costs of remoteness: Evidence from German division and reunification. *American Economic Review*, 98(5), 1766–97.
- Redding, S. J., Sturm, D. M., & Wolf, N. (2011). History and industry location: evidence from German airports. *Review of Economics and Statistics*, 93(3), 814–831.
- Redding, S. J., & Turner, M. A. (2015). Transportation costs and the spatial organization of economic activity. In *Handbook of regional and urban economics* (Vol. 5, pp. 1339–1398). Elsevier.
- Rossi-Hansberg, E., & Wright, M. L. (2007). Urban structure and growth. *The Review of Economic Studies*, 74(2), 597–624.
- Sanderson, E., & Windmeijer, F. (2016). A weak instrument F-test in linear IV models with multiple endogenous variables. *Journal of Econometrics*, 190(2), 212–221.
- Schilling, H. (2004). *Die Stadt in der Frühen Neuzeit*. De Gruyter Oldenbourg.
- Sensch, J. (2007). histat-Datenkompilation online: Bevölkerungsstand, Bevölkerungsbewegung, Haushalte und Familien in der Bundesrepublik Deutschland 1947 bis 1999.

- Siodla, J. (2015). Razing San Francisco: The 1906 disaster as a natural experiment in urban redevelopment. *Journal of Urban Economics*, 89, 48–61.
- Soo, K. T. (2005). Zipf's law for cities: a cross-country investigation. *Regional Science and Urban Economics*, 35(3), 239–263.
- Spiegel, D. (1959). Das Wunder von Hannover. *Spiegel Verlag, Hamburg*.
- Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear iv regression. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, 80.
- Stoob, H. (1985). *Die Stadt: Gestalt und Wandel bis zum industriellen Zeitalter*. Böhlau Verlag, Köln.
- Storeygard, A. (2016). Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa. *The Review of Economic Studies*, 83(3), 1263–1295.
- Swärd, C. (2016). Folknet Database 1810-1990.
- Taylor, M. P., Peel, D. A., & Sarno, L. (2001). Nonlinear mean-reversion in real exchange rates: toward a solution to the purchasing power parity puzzles. *International Economic Review*, 42(4), 1015–1042.
- Tessin, W., Knorr, T., Pust, C., & Birlem, T. (1983). Umsetzung und Umsetzungsfolgen in der Stadtsanierung: Die individuellen Auswirkungen erzwungener Mobilität.
- Tucker, S. C. (2003). *The Second World War*. Red Globe Press.
- Tuckermann, W. (1906). *Das Gewerbe der Stadt Hildesheim bis zur Mitte des fünfzehnten Jahrhunderts*. E. Ebering Berlin.
- Villette, J., & Purves, R. (2018). Exploring microtoponyms through linguistic and geographic perspectives. *21th AGILE Conference on Geographic Information Science, Lund*.
- Villette, J., & Purves, R. S. (2019). From Microtoponyms to Landscape Using Semantics, Location, and Topography: The Case of Wald, Holz, Riet, and Moos in St. Gallen, Switzerland. *The Professional Geographer*, 1–12.
- Volkman, H.-E. (1995). Ende des Dritten Reiches, Ende des Zweiten Weltkriegs: Eine Perspektivische Rückschau. *Piper*.

- Zhang, N. (2011). *Stadterneuerung in China - von der Flächensanierung zur behutsamen Stadterneuerung* (Unpublished doctoral dissertation). Universität Stuttgart.

Nederlandse samenvatting

Vaak moet je naar het verleden kijken om het heden te kunnen begrijpen. Dit is zeker waar als je kijkt naar oude steden. Veel van onze huidige steden zijn in een ver verleden als kleine nederzetting ontstaan. Waarom zijn sommige van die dorpjes uitgegroeid tot steden en zijn andere klein gebleven? Is het een vast gegeven dat de grootste steden altijd groot blijven, of kan een grote ramp een stad ten onder laten gaan en nieuwe steden doen opkomen? En kunnen beleidsmaatregelen om de eeuwenoude transportinfrastructuur van een stad te verbeteren de grotere vraag naar binnenstedelijk transport in een meer verstedelijkte en gemotoriseerde wereld helpen verlichten? In drie empirische hoofdstukken beschrijven en gebruiken wij in dit proefschrift een aantal nieuwe methodologische benaderingen voor deze en andere vragen over de historische oorsprong van steden.

In hoofdstuk 2 kijken we naar het landschap van vroegmiddeleeuwse nederzettingen in en rondom het huidige Duitsland, omdat de meeste huidige steden rond die tijd zijn ontstaan. We kijken naar de rol van verschillende geografische omstandigheden waar nederzettingen zich in de middeleeuwen en de moderne tijd al dan niet hebben ontwikkeld tot steden, en naar welke omstandigheden hun relevantie hebben behouden. Dit doen we via een unieke benadering: we beschouwen plaatsnamen (toponiemen) als bron van geografische informatie die de directe geografische omgeving in de

beginfase van een nederzetting weergeeft en ons daarmee een unieke blik op het verleden gunt. Bovendien bieden deze toponiemen inzicht in de meer diverse geografische kenmerken rondom een nederzetting. We laten zien dat nederzettingen in de buurt van geografische kenmerken die bevorderlijk zijn voor handel via land, en nederzettingen in de buurt van kastelen, zich in de middeleeuwen sneller ontwikkelden tot steden dan nederzettingen in de buurt van kenmerken die worden geassocieerd met handel via waterwegen. Pas later werden kenmerken van watertransport belangrijker, zoals te zien is aan het formaat van deze steden in modernere tijden.

In hoofdstuk 3 richten we ons op recentere tijden. De verwoesting van Duitse steden in de Tweede Wereldoorlog biedt ons een unieke mogelijkheid om te zien

of de inwonersaantallen van steden zich na grootschalige vernietiging uiteindelijk weer herstellen of voor altijd aangetast zijn. Dit hoofdstuk is niet alleen interessant voor economen die steden onderzoeken, maar vooral ook voor beleidsmakers in regio's die door oorlog of natuurrampen zijn geteisterd. Wat ons onderzoek anders maakt dan eerdere studies, is dat wij ons meer richten op mogelijke andere ontwikkelingen die zich ook na de Tweede Wereldoorlog voordeden, zoals het afnemende gebruik van steenkool en de toenemende motorisering. Wij houden rekening met dit soort mogelijke ontwikkelingen en bieden verklaringen voor sommige afwijkende conclusies die in de literatuur worden getrokken over de wederopbouw van steden na grote rampen. Wij concluderen dat de meeste Duitse steden inderdaad geen blijvende schade hebben ondervonden, maar dat er wel een verschuiving heeft plaatsgevonden in de richting van grotere steden in agglomeraties ten koste van steden in meer afgelegen gebieden.

In hoofdstuk 4 kijken we ten slotte in hoeverre de Tweede Wereldoorlog heeft bijgedragen aan het ontsluiten van de eeuwenoude infrastructuur van steden. In het bijzonder onderzoeken we de invloed die de verschuiving richting een autovriendelijke binnenstad na de oorlog heeft gehad op suburbanisatie. Grootschalige verwoesting vond niet alleen plaats tijdens de Tweede Wereldoorlog, maar ook vele jaren later tijdens de 'Flächensanierung' (gebiedssanering) waarmee in West-Duitsland ruimte werd gemaakt voor bredere straten om plaats te bieden aan het toenemende gemotoriseerde verkeer. hoe meer er in de oorlog was verwoest, hoe makkelijker planologen het publiek vaak konden overtuigen. In theorie zouden de autovriendelijkere binnensteden met bredere wegen de reistijd en -kosten van en naar het werk verlagen en het dus aantrekkelijker maken om te verhuizen naar een buitenwijk of voorstad. We hebben inderdaad gezien dat de Flächensanierung heeft geleid tot een flinke decentralisatie van de bevolking dankzij de toename van 'verkeersaders'; in de binnenstad, maar niet dankzij veranderingen in de dichtheid van 'gewone' straten. Onze schattingen liggen ook in de buurt van de geschatte invloed van de aanleg van snelwegen vlak buiten steden in andere landen. Hoewel de nadruk in de jaren 1980 verschoof van autovriendelijke steden naar het behoud van met name historische objecten, bieden

deze bevindingen een bruikbaar alternatief hulpmiddel aan beleidsmakers, met name in landen die te maken hebben met de andere effecten van urbanisatie en motorisering, of steden waar de infrastructuur sowieso aan een update toe is.