Essays on International Trade, Geography and Borders

Marta Santamaría

TESI DOCTORAL UPF / year 2019

THESIS SUPERVISOR
Jaume Ventura
Department d'Economia i Empresa

Universitat
Pompeu Fabra
Barcelona

To my parents, Marián and Miguel, for sharing with me the wonder of learning.

Acknowledgements

This thesis is the result of a collective effort and I would like to thank many people for it. First of all, I am incredibly grateful to my advisor, Jaume Ventura, for his continuous support and mentoring throughout these years. Jaume always encouraged me to keep an open mind and ask the puzzling and tough questions. For this, he will always be a source of inspiration for me. Second, I want to thank the faculty at CREi and Pompeu Fabra University, for their brilliant feedback and constant encouragement. Giacomo Ponzetto, David Nagy, Joan Monras, Manuel García-Santana and Edouard Schaal have all significantly contributed to the development of this dissertation. Also, I have to thank the International Economics group at the London School of Economics, where I spent one year during my PhD. Veronica Rappoport, Thomas Sampson and Catherine Thomas guided my research during these months through many enlightening discussions. Finally, I thank Marta Araque and Laura de Agustí for their constant help and patience.

On the personal side, I am grateful for the friends that shared this journey with me. Having a group of friends to share a coffee break or dinner helped me through the tough moments and allowed me to enjoy Barcelona during the good ones. Thanks to Federica, Julia, Francisco, Ana, María, Cristina and Ana Moreno. During my time in London, I had the chance to keep in touch with Julen and Claudia, old friends that were always ready for a long chat, a walk in the park or a pint.

Needless to say, I owe everything to my family, that has supported me through every exam, deadline and presentation. I am deeply grateful to my parents, for the sacrifices they have made and for *always* being there for me, day or night. Second, I am grateful to my siblings, Clara and Ignacio, for being so much fun and for putting up with their elder sister. Third, I thank my grandparents that have always been proud of me, with a special mention to María and Conchita. Last but not least, I thank Jagdish for keeping me calm and making me smile every time, and for always believing in me. Thank you!

Abstract

This thesis explores how different types of trade costs shape trade patterns and the distribution of economic activity across space. In the first chapter, I study how the choice of transport infrastructure after the division of Germany, in 1949, shaped aggregate outcomes and welfare in the following decades. To this end, I develop and calibrate a quantitative trade model with endogenous infrastructure choice. I find that the West German government reshaped the highway network after division increasing real GDP by 0.7% to 2%. In addition, my results show that pre-division highways were a constraint for the government. Removing this constraint would have increased real GDP by 1.85%. In the second chapter, I explore a different type of trade costs: national borders. I use a new dataset of trade across 269 European regions to estimate the "border effect" in Europe. I find that borders in Europe represent a cost equivalent to a 32.5% tariff. This result is robust to different specifications, controls and to a new estimation strategy that identifies the border effect only from border regions.

Resumen

Esta tesis explora los diferentes tipos de costes que moldean la distribución espacial de la actividad económica. En el primer capítulo, estudio cómo la construcción de carreteras después de la división de Alemania, en 1949, afectó a variables económicas agregadas y al bienestar. Con este fin, desarrollo y calibro un modelo de comercio cuantitativo con elección endógena de infraestructura. Encuentro que el gobierno de Alemania Occidental reformó la red de autopistas después del shock generando un aumento del PIB real de entre un 0,7 % y un 2 %. Sin embargo, mis resultados muestran que las autopistas construidas antes de la división crearon una restricción para el gobierno. La eliminación de esta restricción habría aumentado el PIB real un 1,85 % adicional. En el segundo capítulo, exploro un tipo diferente de costes comerciales: las fronteras nacionales. Utilizo datos de comercio entre 269 regiones europeas para estimar el "efecto de frontera" en Europa. Encuentro que las fronteras en Europa representan un coste equivalente a un arancel del 32,5 %. Este resultado es robusto a diferentes metodologías de estimación y a una nueva estrategia empírica que identifica el efecto usando regiones fronterizas.

Preface

This thesis explores how different types of trade costs shape trade patterns and the distribution of economic activity across space. The first chapter quantifies the gains from infrastructure investments and shows that reshaping the highway network after a large economic shock, the division of Germany, had positive welfare and income effects. To address the endogeneity between infrastructure and economic outcomes, I develop a multi-region quantitative trade model where infrastructure is chosen by the government to maximise welfare. I calibrate the model to the prewar German economy and estimate the key structural parameter of the model using the prewar Highway Plan. I exploit the division of Germany, a large-scale exogenous shock to economic fundamentals, to show that the model can predict changes in highway construction after the division. Using newly collected data, I document that half of the new highway investments deviated from the prewar Highway Plan. I find that the reallocation of these investments increased real income by 0.6% to 2% each year, compared to the construction of the original prewar Plan. Finally, I find a large cost of path-dependence: the ability to reshape the full network after the division could have increased real income by an additional 1.5%.

The second chapter explores whether national borders are still an impediment to international trade. We exploit a rich micro-dataset of shipments of goods by road across 269 European regions to construct a matrix of inter-regional trade flows in nine manufacturing industries. Our data comes from the European Road Freight Transport (ERFT) survey and covers 24 European countries and 9 industries from 2011 to 2017. We use this data to estimate the effect of national borders on spatial trade flows following the standard gravity model. Our results show that the existence of a national border between two regions is an impediment to trade similar in magnitude to a 32.5% bilateral tariff, on average. The magnitude of the "border effect" is robust to using a new identification strategy based on comparing only border regions that are contiguous. Our findings suggest that borders are still an important determinant of trade flows in Europe.

Contents

List of Figures

List of Tables

Chapter 1

The Gains from Reshaping Infrastructure: Evidence from the division of Germany

1.1. Introduction

In 1939, a 500 kilometre-long highway connecting Cologne, in the west of Germany, with Berlin, in the east, was about to be completed. It was part of a highway plan designed by the Nazi Government to endow Germany with a modern highway network (Voigtländer and Voth, 2015).[1] No one at that time would have predicted that a border would divide West Germany from East Germany only ten years later.

Figure 1.1 shows the outline of the 1934 highway plan of four thousand kilometres (Panel A) and the layout of the 1974 highway network, of five thousand kilometres (Panel B). By the time of the Division, in 1949, two thousand kilometres had already been built. Did the West German government reshape the highway network after the Division or complete the prewar Highway Plan? I use newly digitised data of the West German highway network to document that there was considerable reshaping after the Division: half of the highway investments made between 1950 and 1974 deviated from the original prewar Highway Plan.[2] Specifically, the government did not build highways that were

---

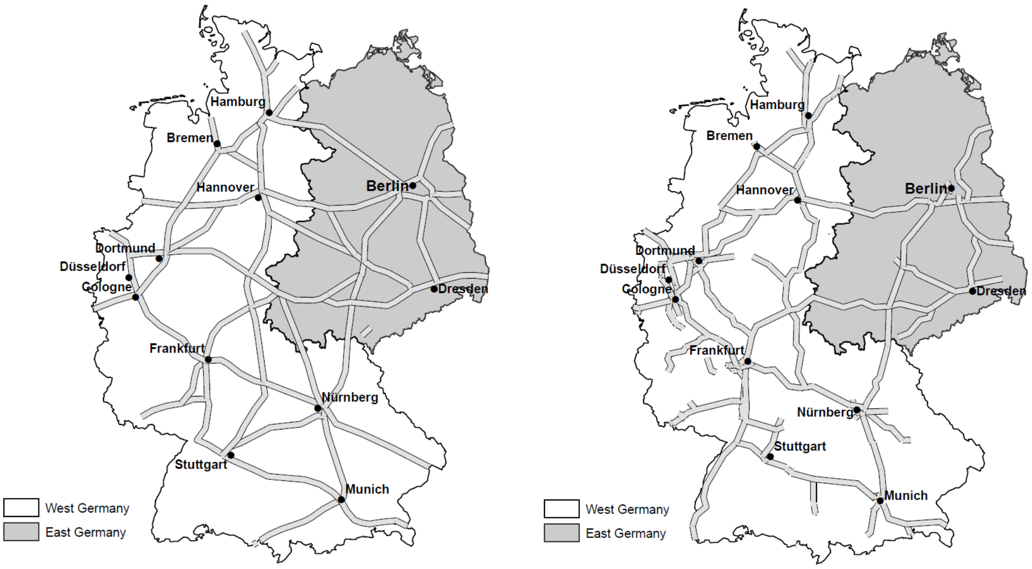
[1] In The 1920s German politicians discussed the construction of a modern highway system. When Hitler appointed Fritz Todt to design the highway system, he traced a plan heavily inspired by the previous plans designed in the 1920s. (Zeller and Dunlap, 2010)

[2] I digitise the German highway network in several years before, during and after the Division using historical maps.

planned to pass across the new border. On the contrary, new highways appeared in the West and close to the border with France.

In this paper I quantify the gains from reshaping the highway network after a large-scale exogenous shock to economic fundamentals: the Division of Germany in 1949. Despite the importance of infrastructure for the movement of goods and people, quantifying the gains from these investments is challenging because infrastructure and economic outcomes affect one another (Redding and Turner, 2015). For example, a government may allocate infrastructure to already fast-growing regions, creating a positive relationship between infrastructure and economic outcomes.[3] Quantitative models or empirical estimations not taking into account the endogeneity of infrastructure will accomplish biased calibrations and estimates.

Figure 1.1: The Division of Germany in 1949 and the Highway network



A) 1934 Highway plan **B) 1974 Highway network**

Notes: The figure displays the countries of West Germany (in white) and East Germany (in grey). Panel A shows the 1934 highway plan. Panel B shows the highways built by 1974. Source: Created by the author from newly digitised historical data.

To address the endogeneity between infrastructure and economic outcomes, I develop

---

[3] Additional challenges, in addition to the endogeneity of infrastructure and economic outcomes, (Fogel, 1962) are the need for a general equilibrium set-up to account for spill-overs (Redding and Turner, 2015) or the complication of choosing an adequate counterfactual.

a quantitative spatial trade model with endogenous infrastructure investments. In the model, trade happens across many regions that are linked through the transport network. Each region has an exogenous productivity level and produces a set of tradable varieties as in Krugman (1980). Infrastructure improvements along a group of regions reduce the shipping costs between these regions, facilitating consumption of non-local varieties. Transport costs are determined by the least-cost path problem of shipping across regions. The equilibrium is driven by workers that choose where to live and how much to consume to maximise utility similar to Redding (2016). This framework builds on the quantitative spatial models reviewed by Redding and Rossi-Hansberg (2017). The novel feature of my framework is a government that, given the decentralised choices of workers and firms, decides how to invest in infrastructure to maximise aggregate welfare.

The model nests a spatial equilibrium into the government's maximisation problem and can be solved backwards. First, given the initial transport network, the matrix of bilateral transport costs is determined by applying the least-cost path algorithm to all region-pairs. With the bilateral transport costs and the initial parameter values, the spatial equilibrium is given by the vector of wages, labour allocations, prices and rents for which the equilibrium conditions hold. The spatial equilibrium determines the expected utility of the economy that is used as a proxy for aggregate welfare (Redding and Rossi-Hansberg, 2017). Then, the government chooses the infrastructure investment allocation that maximises aggregate welfare, given the decentralised spatial equilibrium conditions.

In the solution of the model, optimal infrastructure in a region depends positively on two factors. First, how *large* is a region in terms of trade flows. Second, how *central* is a region in terms of trade transit.[4] The initial highway plan of 1934, in Panel A of Figure 1.1, can serve as an example. Cologne and Berlin are remote cities, located at the edge of the German territory. However, they are included in the highway network because they are large, and thus, they trade intensely with the rest of German districts. Contrary to Berlin or Cologne, Nuremberg, in the south-east, is not a large city but is central: it is located between Munich and Berlin. Because of its centrality, it is also included in the network. These predictions suggest that when the volume of trade or transit in a region changes permanently, as happened after the Division of Germany, the government would like to reshape infrastructure investments.

---

[4] The key driver of these results is that the marginal benefit of investing in infrastructure is higher in large and central regions.

I follow a two-step strategy to quantify the model. First, I calibrate the parameters relevant for the spatial equilibrium, keeping the infrastructure network constant. Specifically, I take the model to data on Germany's population and road network in 1938, eleven years before the German Division. I employ the 1938 population distribution to calibrate the district-specific productivity parameters and the road network to calculate the initial shipping costs. I then assess the fit of the model by comparing the model's trade predictions with domestic good shipments in 1938. Second, given the parameters calibrated in the first step, I estimate the key structural parameter of the model: the returns to highway investments. This parameter determines whether highway investments have decreasing returns and, therefore, shapes the concentration of highway investments in each district. Using the Simulated Method of Moments I estimate this parameter to minimise the difference in the concentration of highway investments between the model with endogenous infrastructure and the 1934 Highway Plan.

Given the quantification of the model I assess its ability to predict infrastructure investments. First, I test the fit of the model with the cross-section of highway investments allocated in the 1934 prewar Plan. To do so, I solve for the *optimal infrastructure* network for Germany before the Division. Comparing the optimal infrastructure allocation with the 1934 Highway Plan, Panel A in Figure 1.1, confirms that the model accounts for the main patterns in the Plan. Besides, the districts where the model predicts the highest value of highway investments coincide with the districts where the first highways were built before Division.

Second, I exploit the Division of Germany to test the ability of the model to predict new highway investments. The Division of Germany provides a unique empirical set-up because it was a large-scale unexpected shock to economic fundamentals: the new border re-defined the country's boundaries and stopped all trade and worker flows between East Germany and West Germany (Redding and Sturm, 2008). To predict the endogenous response of infrastructure to the Division shock, I assume that trade costs between East Germany and West Germany become prohibitive and re-compute the optimal infrastructure allocation for West Germany. This solution takes as fixed the investments made until 1950.

Using data on the construction of highways between 1950 and 1974, I test whether the model can predict new highway investments.[5] Focusing on the additional highway

---

[5] I use investments until 1974 because by then the network was as large as the prewar Highway Plan.

construction between 1950 and 1974 allows me to control for time-invariant specific characteristics that affect both the prewar fundamentals used in the calibration and the amount of infrastructure investments. Examples of such factors are geographical advantages such as being a port city. My estimation shows that the model has good predictive power in explaining the district-level increase in highway construction between 1974 and 1950: the predicted changes account for 19% of the variation in new highway construction. My estimates suggest that a predicted increase of one kilometre in the model explains an increase of 0.32 kilometres in the data (statistically significant at 1 % confidence level). According to these results the reshaping of highways by the West German government can be explained, to a large extent, by a quantitative model of endogenous infrastructure.

The main threat to identify the explanatory power of the model would be changes in other factors, in addition to geography, happening after the Division of Germany and affecting the returns to highway construction unevenly across the West German geography. One of such factors is the process of European integration during which tariffs to international trade were eliminated between Belgium, Italy, Luxembourg, Netherlands and West Germany.[6] To account for this, I extend the model to allow for international trade with other West European countries and solve for the optimal highway construction with both domestic and international trade. The extended model still has good predictive but it predicts a smaller share of the variation in the data than the model without international trade.[7]

Finally, I document to what extent the additional three thousand kilometres built between 1950 and 1974 deviated from the prewar Highway Plan. I find that half of the new highways deviated from the prewar Highway Plan. This considerable reshaping of the highway network suggests that the Government reacted to the change in geography caused by the Division, as predicted by the model.

In 1974, 5000 kilometres had been built while the Plan had a length of 4300 kilometres. Due to limitations in the availability of historical highway maps I use 1974 as the best possible approximation to the length of the Plan.

[6] This process started with creation of the European Economic Community in 1957 with the Treaty of Rome.

[7] One explanation for why the model with international trade seems to be further away from the data could be that in the first years of the european integration process international trade was not so large in magnitude. For example, in 1982, around 85% of all tonnes-kilometres shipped in West Germany by road were domestic shipments while only 15% were cross-border (international) shipments.

I use the model to quantify the gains from reshaping the highway network. First, I find that the reshaping of one third of the network by the government increased real income by 0.64% (compared to the construction of the prewar Highway Plan) in the model with domestic trade and by 2% in the model with international trade. These gains, that increased the level of real income permanently, are obtained only from reshaping the infrastructure, without changing the budget of the government. My results indicate that upgrading infrastructure following the new economic fundamentals, rather than the prewar fundamentals, had very large income effects. Finally, this set-up allows me to quantify the cost of path-dependence. Because highway construction started in the 1930s, the initial investments did not anticipate the Division shock. I find that the ability to reshape the initial highway investments would have increased real income by 1.5% compared to the optimal network constrained by the prewar highway construction.

Relation with the Literature Transport infrastructure projects represent a substantial fraction in the budget of governments and international institutions.[8] The increasing availability of spatially disaggregated data has rekindled the interest of policy-makers and academics on how to better allocate resources to infrastructure upgrading. It is not surprising that a very recent strand of the literature has focused on studying the endogenous choice of transport infrastructure. There have been two approaches so far: (1) to model endogenous infrastructure arising from decentralised decisions like Allen and Arkolakis (2017) and (2) to model endogenous infrastructure as arising from the decision of a government or planner such as Felbermayr and Tarasov (2015), Fajgelbaum and Schaal (2017) and Gallen and Winston (2018).[9] While Gallen and Winston (2018) investigate the choice of infrastructure in a general equilibrium model where infrastructure is a capital investment good that benefits all firms, in my framework the government chooses infrastructure investments in a spatial set-up and is, therefore, closest to Felbermayr and Tarasov (2015) and Fajgelbaum and Schaal (2017).

Felbermayr and Tarasov (2015) endogenise the investment decision in a stylised framework that features two countries located along a line. My model, on the contrary, embeds the government decision in a many-region spatial framework amenable to quan-

---

[8] Between 1995 and 2005, upgrades to the transportation network constituted around 12% of total World Bank lending (Asturias et al., 2014)

[9] Allen and Arkolakis (2017) allow for the emergence of endogenous trade costs due to decentralised shipping choices of traders along the network

titative exercises, like Fajgelbaum and Schaal (2017). There are two main differences between the framework developed by Fajgelbaum and Schaal (2017) and mine. First, I model a government that chooses how to invest in infrastructure subject to the decentralised decisions of workers and firms while Fajgelbaum and Schaal (2017) model a social planner that solves for the allocation of consumption and production together with the optimal infrastructure investment. The structure of my framework, that nests a state-of-the-art spatial model in the goverment's maximisation problem, allows me to use the standard solution and calibration techniques in the trade and urban economics literature. In addition, the spatial equilibrium part of my model can be easily adapted to feature Ricardian trade as in Eaton and Kortum (2002) or to include commuting flows as in Ahlfeldt et al. (2018).[10] On the contrary, Fajgelbaum and Schaal (2017) use a general neoclassical economy model that can accommodate the Armington, Ricardian and factor-proportions models but with a discrete number of goods/sectors and exploit solution techniques developed in the transport literature to solve the optimal infrastructure problem.[11] Second, Fajgelbaum and Schaal (2017) solve for infrastructure investments and trade flows at the link level, leaving the origin and destination of good flows undetermined, while I solve for infrastructure investments at the regional level and I can track both trade flows across the network and the origin and destination of the flows. My paper contributes to this literature in two ways. First, with a new quantitative model that extends the state-of-the-art spatial framework (for example Redding (2016)) by explicitly modelling infrastructure choice. Second, this is the first paper to test the ability of a quantitative spatial model to explain changes in the infrastructure network exploiting an exogenous shock to economic fundamentals. The use of a shock such as the Division of Germany allows me to test the model's predictions exploiting time-variation and, thus, controlling for time-invariant location-specific factors while the previous models in the literature are tested using cross-sectional data.

The results of this paper contribute to the extensive literature about the economic effects of infrastructure investments. This literature can broadly be divided in two categories: First, papers studying the effect of infrastructure access on local outcomes (for example Donaldson (2018) on prices, Michaels (2008) and Duranton et al. (2014)

---

[10] Allen et al. (2014) show the close relation between the structural parameters in many trade and economic geography models that feature a gravity structure.

[11] Specifically, Fajgelbaum and Schaal (2017) re-write the problem as an optimal flow problem in the transport literature, where infrastructure can be solved for to reduce the price differentials across regions.

on specialisation, Banerjee et al. (2012) and Faber (2014) on output).[12] These papers rely on exogenous variation in the construction of infrastructure for identification of local effects. Second, papers studying the aggregate effects of infrastructure investments (for example Donaldson and Hornbeck (2016), Allen and Arkolakis (2017) and Nagy (2016) for the US, Alder and Kondo (2018) for China, Asturias et al. (2014), Donaldson (2018) and Alder (2014) for India, Tsivanidis (2018) for Colombia and Morten and Oliveira (2018) for Brazil).[13] These studies develop rich general equilibrium models to quantify the aggregate effects of transport infrastructure projects. This paper belongs to this second category and is the first paper to quantify the gains from reshaping a fraction of the infrastructure network after a large shock. In addition, my calibration takes into account the endogeneity of infrastructure investments while the rest of frameworks does not consider it.

Finally, this paper contributes to the literature on the role of geography and history in shaping economic activity, for example Davis and Weinstein (2002) on the effects of the Second World War bombings in Japan for city size, Redding and Sturm (2008) on the effects of the Division of Germany for city growth and, most recently, Ahlfeldt et al. (2018) on the effects of the Division of Berlin for agglomeration externalities.[14] By estimating the endogenous response of the infrastructure network to the Division of Germany, I show that infrastructure reshaping is an important mechanism that can exacerbate or atenuate the effects of shocks into the future. Finally, I provide the first estimate of the aggregate cost of path-dependence from inherited highway investments.

This paper is organised as follows: Section 1.2 describes the historical background that serves as a set-up to the paper and the historical data sources. Section 1.3 develops a new theoretical framework with endogenous infrastructure choice and section 1.4 explains the calibration of the model to the Pre-Division economy. Section 1.5 tests the ability of the model to explain the 1934 highway Plan and the new highway construction

---

[12] Other examples are Baum-Snow et al. (2015) on output, Ghani et al. (2016) and Atack et al. (2008) on firm size, Moller and Zierer (2018) on employment, Duranton and Turner (2012), Baum-Snow (2007), Garcia-Lopez et al. (2015), Baum-Snow et al. (2017) on population growth.

[13] Other relevant studies are Balboni (2017) on the interaction between climate change and infrastructure for Vietnam, Fretz et al. (2017) on the effects of highway on spatial sorting for Switzerland and Heblich et al. (2018) on the impact of the railway system for commuting patterns and urbanisation in London.

[14] Other related studies include Brulhart et al. (2012) on the effects of the Fall of the Iron Curtain for the adjustment of wages and employment in Austria and Redding and Sturm (2016) on the effects of the London Blitz for local economic outcomes at the neighbourhood level.

after the Division shock. Finally, section 1.6 reports the quantification of the economic effects of infrastructure and section 1.7 concludes.

1.2. Historical Background

1.2.1. The division of Germany in 1949

In the aftermath of the Second World War the territory of Germany became divided into four parts: two central ones (enclosing nowadays Germany) would be occupied by foreign powers and the other two, the most eastern territories, were annexed to Poland and Russia. Figure 1.2 shows the territory that constituted the new German state under the occupation of the United States, Great Britain, France and Russia, with the most significant cities at the time.

Four zones of occupation were agreed upon by 1945, with each zone under the control of one foreign power to supervise the German de-militarisation. The eastern part remained under Russian control while the western part remained under the control of the Western allies. The delimitation of the East and Western zones followed some pre-existing pattern mostly characterised by features of natural geography (Wolf, 2009). Following the deterioration of the political relations between the Western allies and Russia, with the onset of the Cold War, the two zones of occupation crystallised into two independent countries, West Germany and East Germany, in 1949. Figure 1.2 plots the different territories that were constituted after the Second World War: West Germany, East Germany and two east-most territories that were integrated into Poland and the Soviet Union.

West Germany was the largest territory with 53% of the former German territory and 58% of the population (40 million in 1939).[15] East Germany contained around 23% of the area and 22 % of the population. The former German capital, Berlin, was located within East Germany and was also divided into West and East Berlin. It was the largest city in Germany, with 4 million inhabitants in 1939.

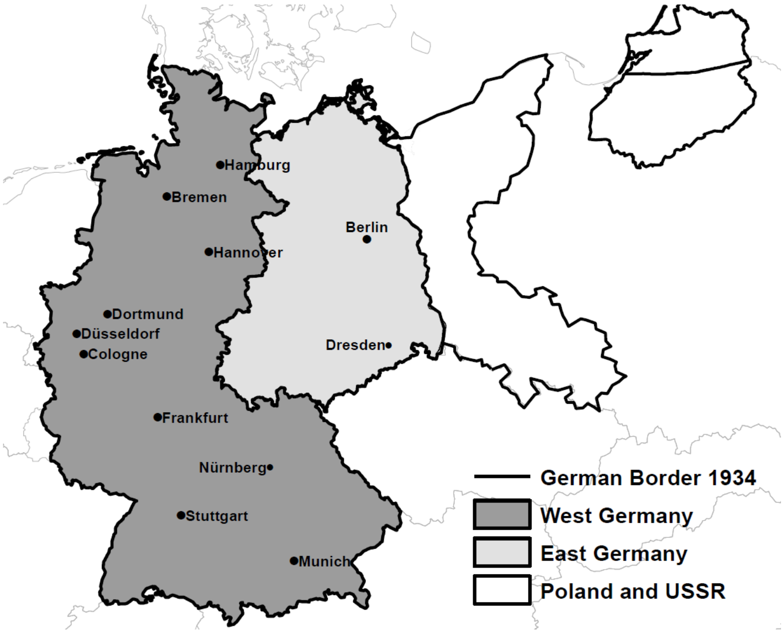
In the initial years after the division, in 1949, there were some economic and political ties between the two states. Yet, the border became sealed from the Eastern side in 1952

---

[15] All numerical figures in this section are taken from Redding and Sturm (2008), and come from the 1952 edition of the Bundesrepublik statistical yearbook.

to prevent migrations to West Germany and all trade relations halted soon after. With the construction of the Berlin Wall in 1961, all population mobility between East and West Germany stopped as well. The division of Germany was recognised by the international community and was generally believed to be permanent.[16]

Figure 1.2: Germany before and after World War II



The division of Germany separated territories integrated for centuries, with origins in the Kingdom of Germany around the year 1000. The foundation of the German Empire in 1871 was the culmination of decades of different levels of economic and political integration. Internal integration of Germany improved substantially after World War I and the German territories were an economically well-integrated area by 1933 (Wolf, 2009). The division, therefore, constituted an important shock that stopped all movement of people and goods between the two states and changed the geographic configuration of West Germany.

Regarding the transportation network, the former German Empire was well connected by a railway system completed in the 1910s. This was the main mode of transportation in the XIXth century. After World War I the construction of a highway network was

---

[16] The two German states became UN members in 1972, the perceptions of the West German population was that reunification was very unlikely even in 1980 (Gerhard Herdegen, 1992)

discussed but finally rejected by the German parliament.[17] The ascent to power of Hitler marked the beginning of the construction of a German-wide highway network that became one of the star policies of the Nazi party. This massive infrastructure project was intended as a way to decrease unemployment and to gain attention from the International press. Fritz Todt, appointed by Hitler as the Inspector General of German Road Construction, traced a plan for the Highway network in 1934 heavily inspired by the previous plans designed in the 1920s.[18]

Transit grew fast along the new highways. In 1955 short-distance shipments by truck were already three times larger than shipments by railway while long-distance truck shipments were one-third of railway shipments.[19] By 1970 short-distance shipments by truck were five times larger than long-distance shipments by rail and long-distance truck shipments in tons became larger than railway shipments by 1985.

1.2.2. Historical data sources

In order to analyse how the division of Germany affected infrastructure investments I need three different sets of data. First, information related to the evolution of the highway network including the outline of the 1934 Highway Plan. Second, information about economic outcomes that will serve to calibrate the model and test its predictions. Finally, additional data to use as controls in the empirical application related to the geography of Germany. The unit of observation through out the analysis will be the district (*Kreise*).[20] This subsection provides an overall description of the data sources employed, further details can be found in section 1.11 of the Appendix.

The first contribution of this paper will be to document the evolution of the West German highway network and the deviations of this network from the 1934 Highway plan. To do this, I collect and geo-reference data about the 1934 Plan and the Highway network.

---

[17] In the 1920s German politicians discussed the construction of a modern highway system. They formed the HAFRABA association that lobbied for the construction of a restricted access motorway connecting Hamburg-Frankfurt-Basel and other connections between major cities (Zeller and Dunlap, 2010).

[18] Zeller and Dunlap (2010)

[19] The data source is the Statistical Yearbook of the Bundesrepublik, multiple years. Figure 1.A1 in the Appendix shows goods traffic in tons by mode of transport.

[20] There are 412 districts between East Germany and West Germany of which 313 districts are in West Germany. For the empirical results the 313 districts are merged according to Mikrocensus regions to account for metropolitan areas.

I digitise and geo-reference the outline of the 1934 Highway Plan using historical maps and compute highway kilometres planned by district. In addition, I collect and geo-reference highway construction data for East Germany and West Germany for the years 1938, 1950, 1965, 1974, 1980 and 1989 from historical maps and road atlases; and from 1950 and 1965 for federal roads. This allows me to document the length and pattern of the network by decade and by district. Figure 1.A2 shows the evolution of the network between these years. Figure 1.A3 in the Appendix displays the pace of construction of the highway network by decade, in kilometres. Finally, I use the EuroGlobal maps dataset, available online, as a source for geo-referenced data of local roads in order to complete the German road network.

To calibrate the theoretical model and test its validity I also require information on historical economic outcomes. I use population data available by decade since 1938 at the district level (*Kreise*) from the historical census. In addition to population data, I collect and digitise data of traffic of goods by road for 18 aggregated traffic districts in Germany, for the year 1939. The traffic data is collected in tons and reported in an aggregated way, as total shipments and total reception by traffic district from the "Statistisches Jahrbuch fur die Deutsches Reich".

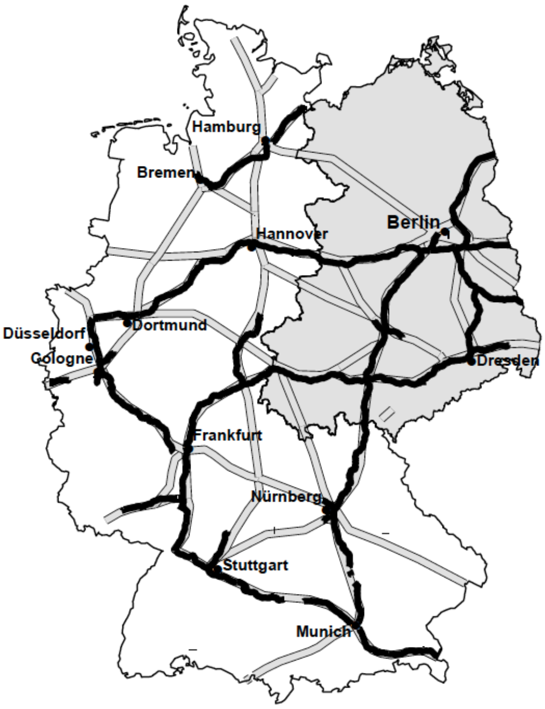
Finally, I collect supplementary geographical data such as area in squared kilometres by district and distance to different geographical boundaries such as the East German border and the border with Western Europe.

1.2.3. Reshaping of the Highway Network after division

In the remaining of this section I will document the construction of the German network of highways up until the division and the completion of the network in West Germany over the following decades. Figure 1.3 shows the 1934 Highway plan over the territories of West Germany and East Germany. As we can see, many of the planned highways were cut by the new border or passed very close to it. The existence of this Pre-division plan represents the initial design of the German government to connect the German territory before the division. The outline of this Highway Plan serves as a counterfactual for the network that would have been built given the economic fundamentals of 1934.

Construction was fast: half of the 6000 kilometres planned were built between 1934 and the beginning of World War II. Figure 1.3 depicts in dark grey the highways that had been built by the year 1946 over the outline of the 1934 Highway Plan. As it is clear in the figure, the construction of highways up until the division followed the pattern of the Plan with almost no deviation.After World War II construction resumed and by 1974, thirty years after the beginning of construction, 5000 kilometres had been completed.

Figure 1.3: Highway plan of 1934 and Highway construction before division



Notes: The figure plots the territories of West Germany and East Germany, delineated in black. The outline of the 1934 Highway plan is plotted in light grey. The highway links that had been built by the year 1946 are plotted in black.

To document whether this additional 3000 kilometres were built following the 1934 Highway outline I classify the old and newly constructed highways, into investments that were planned and investments that were reshaped, and allocated to a different district. I find that the highways built until 1950 followed the 1934 Highway Plan (95% of kilometres were planned) while I find considerable reshaping after the division. Only 47.2% of the kilometres built between 1950 and 1970 followed the 1934 Highway Plan while 52.8 % of the kilometres deviate from the 1934 planned allocation.

Table 1.1: Highway investment allocation (in %)

	Included in the 1934 Highway Plan		
Network	Yes	No	Total
Highway km 1950 (2128 km)	95	5	100
Highway km 1950 to 1974 (3015 km)	47.2	52.8	100

Notes: Share of highway investments (in kilometres) allocated according to the 1934 Highway Plan. Column 1 represents the share of kilometres that were included in the 1934 Highway plan while column 2 represents the share of kilometres that were not included in the 1934 Plan, and were reshaped. The first row refers to the share of kilometres built until 1950 while the second row refers to the new kilometres added between 1950 and 1970.

This decomposition shows that the highway network in 1974 presents several deviations from the original plan. In the next section I build a multi-region spatial trade model with endogenous infrastructure investments to analyse the sources of these deviations, and to quantify to what extent they can be explained by the change in economic fundamentals that followed the division of Germany.

1.3. A theoretical model of endogenous infrastructure choice

In this section, I outline a spatial trade model with endogenous transport infrastructure.[21] I first characterise the spatial equilibrium of the model given an initial infrastructure network. Then I introduce a government that chooses how to invest in the infrastructure network to maximize aggregate welfare. The solution of the model characterises the optimal infrastructure investment, defined as the upgrade in the infrastructure network that maximises welfare. Finally, I use the model to derive qualitative predictions about the response of infrastructure to a shock such as the division of Germany in 1949. The framework features many locations that produce an endogenous measure of differentiated varieties like in Krugman (1980). These varieties can be traded across space subject to transport costs. Workers move across locations to maximize their expected utility that depends on real income and heterogeneous preferences for locations. The model builds on the family of quantitative spatial models reviewed by Redding and Rossi-Hansberg (2017) and is specially close to Redding (2016). I make

---

[21] A detailed exposition of the theoretical framework is contained in section 1.9 of the Appendix.

two contributions with respect to this framework. First, I introduce a new transport cost function that includes infrastructure quality. Second, infrastructure quality is chosen by the government to maximise aggregate welfare.

1.3.1. Model Set-up

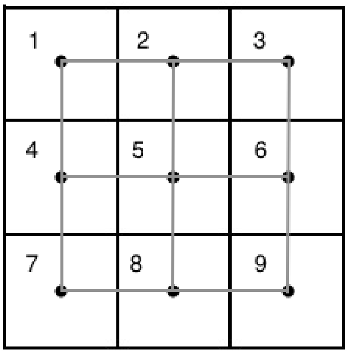
Workers The model features costly trade across many districts, $i = 1...N$, endowed with exogenous labour productivity, A_i. There is a measure L of workers in the economy. Workers derive utility from the consumption of differentiated varieties of the tradable good, from the consumption of housing and from the taste they have for the district they choose to live in. Workers spend a fraction α of their income on the available differentiated varieties and have CES preferences across varieties, with elasticity of substitution $\sigma > 1$. The remaining income share (1-α) is spent on housing. Finally, workers have heterogenous preferences across different districts. These preferences are modeled as an idiosyncratic taste component b_i. Worker ω draws a vector of N realisations $\{b(\omega)_i\}_{i=1...N}$ from a Fréchet distribution with shape parameter ϵ, that governs the dispersion of preferences across workers for different districts.[22]

Firms Production of the differentiated varieties takes place under monopolistic competition, following Krugman (1980). Firms pay a fixed cost of production as well as variable costs in terms of labour, so each firm produces a single differentiated variety in equilibrium. Firms maximise profits by charging a constant mark-up over the marginal cost of production equal to $\frac{1}{\sigma-1}$. Production uses labour as the only input and labour productivity is determined by the district-specific productivity level A_i. The free entry condition drives profits down to zero and pins down the scale of production of each firm. The labour market clearing condition can be solved for the total number of varieties produced in a district and will be a function of the size of the district in terms of population.

---

[22] The parameter ϵ governs the dispersion of heterogenous preferences across workers. A large ϵ implies a low dispersion of the distribution (low standard deviation). Thus, the idiosyncratic preferences are more similar across districts for all workers. Workers have resembling tastes so they react more strongly to changes in real incomes. On the contrary, when ϵ is small the dispersion in preferences is large, and workers are very heterogenous in their taste.

Figure 1.4: Example of simple geography



Notes: Geography with 9 regions; dots are the population centres. In grey the initial transport network, with the same initial quality.

Housing Residential land is assumed to be in fixed supply, as a function of land endowments. I denote the endowment of residential land in district i by H_i, that can be used for housing. Each agent spends $(1-\alpha)$ share of her income on renting residential land. Expenditure on land in each location is redistributed lump-sum to the workers residing in that location as in Redding (2016). This implies that total income in district i, denoted by $v_i L_i$, will equal total labour income plus expenditure on residential land: $v_i L_i = w_i L_i + (1-\alpha) v_i L_i = w_i L_i / \alpha$. This assumption minimises the effects of introducing a housing market in the model while still allowing for a dispersion force that motivates workers to spread across locations because they "dislike" paying high rents. [23] The land market clearing condition will pin down the equilibrium land rent, r_i, in each location.

Geography The geography of the framework is as follows. Districts have some geographic surface of similar size. Workers are concentrated in the centre of the district where consumption and production happen. The set of districts, $i = 1...N$, are located on a finite plane of generic shape. Each district is connected to the adjacent locations by the infrastructure network[24]. These network links can be transited freely by workers but moving goods is costly. The cost of transit depends on the distance that has to be covered and the quality of the infrastructure along that geographic distance. Figure 1.4

---

[23] The real income in location i will be $w/(P^{\alpha} r^{1-\alpha})$

[24] In the calibration of the model the connexion to adjacent districts will be given by the assumed underlying network that is constructed from the existing local roads and federal roads (Bundesstrasse)

provides an illustrative example. This geography can be represented by a graph of edges (infrastructure links) and vertices (population settlements). The set of settlements in this network is fixed, so there is no city creation or destruction. The set of links is also taken as given.[25] The quality of the links, on the contrary, can be improved by investing in infrastructure.

1.3.2. Transport Costs

Consuming non-locally produced varieties is costly because of the dispersion of production and population across the grid of districts.[26] The price of variety i consumed in district n is given by the production price, p_i, and the transport cost of shipping between district i and district n:[27]

$$p_{i,n} = p_i T_{i,n}. \tag{1.1}$$

I will now define how the matrix of transport costs $\{\mathrm{T}_{i,n}\}_{i,n=1...N}$ is determined.

Cost of transit The cost of shipping a good along district i will depend on geography, D, which determines the distance cost in ad-valorem terms that has to be covered across the district. In addition, it will depend on the quality of infrastructure, Φ, that affect how costly (fast) can this distance be transited. The ad-valorem cost of shipping across district k is defined as the cost of transit:

$$\text{Cost of transit}_k = \frac{D_k}{\phi_k^{\gamma}}, \tag{1.2}$$

where I use ϕ_k to denote the district-level infrastructure in k and Φ to denote the vector of infrastructure allocations. I assume that the quality of infrastructure is homogeneous

---

[25] The assumption of a fixed network of links that can be upgraded in terms of quality is also present in related papers in the literature. This constitutes an important difference with the literature about banking, social and business networks where the links are endogenous. Allen and Arkolakis (2014), on the contrary, consider the continuum of space as the domain for the transport cost function that is defined at every point of the plane (instantaneous trade costs). The existence of transport network changes the cost of transit over specific points of the plane.

[26] I assume a domestic closed economy. Thus, there are no tariffs or other trade costs in addition to transport costs.

[27] Without loss of generality, I denote the origin of a trade flow with subscript i and the destination of a trade flow with subscript n.

within a district.[28] This specification of transit costs means that D_k/ϕ_k^γ units of the good shipped will be paid for shipping 1 unit of any good across district k. As we can see, a higher infrastructure investment will reduce the ad-valorem cost of transiting a district. In the quantitative exercise ϕ_k will be the quality of the road and its empirical counterpart will be highway construction.

I assume that $\phi_k \geq 1$, so that the transport cost will always be bounded *above* by the physical geography. Parameter γ is the returns to infrastructure investments. It measures the elasticity of the ad-valorem transit cost to infrastructure investments. I assume it to be positive, so that the cost of transit is decreasing on infrastructure investments. It determines whether infrastructure has increasing returns ($\gamma>1$) or decreasing returns ($\gamma<1$).

We can think of this geography of regions linked by the transport infrastructure as a network composed of vertices (regions) and edges (transport links). Each edge, as in the network literature, will be associated with a cost of transit defined in ad-valorem terms. Since I have specified the cost of transit at the district level, I need to define a way to aggregate district specific transit costs into bilateral costs. I define the cost of shipping across a link i,n as $\omega_{i,n}$:

$$\omega_{i,n} = \sum_j 0.5 \times \frac{\text{distance cost}_{i,n}}{\phi_j^\gamma} \tag{1.3}$$

This functional form implies that the cost of transiting across two districts with different levels of infrastructure will be averaged across the different infrastructure levels and distance-costs of the two districts and that infrastructure qualities are not perfect substitutes. Once we have identified the cost of shipping across links, we discuss how the shipping decision is made.

Least-cost path problem Given the cost of transit along all links, what is the transport cost between district i and district n? In this network economy there will be many alternative paths to ship goods between districts i and n. To define the transport cost matrix, I make the following assumptions.

First, I normalize the transport cost within a region to 1, so that $T_{n,n} = 1$ for all regions.

---

[28] In the real world a district may have one very high-quality highway and one very low-quality road. Therefore, we may think of ϕ_i as the average quality of the infrastructure stock in district i.

Second, I define the transport cost across any two regions as 1 plus the smallest-possible cost (least-cost path) to transit between these two districts in ad-valorem terms.

Figure 1.5 will help us illustrate the transport cost function with a simple example. Consider the three German districts in the figure (A, B and C) located sequentially. The transport cost between A and B will be $T_{A,B} = 1 + \omega_{A,B}$ because these two regions are contiguous. The transport cost between A and C will be:

$$T_{A,C} = 1 + \omega_{A,B} + \omega_{B,C} = \sum_{j=A,B} 0.5 \times \frac{\text{distance cost}_{A,B}}{\phi_j^{\gamma}} + \sum_{j=B,C} 0.5 \times \frac{\text{distance cost}_{B,C}}{\phi_j^{\gamma}} \quad (1.4)$$

Notice that $\omega_{A,B}$ and $\omega_{B,C}$ are functions of the infrastructure quality in B. Therefore, the infrastructure quality in B affects the transport cost between A and C. In standard trade models, this is normally not the case because transport costs are purely bilateral. This implies that the cost of shipping between any origin and any destination only depends on origin and destination-specific parameters. However, to study road transportation we need a more general specification of the transport cost function that takes into account the spatial nature of transport costs.

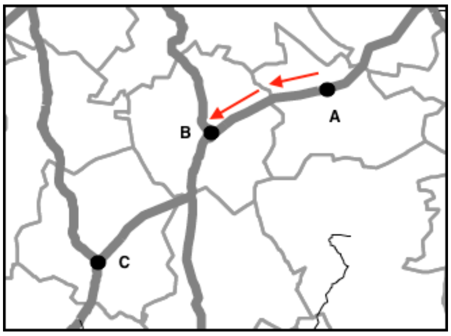


Figure 1.5: Transport costs: Illustration

It is important to point out that this functional form assumes that the triangle inequality is always strict. Shipping from A to B ($T_{A,B}$) and then from B to C ($T_{B,C}$) will always be more expensive than direct shipping ($T_{A,B} \times T_{B,C} > T_{A,C}$).

Finally, I assume goods are always shipped along the least cost path. This is similar to modelling a shadow transport sector that operates under perfect competition, and therefore, ships goods at the minimum costs.[29]

---

[29] Formally, define the set of vertices in the network as V. We can define a path P as a sequence of

1.3.3. Location choice and Spatial equilibrium

Worker location choice and Welfare Workers choose where to live by maximising indirect utility, given by real income and the idiosyncratic preference taste. The distribution of indirect utility is also Fréchet and, given the properties of this probability distribution, we can write the share of workers that choose to live in district i as:

$$\frac{L_i}{L} = \frac{(v_i/P_i^{\alpha} r_i^{1-\alpha})^{\epsilon}}{\sum_{n=1}^{N}(v_n/P_n^{\alpha} r_n^{1-\alpha})^{\epsilon}}, \tag{1.5}$$

where $v_i = w_i/\alpha$ denotes total income in location i. Expected utility for a worker across locations is given by:[30]

$$\tilde{U} = \delta\left[\sum_{i=1}^{N}(v_I/P_i^{\alpha} r_i^{1-\alpha})^{\epsilon}\right]^{1/\epsilon}, \tag{1.6}$$

where $\delta = \Gamma\left(\frac{\epsilon}{\epsilon-1}\right)$ and $\Gamma(.)$ is the gamma function. We impose $\epsilon > 1$ to ensure a finite value of the expected utility. Because indirect utility follows a Fréchet distribution the expected utility conditional on living in district i is the same across all districts and equal to the expected utility of the economy as a whole.[31] Following Redding (2016), I use this measure of expected utility as a proxy for aggregate welfare.

Spatial equilibrium For a given initial transport network defined by $\{D, \Phi\}$ and the transport cost function specified in subsection 1.3.2, and exogenous land endowments $\{H_i\}_{\forall i \in N}$ and productivities $\{A_i\}_{\forall i \in N}$, the spatial equilibrium is a combination of wages, price indices, rents and labour allocations, $\{w_i, P_i, r_i, L_i\}$ such that for all districts the goods and housing markets clear in each district, the domestic labour marker clears domestically and expected utility is equalised across all workers. The equilibrium trade shares and rental rates can be solved as a function of these four equilibrium variables. The following equations define the equilibrium vector $\{w_i, P_i, r_i, L_i\}$:

---

vertices (regions) $P = (v_1, v_2, ..., v_n) \in V \times V \times ...V$ such that v_i is adjacent to v_{i+1} for 1≤ i < n. The path P has length n-1 and goes from v_1 to v_n. Let $\omega_{i,j}$ be the cost of shipping across vertices v_i and v_j. The shortest path from v to v' is the path $P = (v_1, v_2, ..., v_n)$ where $v_1 = v$ and $v_n = v'$ that over all possible n minimizes $\sum_{i=1}^{n-1} \omega_{i,i+1}$

[30] See part 1.9 of the appendix for derivation details

[31] Because more productive districts attract more workers despite their preference taste the expected value of indirect utility, E$(b_n w_n/P_n^{\alpha} r_n^{1-\alpha})$ will equalise across locations.

The goods market clearing is given by the balanced trade condition:

$$w_i L_i = \sum_j \frac{L_i}{\sigma F}\left(\frac{\sigma}{\sigma - 1}\frac{w_i}{A_i}T_{i,j}\right)^{1-\sigma}(P_j)^{\sigma-1}w_j L_j, \forall i. \tag{1.7}$$

The Price index in district i given by:

$$P_i^{1-\sigma} = \sum_j \frac{L_j}{\sigma F}\left(\frac{\sigma}{\sigma - 1}\frac{w_j}{A_j}T_{j,i}\right)^{1-\sigma}, \forall i. \tag{1.8}$$

The rental rate is given by the clearing of the housing markets:

$$r_i = \left(\frac{1-\alpha}{\alpha}\right)\frac{w_i}{L_i}, \forall i. \tag{1.9}$$

The fraction of workers that chooses to live in district i is determined by the worker's utility maximisation problem and implies the following workers' residential choice equation:

$$\frac{L_i}{L} = \frac{(v_i / P_i^{\alpha} r_i^{1-\alpha})^{\epsilon}}{\sum_{n=1}^{N}(v_n / P_n^{\alpha} r_n^{1-\alpha})^{\epsilon}}. \tag{1.10}$$

Given the model's parameters and transport infrastructure $\{\Phi\}$, equations (1.7), (1.8), (1.9) and (1.10) can be solved for the equilibrium vector $\{w_i, P_i, r_i, L_i\}$. Lastly, the equilibrium level of expected utility, $\tilde{U}$, is implicitly determined by the domestic labour market clearing, $\sum_i L_i = L$.

Existence and Uniqueness As shown in Redding (2016) the condition for the existence and uniqueness of the spatial equilibrium will hold if the elasticity of expected utility to the labour share in a district is negative, this is, if the dispersion forces are stronger than the agglomeration forces of the model.[32] In the kind of models with housing and imperfect labour mobility, the condition for existence and uniqueness of the equilibrium is

$$\sigma\left(1 - \frac{\alpha}{1 + \frac{1}{\epsilon}}\right) > 1. \tag{1.11}$$

---

[32] The proof follows the same structure as in Allen and Arkolakis (2014)

1.3.4. Problem of the Government: Choice of Infrastructure Investment

I model the choice of infrastructure as a Stackelberg game between the Government and the economic agents in the economy (workers and firms). The Government is the leader and thus has the advantage to choose first in the game. The game is solved by backward induction. Thus, the Government chooses infrastructure to maximise expected utility, $\tilde{U}$, constrained by the choices of workers and firms, given by the decentralised equilibrium allocation. This set-up is similar to a Ramsey problem with a Government that maximises welfare replacing the FOCs from the problems of consumers, firms and workers into the constraints.

I assume that the Government can choose how to allocate a fixed amount of resources to improve infrastructure across all the districts in the economy. This budget, that I denote by Z, is modelled as an endowment of the government and thus, is assumed to be exogenous to the government's decision. The budget is not charged to consumers via taxes or any other public finance tools. The cost of investing in district i is $c_i \phi_i$ and the budget constraint of the government is:

$$\sum_i c_i \Phi_i \leq Z. \tag{1.12}$$

The marginal cost is equal to c_i, that is allowed to differ across districts.

Government's problem We can write the problem of the Government as follows:

$$\underset{\{\Phi_j\}}{\text{Max}}\ \delta \left[\sum_{i=1}^{N} (v_i(\Phi)/P_i(\Phi)^{\alpha} r_i(\Phi)^{1-\alpha})^{\epsilon} \right]^{1/\epsilon},$$

subject to:

1. *Goods market clearing*

$$w_i L_i = \sum_j \frac{L_i}{\sigma F} \left(\frac{\sigma}{\sigma - 1} \frac{w_i}{A_i} T_{i,j} \right)^{1-\sigma} P_j^{\sigma-1} w_j L_j, \forall i. \tag{1.14}$$

2. *Labour market clearing*

$$\frac{L_i}{L} = \frac{(w_i/P_i^{\alpha} r_k^{1-\alpha})^{\epsilon}}{\sum_{k=1}^{N}(w_k/P_k^{\alpha} r_k^{1-\alpha})^{\epsilon}} \text{ and } \sum_i L_i = \bar{L}. \tag{1.15}$$

3. *Transport cost matrix {T}* specified as in subsection 1.3.2, determined by the least-cost path shipping.

4. *Government's budget constraint*

$$\sum_i c_i \Phi_i D_i \leq Z, \tag{1.16}$$

where $P_i = \left[\sum_j \frac{L_j}{\sigma F}\left(\frac{\sigma}{\sigma-1}\frac{1}{A_j}T_{j,i}\right)^{1-\sigma}\right]^{1/(1-\sigma)}$, $r_i = \left(\frac{1-\alpha}{\alpha}\right)\frac{w_i}{L_i}$ and $\delta = \Gamma(\frac{\epsilon}{\epsilon-1})$, where Γ is the gamma function.

Intuition for Welfare effects To build some intuition about the effects of infrastructure investment on welfare let us consider the same problem but without allowing the matrices of least-cost paths to change and in a model with no housing ($\alpha = 1$). This assumption avoids a response of the shipping decision to a change in infrastructure upgrading and abstracts from the response of rents to new investments.[33] Holding the shipping path between every pair of districts constant, the first order condition with respect to ϕ_j is:[34]

$$\frac{\partial \mathcal{L}}{\partial \phi_j} = 0 : \underbrace{\delta^{1/(1-\epsilon)} \sum_i \sum_k (\frac{U}{v_i})^{1-\epsilon} \frac{X_{k,i}}{P_i} \gamma \mathbb{I}_{k,i}^{j} \frac{D_j}{\phi_j^{\gamma+1}} p_k}_{\text{Direct effect}} + \underbrace{(\sigma - 1) \sum_i \sum_k \eta_i X_{ki} \left(\frac{\partial T_{ki}}{\partial \phi_j}\frac{1}{T_{ki}} - \frac{\partial (P_k)}{\partial \phi_j}\frac{1}{P_k}\right)}_{\text{Response of wages}}$$
$$+ \underbrace{\epsilon \sum_i \lambda_i \frac{L_i}{L}\left(\frac{\partial P_i}{\partial \phi_j} - \sum_k \frac{\partial P_k}{\partial \phi_j}\frac{1}{P_k}\right)}_{\text{Response of Labour}} = \underbrace{\mu c}_{\text{Marginal cost}}, \tag{1.17}$$

---

[33] Allowing for changes in the shipping path would just add an additional term to the expression below, accounting for how the shipping path matrix will change after an infrastructure upgrading. This effect is not quantitatively very large.

[34] The derivations use Roy's inequality to get the direct effect and can be found in the part 1.9 of the Appendix.

where v_i denotes now real income in district i, $v_i = \frac{w_i}{P_i^{\alpha} r_k^{1-\alpha}}$ and $\mathbb{I}_{k,i}^{j}$ is an indicator matrix that captures whether region j is in the least-cost path to move goods between i and j. When $\mathbb{I}_{k,i}^{j} = 1$ this indicator reflects the fact that the infrastructure quality in region j will have an direct effect on the trade flows between k and i, through the transport costs.

Equation (1.17) shows that infrastructure investment is chosen so that the marginal benefit of investing in a district, left-hand side of the equation, equates the marginal cost of building infrastructure in that district. Notice that the marginal benefit of investing in district j includes the sum of the gains from investing in district j across all districts in the economy that may benefit from the change, similar to the Samuelson rule for the allocation of public goods.

The marginal benefit is composed of a direct effect and the indirect effect coming from the response of wages and population. As we can see, the direct effect is just the partial equilibrium effect of an upgrade of infrastructure on aggregate welfare, before the adjustment of wages and population. This effect is the largest, quantitatively, and affects real income through the effect of infrastructure investments on the Price index of tradable goods. In addition to the partial equilibrium effects the change in infrastructure quality will trigger a response of wages and population that will adjust in response to the new infrastructure quality. In quantitative terms the response of wages and population is small compared to the response of the Price index. The response of wages is increasing on the elasticity of substitution across goods, σ, and the response of population is increasing on the degree of homogeneity across workers, ϵ (recall that $\epsilon \to \infty$ is the case with perfect worker mobility).[35]

We can build intuition about the optimal infrastructure allocation by approximating infrastructure investment in district j with the *partial equilibrium* effects. Rearranging the terms of equation (1.17) we can write an expression for the infrastructure investment level in district j:

$$\phi_j^{\gamma+1} \approx D_j C \left(\underbrace{\sum_k \left(h_{k,j} + h_{j,k}\right)}_{\text{Size=Importance as } \textit{source} \text{ of trade}} + \underbrace{\sum_{i \neq j} \sum_{k \neq j} \mathbb{I}_{k,i}^{j} h_{k,i}}_{\text{Centrality=Importance as a } \textit{hub} \text{ for trade}} \right), \tag{1.18}$$

---

[35] Notice that changes in the transport costs will trigger a larger response of wages if goods are very good substitutes, thus triggering a large trade response to changes in prices, and if workers are very mobile, thus triggering a large response of workers to changes in real income.

where the function $h_{k,j} = v_k^{1-\epsilon}\frac{X_{j,k}}{P_k}p_j^{-\sigma}$ is increasing in exports from k to j. Equation (1.18) shows that infrastructure investments will be higher in districts that trade more, first term in the parenthesis, and in districts that are transited by large trade flows, second term in the parenthesis.

This expression is a non linear function of the weighted sum of total trade flows that are exported by j, are imported by j or transit j, with weights being a function of real incomes. To see why notice that $\mathbb{I}^j_{k,i}$ will be one every time $i = j$ or $k = j$ and when j is transited by a trade flow between a pair of regions not including j. For all other trade flows the indicator $\mathbb{I}^j_{k,i}$ will be zero and will not show up in this expression. Thus, the infrastructure level in district j will be a function of the total trade flows of region j and of the trade flows that transit j. The labels indicate that we can think of this expression as a function of the importance of district j in terms of size (importance as a *source* of trade) and in terms of centrality (importance *hub* for trade flows).

1.3.5. The division of Germany in 1949

The theoretical framework developed in this section helps us understand what is the optimal infrastructure pattern across regions in a general equilibrium framework. As indicated in equation (1.18) infrastructure investment will be higher in districts that are an important *source* of trade flows, and in districts that are an important *hub* for trade flows. In this framework, a permanent change in the size of trade flows or trade transit in a district would create incentives to reshape the infrastructure network. Given some infrastructure budget, the new investments would be allocated to maximise aggregate welfare given the new fundamentals and the investments already made.

The division of Germany into East Germany and West Germany in 1949 was a sharp shock to the trade partners of West German districts (Redding and Sturm, 2008). Firstly, it caused a reduction in the domestic exports and imports of West German districts as all trade with East Germany stopped. Besides, the transit going through different West German districts was affected because districts near the Inner German border, very central before division, became remote after division. Finally, it even caused a change in the transportation network because the border cut through some of the already existing highways and roads.

In the next section I take the model to historical data of Germany to simulate the effects of the division on the optimal infrastructure network and test to what extent this quantitative framework can explain the observed highway reshaping documented in section 1.2.

1.4. Calibration of the model

In the previous section I have built a quantitative spatial model that incorporates endogenous infrastructure construction. In this section, I take the model to the data. The goal of this section is to achieve a quantification of the model that captures the economic geography of Germany and can be used as a tool to understand the effects of the division shock on the reshaping of the transport network.

I follow a two-step strategy to quantify the model detailed in the previous section. First, I calibrate the model to the German economy before the division. This calibration, that will target the year 1938, abstracts from the endogenous response of infrastructure investments. The goal of this first step is to test whether the predictions of a spatial equilibrium model with a fixed infrastructure network can match the patterns in German data from the year 1938. I use newly collected data of shipments of goods by road within Germany to assess the fit of the model's predictions given the infrastructure network at that time.

The second step is to use the model to compute the optimal highway network before the division of Germany. The optimal network is the solution to the government's problem presented in section 1.3. I compare the model-generated optimal network with the 1934 highway plan to check the ability of the model to capture the main objectives in highway network planning.

1.4.1. Calibration of the Spatial Equilibrium

The goal of this calibration is to obtain a quantitative model that represents as close as possible the spatial equilibrium of the German economy before the division. To take the model to the data I need to calibrate two sets of parameters: first, the pre-division transport network that will determine the initial transport cost matrix, second, the parameters of the model that will determine the spatial equilibrium.

Initial network and Transport costs

The geography of this model is a graph composed of a set of districts linked by the transport network. This graph represents the underlying geography of Germany and is assumed to be fixed. The quality of the links can be upgraded by investing in infrastructure. I build the underlying graph as follows. I combine the highways, (Autobahns) and all federal highways (Bundesstraße) that existed in 1950.[36] I add the local roads needed to ensure that all districts in Germany are connected to the network. This gives me a network that contains all german districts.[37] Figure 1.A4 in the Appendix displays the roads chosen for the initial network and the graph corresponding to this network.[38]

After building the network, I compute the cost of transporting goods following Combes and Lafourcade (2005). I compute the shipping cost by adding a time-related component and a distance-related component. These costs are the frictions that the government will be able to reduce by investing in infrastructure quality. Finally I convert the computed initial transport costs (in euros) to ad-valorem transport costs by scaling the cost of shipping by the average value of the shipment of a truck in Germany in 1950. Full details about the cost computation can be found in part 1.10 of the Appendix.

Given the graph and the associated transit costs I can compute the initial transport cost matrix by applying a least-cost path algorithm to the network. This calibration gives me the transport cost matrix in 1938, before the German division.

Parameter choice

In addition to the transport cost matrix $\{T_{i,n}\}$, calculated as explained above, the model described in section 1.3 has several additional parameters to be calibrated. First, there are two district-specific vectors of parameters: $A_i\}$, the exogenous productivity of each district and $\{H_i\}$, the land endowment of each district. Then there are standard parameters present in other trade and spatial models. This is the case of α, the share of tradable goods in total expenditure, ϵ, the shape parameter of the Fréchet distribution from which idiosyncratic tastes are drawn and $(\sigma - 1)$, the trade elasticity. Finally, the model has

---

[36] The federal highways are roads with multiple lanes but not limited-access like Autobahns

[37] I provide further details of the construction of the network in 1.10 of the Appendix.

[38] For the network construction I use the Network Analysis toolkit in the geographic information software ArcGIS.

three parameters related to the construction of infrastructure: γ, the elasticity of transport costs to infrastructure investments, Z, the budget of the government for infrastructure upgrades and $\{c_i\}$ the district-specific marginal cost of construction.

Standard parameters calibrated to exogenous values I calibrate $\{\epsilon, \alpha\}$ to existing values in the literature. I set the shape parameter of the Fréchet distribution to $\epsilon = 3$ following the estimated value from domestic migration flows across U.S. counties by Monte et al. (2015). I vary this parameter in robustness checks to $\epsilon = 7$, value estimated for the heterogeneity of worker's preferences governing commuting and location choices within Berlin in Ahlfeldt et al. (2018). I calibrate an expenditure share of tradables of α=0.7, leaving an expenditure share of housing of $(1 - \alpha) = 0.3$ following Redding and Sturm (2008) in their study about the population growth effects of the German division.

Standard parameters calibrated to Germany 1938 The district-level productivities, $\{A_i\}$, are calibrated to match the population distribution of Germany in the year 1938: I compute the productivity level of each district by inverting the spatial equilibrium and solving for the vector of district productivities that, in equilibrium, delivers the population distribution observed in the data.[39] I use population at the district-level for 1938 from the German Census collected in the Statistical Yearbooks of the Federal Republic of Germany.

The district-level land endowments, $\{H_i\}$, are equated to the surface of each district in squared kilometres as measured in the data.

Parameter estimated using the full-structure of the model: Trade elasticity Finally, I use data on shipments by road over 10 distance brackets in 1938 to calibrate the trade elasticity parameter $(1 - \sigma)$. For this estimation I use the full structure of the model with fixed infrastructure. First, I estimate the elasticity of trade shipments to distance using historical data of shipments and obtain an estimate of $\beta = -2.8^{***}$. This estimate is larger than the average magnitude estimated in the literature but transiting through ground-transport means, such as roads, has been shown to yield substantially higher distance coefficients(Disdier and Head (2008)).

---

[39] This calibration technique is explained in the survey by Redding and Rossi-Hansberg (2017)

Under the standard assumptions of the gravity equation this elasticity is the product $\beta = (1 - \sigma) \times \nu$, where $\nu = \partial log(T_{i,j})/\partial log(dist_{i,j})$. The consensus in the literature is to choose $\nu = 0.3$ (Monte et al. (2015) among others). The parameter ν does not have an exact counterpart in my model. It will be a combination of the elasticity of transport costs to distance along different types of road and conditional on transit over the least-cost path.

To compute the value of σ implied by the estimated elasticity of $\beta = -2.8^{***}$ we need the elasticity ν in 1938 as implied by my model. To this end I set the elasticity of substitution to σ=5, following the consensus in the trade literature (for example (Broda et al., 2008)), and compute the implied trade flows across all district pairs conditional on the parameter values chosen above. The elasticity of trade shipments to distance in the model with $\sigma = 5$ is $\beta^{model} = -1.84^{***}$, which implies a value of $\nu = \partial log(T_{i,j})/\partial log(dist_{i,j}) = 0.46$. Given $\nu = 0.46$, I set $\sigma = 7$ in order to achieve an elasticity of trade flows with respect to distance that matches the estimated elasticity in the data of 1938 ($\beta^{model} = -2.8$). Table 1.A1 shows that the elasticity of trade shipments to distance in the model with $\sigma = 7$ is $\beta^{model} = -2.78^{***}$, equal to the elasticity estimated in the data.

1.4.2. Calibration of the infrastructure-related parameters

Parameters related to Infrastructure-choice The three parameters specific to my model that are crucial for the choice of infrastructure are γ, the returns to infrastructure investments, Z, the government's budget to invest on infrastructure and $\{c_i\}$, the district specific marginal construction cost. Recall that the budget constraint of the government is:

$$\sum_{i=1...N} c_i \phi_i \leq Z, \tag{1.19}$$

where $\{\phi_i\}$ is the vector of infrastructure investment allocations. Because the initial underlying grid is constructed from the existing highways and roads in Germany these links will already be, to some extent, equally easy to build on. Thus, I choose $c_i = c, \forall i$ to simplify the computation problem. However, the ruggedness of the terrain or the existence of rivers could be introduced easily in the problem.

I discipline the two remaining parameters $\{\gamma, Z\}$ using the design of the 1934 highway

Plan. The parameter γ, the returns on infrastructure investments, determines whether infrastructure investments have increasing or decreasing returns. Therefore, this parameter will shape how concentrated are investments at the district level. I estimate γ to bring the degree of concentration of investments in the model as close as the concentration in the 1934 Plan. As a measure of concentration of investments I use the skewness of the distribution of highway kilometres in the 1934 highway plan. I estimate this parameter using Simulated Method of Moments (SMM) on a simulated 50-district economy where I discipline the productivity distribution using random draws from a distribution similar to the calibrated productivities in the previous section. This estimation yields parameter $\gamma = 0.84$. Further details are provided in section 1.10 of the Appendix.

To calibrate the budget of the government, Z, I use the share of highway kilometres by district allocated in the 1934 Plan. I compute the average share of kilometres in the data by district $\bar{h} = 2.8$ and set $Z = (\bar{h}) \times N$ so that the average expenditure in the model matches the average kilometres by district in the 1934 highway Plan. N is the number of districts, $N = 395$.

Finally, I impose $\phi_i \geq 1$ to ensure that the government is constrained by the original network and transport costs can only decrease with the choice of infrastructure. This also ensures that infrastructure that is given cannot be disinvested. I adjust the budget to account for the cost imposed by this lower bound restriction.

Given the productivity distribution, the initial transport network and transport costs and the calibrated and estimated structural parameters, I compute the infrastructure allocation that maximises aggregate welfare. The solution to the Government's problem is a 395 vector of the optimal district level investments that I can compare to the 1934 government plan.

1.4.3. Solution of the model: Optimal infrastructure

Solution method Given the parameter values, the underlying transport network and the initial transport costs, I solve for the infrastructure allocation that maximises expected utility. The solution will be a vector of 395 infrastructure investments representing the spatial pattern of infrastructure investments.

The model does not feature congestion costs because transport costs are independent of the quantity shipped ($\frac{\partial T_{i,j}}{\partial X_{ij}} = 0$). This choice simplifies the solution method because given the network and the investment vector I can compute transport costs independently of the equilibrium allocation. However, the lack of congestion makes the government's problem not globally convex.[40] The short-come of modelling transport costs as constant on the quantity shipped (weakly convex) is that I cannot prove that the solution I find is the global optimum of the problem. In my set-up this is not a concern. The local optimum will be the best possible deviation from the initial network. This solution coincides with the problem the government has to solve: how to allocate limited resources to upgrade the highway network. On the contrary, the global optimum may be very far from the initial network and require a much more significant investment. Even without congestion the spatial problem of the economy is convex and features a unique and stable equilibrium (given the calibrated parameter values in the previous section). The transport costs are constant on the quantities traded and convex on the infrastructure investments which makes the problem (weakly) convex.

I can rewrite the problem as an optimisation of the expected utility in equilibrium over the infrastructure investment vector:

$$\underset{\{\phi_j\}_{j\in N}}{\mathrm{Max}} EU^{eq} = f(w^{eq}(\Phi), P^{eq}(\Phi), r^{eq}(\Phi), L^{eq}(\Phi), T^{eq}(\Phi), \Phi), \tag{1.20}$$

where EU^{eq} is the equilibrium expected utility for a given infrastructure network and a given vector of infrastructure investments (Φ). The equilibrium expected utility is a function of the equilibrium wages, $w^{eq}(\Phi)$, equilibrium Price indices $P^{eq}(\Phi)$, equilibrium rents, $r^{eq}(\Phi)$, equilibrium population allocation, $L^{eq}(\Phi)$ and the equilibrium transport cost matrix, $T^{eq}(\Phi)$. To solve for the infrastructure allocation, I start the problem from the initial transport network (assuming all roads are local), and I search for the infrastructure allocation that maximises expected utility using an interior-point algorithm.

1.5. Test of the model

The goal of this section is to show that the calibrated model developed in section 1.3 provides a reasonable model of Germany and can be used to quantify the gains from infrastructure investments. The first part of this section is devoted to the fit of the model before the division of Germany. Since the model has been calibrated to 1930s Germany,

---

[40] See Fajgelbaum and Schaal (2017) for a detailed discussion about convexity in spatial networks

the model should be able to capture static outcomes before the division in 1949. In addition, I will use the model to solve for the optimal infrastructure investment before division and show that the model can also rationalise a large part of the 1934 Highway plan. The second part of this section performs the true test of the model. I will exploit the division of Germany, an unexpected and exogenous change in borders, to proof that model can also predict different outcomes in the post-division period. Notice that the good fit of the model for the 1930s may be expected due to the calibration strategy. However, the performance of the model post-division constitutes an out-of-sample type of test. The ability of the model to capture post-division outcomes will provide us with a lot of confidence that we can use the model to perform counterfactual analysis.

1.5.1. Fit of the Model before division

Spatial Equilibrium

Given the initial transport network summarised by $\{T_{i,n}\}$ and the value of the parameters of the model discussed in the previous section, I solve for the spatial equilibrium defined by equations (1.7), (1.8),(1.9) and (1.10) of section 1.3. With the equilibrium endogenous variables $\{w_i, P_i, r_i, L_i\}$ I construct the matrix of simulated trade flows between all German districts.

As a first check I show that the relation between trade flows and distance is very close in the data and in the model. I collect the total tons of goods shipped by distance bracket from manufacturing shipments and furniture shipments in the year 1938 within Germany.[41] I construct the share of shipments in real terms by distance bins in equilibrium using the calibrated model and compare the model's predictions with the data. Figure 1.6 plots the density of trade flows over distance in the data (continuous line) and in the model (dashed line). The very good fit of the model to the data is not surprising because I calibrate $\sigma = 7$ to match the elasticity of trade flows to distance. However it is reassuring that such a good fit can be achieved by fitting only one parameter.

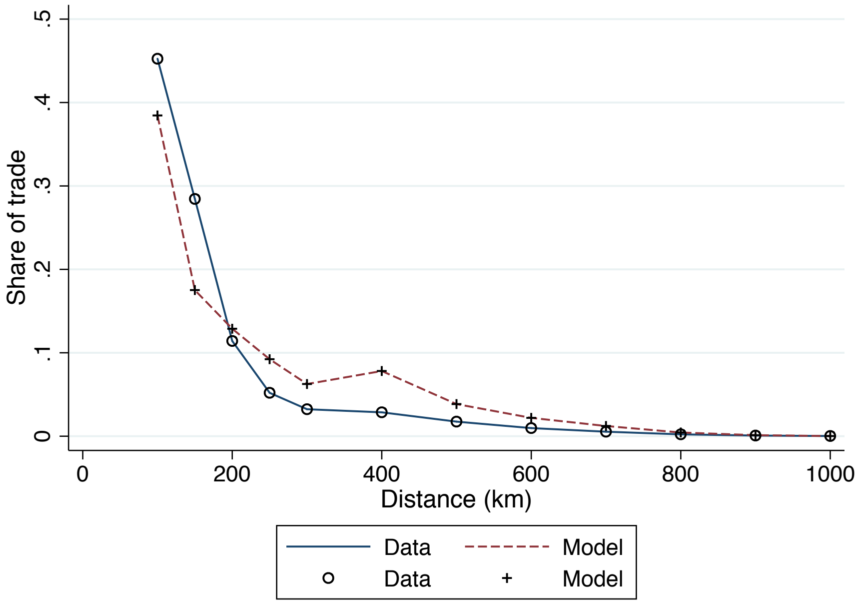
To provide a further check, I aggregate total trade by district into a different classification, traffic-districts (*Verkehrsbezirken*), for which I observe road shipments in the historical data. The data provides a measure of the total tons of goods received by

---

[41] Data collected from Statistical yearbook of the Deutches Reich (1940)

road in any traffic-district from the rest of Germany and of total tons shipped to the rest of Germany. This information is available for 18 traffic-districts (that contain all 412 districts in Germany). I compare the predicted trade in the model with the total imports and total exports of each traffic-district in the data.

Figure 1.6: Model fit of trade flows over distance

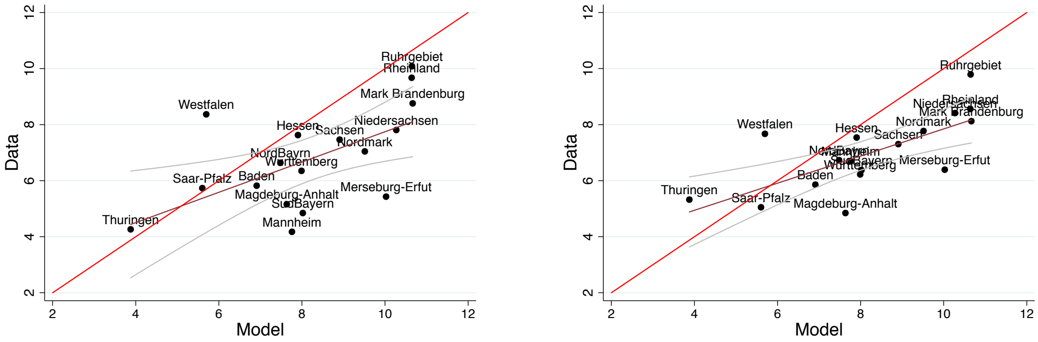


Notes: Share of total weight of good shipments by distance bins. The continuous line plots the density of goods shipments over distance in the goods traffic data while the dashed line plots the density of trade flows over distance in the model. Goods shipments only contain shipments by truck of manufacturing firms and furniture.

Figure 1.7 plots the goods traffic data against the model predictions. The correlation between the model and the data is corr=0.59 for total imports and corr=0.70 for total exports (the model explains 35% of the cross-sectional variation in imports and 50% of the variation in exports, as measured by the R-squared). Thus, the calibration presented in the previous sub-section does an excellent job in replicating trade flows across German districts before division.

These two tests show that the model has the ability to explain the spatial equilibrium of Germany and captures well the trade flows observed in the cross-sectional data.

Figure 1.7: Model fit of domestic trade



A) Domestic Imports, R^2= 0.35, corr=0.59 **B) Domestic Exports, R^2= 0.50, corr=0.70**

Notes: Each dot represents one traffic-district, there are 18 in total. Data comes from the Statistical Yearbook of the Bundesrepublic, year 1940. The road shipment data is collected in tons and split up by tons imported and tons exported to the rest of German districts.

Optimal highways before the division

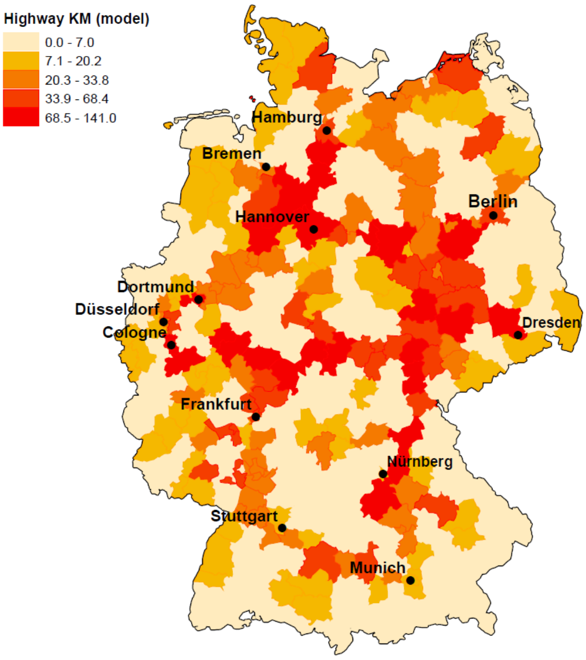
I will now use the model to compute the optimal investment before division. The solution of the model finds the share of the investment budget that should be allocated to each district to maximise aggregate Welfare. I solve for the optimal investment before division and compare it to the 1934 highway plan designed by the Nazi government.[42] Figure 1.8 compares the optimal number of kilometres per district in the model (upper figure) to the number of kilometres per district allocated in the 1934 highway plan (lower figure). The shading represents the number of kilometres of investment predicted by the model or allocated in the plan in each district.

The model is able to predict the main patterns of investment, such as the connections of the main German cities to Berlin (West to East and South to East links). In addition, the model under predicts investments at the border, due to the fact that the baseline model abstracts from international trade and other geopolitical concerns that may have been driving infrastructure investments. I will introduce international trade in the solution to the optimal investments post-division to understand the relevance of trade as a driver of infrastructure construction.

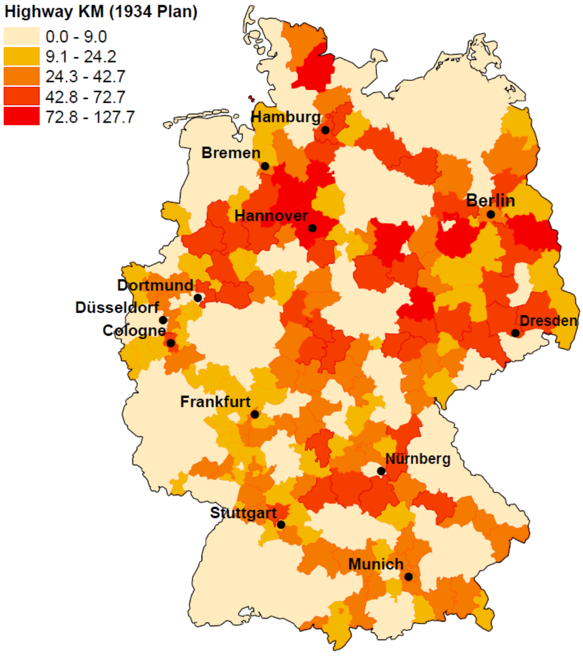
---

[42] To a reasonable comparison, I convert the share of investments in the model into highway kilometres by assuming that the number of kilometres built in the model is the same as the total kilometres that the government planned to build in the 1934 Highway plan

Figure 1.8: Simulated Infrastructure investments before the division shock



A) Model



B) 1934 Highway Plan

Notes: The shading represents the number of highway kilometres by district (darker, more kilometres). The upper panel displays the highway kilometres predicted by the model while the lower panel displays the highway kilometres allocated to each district in the 1934 Highway Plan.

The model is also able to capture the importance of the highway investments. In figure 1.A7 in the Appendix we can observe the comparison between the model's prediction and the first highways that were built before division. The model's predicted intensity of investment coincides as well with the highway construction timing. In the figure, the districts with higher marginal benefits of highway investments (darkest shades), are also the ones where first highways were built.

To test the predictive power of the model-generated investments, I run the following regression at the district level:

$$H_{i,Plan} = \alpha + \beta_1 H_{i,OPT} + X'\phi + u_{i,t}, \tag{1.21}$$

where $H_{i,Plan}$ is the total number of highway kilometres planned in district i and year 1934 and $H_{i,OPT}$ is the number of kilometres allocated to district i in the optimal network before division. The vector X' represents district-level controls, such as the distance to the border with Western Europe and the area of a district is squared kilometres. The coefficient of interest is β_1, that captures the relation between the model's predictions and the observed planned highways.

The identification of β_1 could be biased if there are district-specific characteristics that affect population and other fundamentals as well as highway investments before division. For example, being close to a port increases centrality of a region relative to the rest of the world and may have attracted population and trade as well as better infrastructure for decades. To control for some of the geography related time-invariant factors such as the ruggedness of terrain or proximity to the coast I run an additional specification that includes state-level fixed effects. It is clear that these controls will not dissipate the concern about the existence of omitted variables. In the next sub-section, I exploit the change in economic geography caused by the division of Germany to address some of these concerns. Table 1.2 reports the results of this specification and table 1.A3 in the Appendix reports the results with standard errors clustered at Government regions (Regierungsbezirke).[43]

---

[43] Clustering the standard errors at the Government region seeks to control for spatial correlation of errors. There are 30 Government regions.

Table 1.2: 1934 Highway Plan : Cross-sectional variation

OUTCOME:	Highway km Plan (1)	Highway km Plan (2)	Highway km Plan (3)
Optimal highway km	0.4496***	0.3213***	0.3285***
	(0.0568)	(0.0508)	(0.0504)
Pop 1938		3.8656**	4.6497**
		(1.5457)	(2.0711)
Distance to the Border		0.0576***	0.0578***
		(0.0188)	(0.0220)
Elevation		-0.0033	-0.0169**
		(0.0046)	(0.0068)
Area (sqkm)		0.0115***	0.0117***
		(0.0022)	(0.0025)
Constant	10.2302***	-48.3552**	-44.0867
	(1.2306)	(18.7763)	(29.3548)
State FE	No	No	Yes
Observations	331	323	323
R^2	0.265	0.381	0.451
Mean Highway km	20		
SD Highway km	24		

Notes: Robust HAC standard errors, are in parentheses.* significant at 10%, ** significant at 5%, *** significant at 1%. Results for the 395 German districts aggregated into 331 micro-census regions. The dependent variable is the number of highway kilometres in the 1934 Plan for each district. *Optimal highway km* is the highway kilometres predicted by the quantitative model. Distance to West Border measures the distance from the district centroid to the German border with a western European country.

The model-generated optimal network can explain 26.5% of the cross-sectional variation in planned highway across German districts. The coefficient $\beta_1 = 0.44^{***}$ shows the close relation between the predicted investments and the observed investments. This coefficient is reduced when I control for the surface of the district and for the distance to the border to $\beta_1 = 0.32^{***}$. The reduction in the coefficient shows that controlling for geographic factors seems to be very important to assess the actual predictive power of the model. The coefficient remains stable after including state-level fixed effects. The coefficients on Area and Distance to Border are positive showing that the 1934 Highway Plan allocated more kilometres to larger districts and to districts further from the country's border.

There are a few discrepancies between model and data that may worsen the fit of the model. These discrepancies are clear when looking at the spatial distribution of investments in Figure 1.8. First, the model's measure of investment is smooth and thus the investment pattern is more gradual than the highway pattern. Second, the model fails to capture the investments near the Eastern border of Germany. In 1934 the German territories extended further East than nowadays borders and, thus, in figure 1.8 we observe several planned highways that crossed the border with Poland, (former) Czechoslovakia and Austria. Finally, we do see that the plan invests heavily in the centre of the country while the model's solution is more radial.

In addition, the use of a linear regression to compare the two maps is based on comparing district-level investments but does not capture the spatial distribution of investments. To capture the spatial nature of the data, I compare the changes in bilateral transport costs that each district would enjoy after the construction of the model predicted investments and the (counterfactual) construction of the 1934 Highway Plan. I run the following regression,

$$\Delta T_{i,j}^{Plan-1938} = \beta_1 \Delta T_{i,j}^{OPT-1938} + v_{i,j}, \tag{1.22}$$

where $\Delta T_{i,j}^{Plan-1938}$=$TC_{ij}^{Plan} - TC_{ij}^{1938}$, the fall in bilateral transport costs after building the 1934 Plan, and $\Delta T_{i,j}^{OPT-38}$=$TC_{ij}^{Model} - TC_{ij}^{1938}$, is the fall in transport costs after the construction of the optimal network. The results of this regression, Table 1.A4 in the Appendix, show that the model captures very well the change in transport costs that would have taken place if the 1934 Highway Plan had been built. This shows that the connections to which the model gives priority coincide, to a large extent, with the goals of the government.

In all, these results suggest that a model with endogenous infrastructure investments performs well in predicting the cross-section of highway investments.

1.5.2. Test of the model after division

Population and Trade in changes after division

I now provide two additional tests of the ability of the spatial model with endogenous infrastructure networks to explain economic outcomes. To this end, I exploit the unexpected appearance of the border between East and West Germany after the Second World War. The division of Germany happened as a result of the increasing tensions between the United States and the Soviet Union. While the division was supposed to be temporary, it became permanent once the conflict between this two countries escalated.

The unexpected division of Germany can be used to test whether the structural model can capture the reaction of the economy to this division shock, that exogenously changed the Market Access and Centrality of all West German districts[44].

I perform two different tests. First, I compare the district-level population change between 1950 and 1974 in the data with the model's prediction. Second, I compare the change in domestic trade flows across German states between 1950 and 1974. Notice that both these tests assess the performance of the model to make out-of-sample predictions since the only data that has been used in the calibration of the model are $pre-division$ variables.

For these tests I need to simulate the division of Germany in the model. Following the division of Germany in 1949, West Germany suffered an important shock to trade and population mobility coming from the establishment of the inner German border. I simulate the division of Germany in the calibrated model by assuming that trade costs between East Germany and West Germany became prohibitive.

Population changes The first test is to simulate the division and predict the reallocation of population in West Germany in 1974 relative to 1950. To do this, I simulate the 1974 equilibrium population allocation using as input the observed 1974 highway network and the population level in 1974. To proxy for the effect of the division I compare the change in log-population of different districts relative to their proximity to the new border, following Redding and Sturm (2008):

$$\Delta ln(pop)_i = \alpha + \beta \text{dist to Border}_i + e_i \tag{1.23}$$

where i is district and $\Delta ln(pop)_i = ln(pop_{1974}) - ln(pop_{1950})$, is the change in log-population between 1950 and 1974. Table 1.3 reports the results.

---

[44] Redding and Sturm (2008) show that an economic geography model like the one I use in this paper can successfully capture the population response to the division shock of the 100 largest West German cities. I do a similar exercise including all districts and I extend the analysis to trade flows

Table 1.3: Out-of-sample test: Division shock on Population

$\Delta ln(pop)_{1974,1950}$	All micro-regions		Large regions		Medium-Small regions	
	Data	Simulation	Data	Simulation	Data	Simulation
Dist to Border (km)	0.0011***	0.0023***	0.0011***	0.0016**	0.0011***	0.0022***
	(0.0001)	(0.0003)	(0.0003)	(0.0007)	(0.0001)	(0.0002)
Constant	0.0035	-0.5517***	0.0718	-0.0446	-0.0161	-0.6942***
	(0.0231)	(0.0480)	(0.0477)	(0.1268)	(0.0260)	(0.0351)
Observations	252	252	70	70	182	182
R^2	0.220	0.163	0.189	0.068	0.226	0.275

Notes: Robust HAC standard errors, are in parentheses.* significant at 10%, ** significant at 5%, *** significant at 1%. Large micro-regions are micro-regions with more than 200,000 inhabitants, while the remaining regions are defined as small-medium. This threshold is chosen taking the mean population and adding half a standard deviation.

Comparing column 1 (data) with column 2 (simulated data) we can see that the model captures the negative effect of the border shock on population. According to my estimation, a district 100 kilometres further away from the border benefited from an increase in population (of 10% according to the data, $\beta = 0.001$) and that this pattern is also captured in the model (increase of 22%). Columns 3 and 4 repeat the estimation for large cities while columns 5 and 6 display the result for medium to small sized districts. As we can see, the model predicts better the response to the shock observed in large cities while it overpredicts the population response in small and medium-sized districts.

Changes in trade flows The second test is to compare the predicted change in domestic trade flows between 1960 and 1989 with the changes in the data. This test seeks to show the validity of the calibration of the model relative to the effect of highway construction. The goal is to quantify whether, using as an input to the model the change in highway construction after division, it can capture the change in aggregate trade flows. Because of lack of inter-state bilateral data for before 1960 I focus on the decades after the division. To do this, I compute trade flows between west German states given the highway network in 1960 and then recompute the trade flows for the highway network of 1989.

The change in trade flows is computed as the growth rate of trade flows across German states (log-change)[45]. To assess the performance of the model, I compare the

---

[45] The data is in tons while the simulated data from the model is in nominal value of flows

trade data with the model simulation. I use the model to predict the change in trade flows between 1960 and 1980, by computing trade flows after the division (1960) and trade flows just before reunification (1989). This counterfactual is calculated by feeding into the model the 1980 population and the changes in highway construction observed until then.

Table 1.4 shows the reduced form relation between log-changes in trade flows and changes in transport costs. As we can see, the model is able to capture remarkably well the response of trade flows to reductions in transport costs (column (1) vs column (2)). According to these results, a pair of states that benefited from a reduction in the ad-valorem transport cost similar to the mean (-1), saw an increase in trade flows of 19.5% in the data and of 14.5% in the model.

Table 1.4: Out-of-sample Test: Change in trade flows 1950 to 1980

Dep. Var: $\Delta log(Trade_{i,j})$	Data	Simulation
	(1)	(2)
Change in Transport costs (1980-1950)	-0.195***	-0.145**
	(0.0542)	(0.0613)
Constant	0.714***	0.295***
	(0.0597)	(0.0711)
Obs.	90	90
R-Squared	0.117	0.0435
Mean Change	-0.995	
St. Dev Change	0.51	

Notes: Robust HAC standard errors, are in parentheses.* significant at 10%, ** significant at 5%, *** significant at 1%. The Change in Transport costs is computed as the change in ad valorem transport costs from district to district, for the transport network in 1950 and the network in 1980. This variable is always negative going from 0 to -2. The changes in transport costs are aggregated by taking simple means to compute the state-to-state change in transport costs. The Data variable is log-change in tons shipped across each pair of states. The Model simulations use the log-change in nominal value of state-to-state trade.

Finally, I show that the model with infrastructure explains a larger part of the change in transport costs than other alternative counterfactuals. I consider two alternative simulations. First, I only consider the change in population, from 1960 to 1980 levels, while keeping infrastructure fixed as it was in 1960. In the second additional counterfactual I keep the infrastructure fixed as in 1960 and I allow for productivity to change so that the 1980 population distribution is perfectly rationalised. I plot the trade share changes in the data over the simulated trade changes. Figures 1.A5 and 1.A6 in the Appendix confirm the good performance of the model with infrastructure growth compared to the other two counterfactuals. The model with improvements in highway construction explains 11% of the change in trade flows observed in the data, while the other two counterfactuals explain less than 3%. In the two alternative counterfactuals the correlation between the measures is negative.

These two additional tests show that the static model I build in section 1.3 performs strikingly well in predicting long-term changes in population and trade flows. The ability of the model to predict a large part of the population and trade changes over a twenty-five year period should gives us a lot of confidence on the calibration of the model.

Optimal highway construction after the division

The existence of district-specific time-invariant factors (such as geographical advantages or political importance) poses a threat to identifying the predictive power of the model for the cross-section of highway investments. To address this threat, I exploit a large-scale exogenous shock to economic fundamentals: the German division. This exogenous shock allows me to test the model's predictions about new highway construction after the division shock. Focusing on the change in highway kilometres built between 1974 and 1950 controls for time-invariant characteristics affecting the stock of highways in that district. I will now explain how I solve for the optimal highway construction, after the division.

Constrained choice of infrastructure First, after division, the government's choice is restricted to building infrastructure in West Germany. The new government's objective function will be to choose the infrastructure in West Germany to maximise the aggregate welfare of West German districts. Finally, the existing highways that had been built between 1934 and 1950 are also used as constraints in the model (Figure 1.3). Recall that about one half of the six thousand kilometres outlined in the 1934 Highway plan had been built before the division (two thousand kilometres were built in West Germany, while one thousand were built in East Germany). To capture this physical constraint I add to the government's problem explained in section 1.3 the following lower bound constraint:

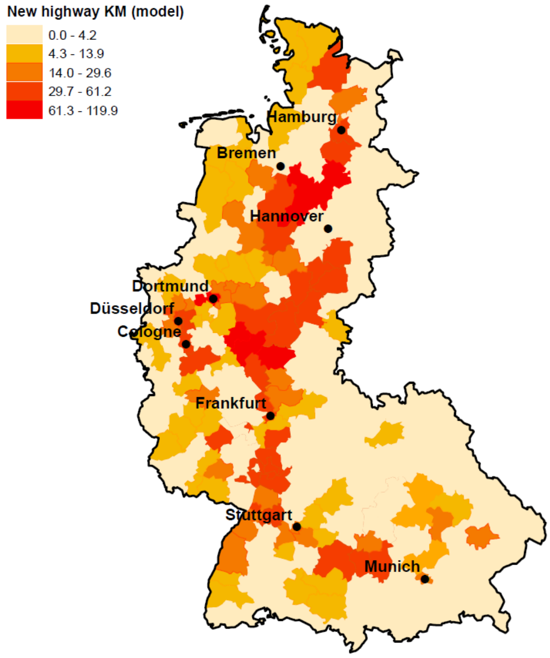
$$\phi_j^{Postdivision} \geq \phi_j^{Predivision}. \tag{1.24}$$

The addition of this constraint will allow us to compare the constrained solution in the model with the constrained solution in the data. To end, I assume that all other structural parameters remain unchanged. In this assumption I follow Redding and Sturm (2008) that interpret the division of Germany as mainly a trade and labour shock. They provide strong evidence showing that other factors such trends in specialisation, the integration with the Western trade partners or the fear of further armed conflict were important but to a much lesser extent compared to the trade shock. Table 1.A2 in the Appendix summarises the calibrated and estimated values for all parameters.

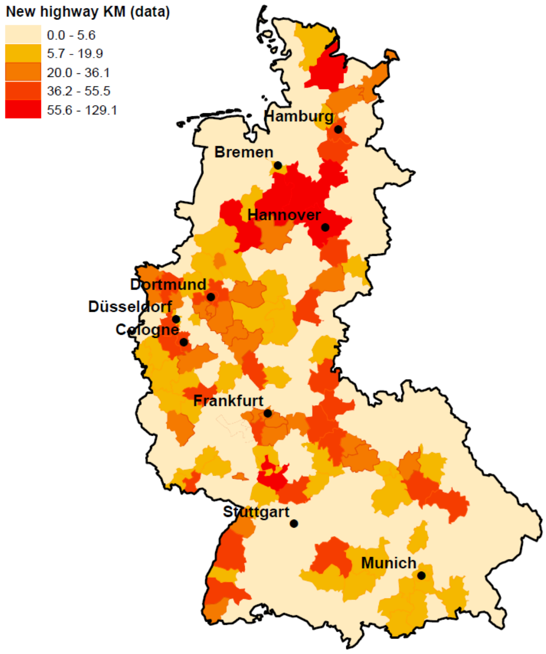
New highway construction in the data and in the model Figure 1.9 plots the spatial distribution of highway investments predicted by the model after simulating the division of Germany. The shades represent the investment allocation predicted by the model, with darker shades representing higher investments. This optimal (constrained) network serves as a benchmark of the highway investments we would observe if the government's choice was driven by the change in economic fundamentals that followed the German division. The figure plots the additional predicted investments, the change between 1950 and 1974. As documented in section 1.2, about half of the new construction deviated from the original 1934 Highway plan. I compare the predictions of the model relative to the change in highways between 1950 and 1974 to show that the model can account for the quantitative changes in the data.

An empirical challenge that we face when comparing the total amount of infrastructure investments across districts is the existence of district-specific characteristics affecting the economic fundamentals of that district, such as population or trade and the highway stock in the district. For example, districts in the mountains could have less population and trade less goods because of their remoteness and, at the same time, receive less highway construction because of the high construction cost. Thus, the elevation of the terrain can create a spurious positive correlation between economic fundamentals and highway construction.

Figure 1.9: Simulated Infrastructure investments after division - Baseline



A) Model (Baseline)



B) Data (1974-1950)

Notes: The shading represents the change in highway kilometres by district (darker, more kilometres). The upper panel displays the changes predicted by the model while the lower panel displays the highway changes observed in the data (new highway construction between 1950 and 1974).

In the cross-section, the finding that rugged districts have few highways would coincide with the low economic activity in these districts but this coincidence would come from the geographical disadvantage rather than from the relation between highways and economic fundamentals. This type of factors would induce a bias in the estimation of the predictive power of the model.

A good solution in the presence of time-invariant characteristics is to perform the analysis looking at changes in the variable of interest. Therefore, to test the ability of the model to predict highway construction I use the change in highway investments. As a measure of highway change I will use $\Delta H_i = H_{i,1974} - H_{i,1950}$, the change in highway kilometres in district i between 1974 and 1950. The model counterpart of this empirical measure will be the $\Delta H_{OPT} = H_{i,OPT} - H_{i,1950}$, the district-level highways built $after$ the division of Germany, taking into account the constraints mentioned before.

To test the predictive power of the model I run the following regression, at the district level:

$$\Delta H_{i,74-50} = \gamma_1 \Delta H_{OPT} + \nu_{i,t}, \tag{1.25}$$

where i denotes districts. The outcome measure, $\Delta H_{i,74-50}$ is the difference in highway kilometres in district i between 1974 and 1950. The main regressor ΔH_{OPT} is the allocation of the additional highway kilometres in the model, subtracting the initial investments realised by 1950. Because this is a first-differences specification I do not include district geographical characteristics or state fixed effects. This specification tests whether the predicted changes in highway investments in the quantitative model can explain the changes in highway construction observed in the data. Notice that the model-generated changes are purely driven by the simulation of the division shock, and therefore, only account for the economic factors affecting highway construction. The coefficient of interest is γ_1 that identifies the relation between the predicted changes and the observed highway reaction in the data. The results are reported in Table 1.5. The results with standard errors clustered at the Government region level are reported in table 1.A5 in the Appendix.

The estimated coefficients show that the model is also successful in explaining the changes in highway construction. We can think of these changes as the innovations implemented after the division shock. The predicted construction in the model explains 19% of the cross-district variation in highway changes (column 1). Furthermore, the model has strong predictive power even controlling important variables that can affect highway construction: Population in 1950, distance to the border, the district's surface and elevation and the number of kilometres allocated to the district in the 1934 Highway Plan. The model helps us explain part of the highway construction after division on top of all these factors.

Table 1.5: Change in highway construction

OUTCOME	Δ Kilometres (1974 – 1950) (1)	Δ Kilometres (1974 – 1950) (2)	Δ Kilometres (1974 – 1950) (3)
Optimal change in Kilometres	0.4101***	0.2307***	0.2350***
	(0.1106)	(0.0676)	(0.0719)
Log Pop 1950		3.9567**	3.7540*
		(1.7505)	(2.0213)
Distance to West Border		-0.0312***	-0.0326**
		(0.0109)	(0.0137)
Area (sqkm)		0.0037	0.0043*
		(0.0023)	(0.0024)
Elevation		-0.0038	-0.0078
		(0.0036)	(0.0064)
Plan 1934 km		0.3923***	0.3858***
		(0.0672)	(0.0701)
Constant	7.1523***	-43.1408**	-33.8805
	(1.2356)	(20.4614)	(26.9219)
State FE	No	No	Yes
Observations	258	257	257
R^2	0.188	0.441	0.447

Notes: Robust HAC standard errors, are in parentheses.* significant at 10%, ** significant at 5%, *** significant at 1%. Results for the 258 West German districts. The dependent variable is the change in highway kilometres between 1950, the end of the division and 1974. *Optimal change km* is the predicted increase in highway kilometres simulated in the quantitative model. highways built. Distance to West Border measures the distance from the district centroid to the German border with a western European country.

I find that an increase of one kilometre in the model after division predicts an increase of 0.41 kilometres in the data. In the last column I allow for state-specific time-deviations

(with state fixed effects) and the relationship between the model predictions and the observed changes remains positive and very significant. The fact that the coefficient is far from unity tells us that the model predicts a stronger reaction to the division shock than the one observed in the data. This attenuated reaction could be driven by constraints or adjustment costs in the real world that are absent in the model. Furthermore, this estimated coefficient of 0.41 could be explained by the existence of other factors driving highway investments such as political interests.

As we did before, I also measure the predictive power of the model looking at changes in the bilateral transport costs between all pairs of districts. I run a similar regression that correlates the change in ad-valorem transport cost between 1950 and 1974 with the model-implied change in ad-valorem transport cost under the assumption that the construction after 1950 follows the model solution:

$$\Delta T_{i,j}^{74-50} = \beta_1 \Delta T_{i,j}^{OPT-50} + v_{i,j}, \tag{1.26}$$

where $\Delta T_{i,j}^{74-50}$ is the change in bilateral transport costs between years 1950 and 1974, between districts i and j, and $\Delta \mathrm{T}_{i,j}^{OPT-50}$ is the change in bilateral transport costs between 1950 and 1974, assuming that construction in 1974 is equal to the model's prediction. Table 1.A6 in the Appendix, displays the results. The model does a very good job at predicting the change in transport costs during this period. And, crucially, the model performs *better* in predicting the change in transport costs than the 1934 Highway Plan.

Threats to Identification

Given the estimation strategy presented above, the main concern that should be addressed to identify the explanatory power of the model are changes in other factors, in addition to geography, happening right after the division of Germany and affecting the returns to highway construction unevenly across the West German geography. Factors affecting all West German regions simultaneously such as the re-construction of cities after the Second World War or the decline in the importance of the railway as the main mode of transport for freight should not bias the estimates presented in the previous subsection. As long as the effect is constant across all regions, these changes create a level effect that would be differenced out in the proposed estimation strategy.

However, changes affecting West German regions in an uneven way could bias the estimation of the predictive power of the model. One of such factors is the process of European integration during which tariffs to international trade were eliminated between Belgium, Italy, Luxembourg, Netherlands and West Germany. This process started with creation of the European Economic Community in 1957 with the Treaty of Rome. To account for this, I extend the model to allow for international trade with other West European countries and solve for the optimal highway construction with both domestic and international trade. The details about the introduction of international trade in the calibration of the model can be found in section 1.10 of the Appendix. Figure 1.A8 in the Appendix plots the optimal change in highway construction as predicted by the model with international trade (above panel) and the observed change in highway construction between 1950 and 1974 in the data (lower panel).

The results from the model extended with international trade are reported in Table 1.A7, counterpart of table 1.5. The optimal highway investment in the model still has good predictive for the actual change in highway construction but it predicts a smaller share of the variation in the data than the model without international trade (the R-squared of column 1 is 13.9% in the extended model compared to 18.8% in the baseline model). One explanation for why the model with international trade seems to be further away from the data could be that in the first years of the european integration international trade was not so large in magnitude as a share of domestic activity. In the year 1982, around 85% of all tonnes-kilometres shipped in West Germany by road were domestic shipments while only 15% were cross-border (international) shipments. However, the political intention to integrate with other West European countries may explain a part of the shift of highway investments towards the Western border of Germany, even if the international trade magnitudes increased very gradually.

Additional factors that we could consider are changes in the industrial composition of West Germany or the shift of the capital from Berlin to Bonn. Extending the model to account for industrial policy or for the benefits of a change in the capital city goes beyond the scope of this paper. However, the interplay between infrastructure policy, industrial policy and the institutional setting is a promising avenue for future research. For example, Bai and Jia (2018) show how the loss and gain of regional capital status in China was accompanied by upgrades in the transport network that resulted in a gain or loss of centrality of a given city in the transport network, creating a link between the political and the economic status of the city.

In this section we have seen that a quantitative model of endogenous infrastructure successfully explains a large part of the West German allocation of new highways between 1950 and 1974. In the next session I quantify the economic impact of highway investments in two quantitative exercises. First I examine the gains from the partial reaction of the government to the shock, comparing the welfare level for the observed highway network with the counterfactual of building the highway plan as it was designed in 1934. Second, I quantify the cost of the constraint imposed by the initial two-thousand kilometres built before the division.

1.6. Quantification: Economic impact of highway choice

In this section, I quantify the aggregate gains of reshaping the infrastructure network. I use the structural model to evaluate the welfare gains from different policy-relevant counterfactuals.

The first question we seek to answer is how important is the specific allocation of infrastructure in aggregate terms. Can governments improve welfare substantially by placing infrastructure in a sensible way (taking into account the economic geography of the country)? Or is infrastructure beneficial for the reduction of transport costs it generates, leading to large gains even across different spatial allocations? To understand this I take a data-driven approach: I quantify the gains in welfare accomplished by the considerable reshaping of the highway network that took place during the first decades of the division (as documented in the previous sections). I use the model as a measuring tool to quantify the gains from highway reshaping observed between 1950 and 1974.

The second question I examine is the cost of path-dependence. Infrastructure projects are long-lived and, therefore, unexpected changes in the economic geography of a country may reduce the value of the infrastructure network if it becomes obsolete. The construction of the first part of the 1934 Highway plan and the subsequent division of Germany is an extreme, but clear, example. In these cases, we would like to know how large are the costs of path-dependence from past infrastructure construction and how costly is it to overcome these loses. I use the structural model to solve for the optimal unconstrained highway network for West Germany, without taking into account construction before the division. I quantify the cost of path-dependence by comparing the optimal constrained highways, predicted by the model, with the alternative solution of the optimal unconstrained highway network. Both these exercises can help us understand how much the decision on where to place infrastructure investments can matter and the different trade-offs that may appear.

1.6.1. The gains from reshaping infrastructure

To understand the importance of the placement of infrastructure I compare the observed highway network in 1974 with the 1934 highway plan. The 1934 Plan can serve as counterfactual of a suboptimal network. Comparing the actual highway network in 1974 with the 1934 Plan has two advantages: First, the counterfactual comes directly from the data, from the digitised historical map. Second, both networks are of the same length, in terms of highway kilometres. Differences in aggregate measures come purely from the reallocation of construction across districts. I use the model as a measuring tool to compute welfare gains from the construction of the 1934 Highway Plan as well as of the 1974 highway network. Details about the construction of the model counterpart 1934 and 1974 network can be found in section 1.10 in the Appendix.

I take as baseline the 1934 Highway Plan that is the highway network that would have been built if the economic fundamentals in Germany had remained constant after 1949. I compare the gains of building the observed 1974 network in two cases, the baseline model with no international trade and the extended model that allows for trade with Western European countries. The results are reported in Panel A and B of Table 1.6.

The government's reshaping of 1600 kilometres (1/6 of the total West German network) increased welfare by 1.07% relative to building the 1934 Highway Plan with no adjustments. The gains in terms of real income were of 0.65% compared to the level under the 1934 Highway Plan. It is important to notice that these gains are an improvement that West German would enjoy permanently after the change. Furthermore these gains come purely from the reallocation of the highway network keeping the budget fixed.

Panel B shows the gains from the observed reshaping of infrastructure if we consider the potential trade flows with European neighbours. In this case the gains increase to 1.86% in terms of welfare and 2.02% in terms of real income. Thus, taking into account international trade uncovers a larger disutility from building the 1934 Plan, that was mostly oriented eastwards. This difference in the quantification highlights the importance of taking into account international trade in models of endogenous infrastructure.

The gains from reshaping infrastructure can also be found when looking at other indicators such as trade openness by district and inequality. Finally, we can also use the model to calculate the gains from highway construction compared to no highway construction. The welfare comparison for the cases of the 1934 highway plan, the 1974 network and the optimal constrained network are plotted in figure 1.A9 in the Appendix. I find that the 1934 highway Plan increased welfare by 14.82% while the 1974 highway network increased welfare by 16.06%. Furthermore, the optimal allocation of highways by the model provides a gain of 24% in terms of welfare, which is an increase of 8% compared to the government allocation. These quantitative results are similar to other studies in the literature such as the construction of the highway network in Brazil where highways were estimated to increase welfare by 13% (Morten and Oliveira, 2018).

Table 1.6: Gains from reshaping the highway network

	Compared to 1934 Highway Plan
Panel A: Highways 1974 - baseline	% Change
welfare	**+1.07**
Real Income	+0.65
Panel B: Highways 1974 - int. trade	% Change
welfare	**+1.86**
Real Income	+2.02
Effects on trade - baseline	
Openness (Intra-district trade)	**+ 14**
Effects on inequality	
Inequality (Variance of real GDP)	- 7

1.6.2. The cost of path-dependence

Finally, this set-up allows me to quantify the aggregate cost of path-dependence. The construction of the German highway network started in the 1930s when the division could not be imagined. In this additional exercise I compare the welfare level of the optimal constrained network, that includes the 2000 initial highway kilometres, with the unconstrained network.

I use the model to compute an additional counterfactual highway allocation that assumes the West German government could have build the highway network from scratch. Table 1.7 reports the welfare and real income gains from building the constrained and the unconstrained network, taking the 1974 highways as reference. Figure 1.A10 in the Appendix reports the welfare and income gains relative to the no highway case. As we can see, the optimal constrained network would create welfare gains of 7.5% compared to the 1974 network (panel A). The construction of the optimal unconstrained network (panel B) would provide gains of 9.1%.

The difference between panel A and panel B tells us about the cost of path-dependence. My counterfactual exercise suggests that the construction of 2000 kilometres before the division of German had a cost of 1.6% of the 1974 welfare level and of 1.675% in terms of 1974 real income. The cost comes from the fixed nature of the initial 2000 highway kilometres. Therefore, pre-division highway investments created a binding constraint for the government and the losses from this constraint were large. It is important to notice that these quantification is relevant for West Germany and does not take into account the re-unification of Germany in 1989.

1.6.3. Discussion of results

Finally, let us put in perspective these welfare and real income gains associated with the different counterfactual networks. The quantifications so far were considering the aggregate gains of different highway network allocations of the same length. As we could see, there are considerable gains from different infrastructure investment patterns.

What are the aggregate economic gains of infrastructure construction? Taking as baseline the 1974 highway network, I find that eliminating all highways would cause welfare to fall by 16% and real GDP by 8%. As we mentioned before, if we change the 1974 Highway network for the 1934 Highway Plan, welfare would fall by 1% and

Table 1.7: The cost of path-dependence

	Compared to Highways in 1974
Panel A: Constrained optimal network	
welfare	**7.512**
Real Income	5.659
Panel B: Unconstrained optimal network	
welfare	**9.119**
Real Income	7.344

real GDP by 0.65%, thus by 6.25% of the total gains from building the 1974 highway network. This shows that the government's response to the division increased the gains from highway construction by a large magnitude. The gains could have been larger, going from the 1974 highways to the optimal constrained network would increase welfare by 8% and real GDP by 6%. Figure 1.A9 and 1.A10 in the Appendix plot the level of aggregate gains for all the counterfactual networks.

The magnitudes found in this quantification depend on the calibration of the parameters of the model. The last exercise I do is to put bounds on the reshaping gains by looking at different values of two very important parameters: the trade elasticity σ and the returns to highway investments γ[46]. Table 1.8 summarises the results. As we can see, the estimates reported below may fall or increase slightly but it seems clear that there are positive welfare gains from reshaping infrastructure and costs of path-dependence, since the intervals are far away from zero.

1.7. Conclusion

Understanding the economic effects of transport infrastructure is essential for policy-makers to take informed decisions about expensive and persistent investments. However,

---

[46] For σ I consider the extremes of the confidence interval estimated when calibrating the trade elasticity using trade flows. For γ I consider 0.5 and 1, following Fajgelbaum and Schaal (2018)

Table 1.8: Welfare Gains: Bounds around quantification results

Welfare Gains	Min	Max	Bounds for $\gamma = \{0.5, 1\}$ σ=7	Bounds for $\sigma = \{6, 8.5\}$ γ=0.84
Reshaping highways	0.68	1.44	(0.84, 1.11)	(0.81, 1.32)
Path-dependence cost	0.69	1.62	(0.97, 1.52)	(1.31, 1.60)

the endogeneity between infrastructure investments and economic outcomes is an important challenge to common empirical and quantitative methods. The reduced-form approach of exploiting exogenous variation in infrastructure construction is appealing to estimate local effects but cannot be employed to quantify the aggregate effects of infrastructure projects.

In this paper I take a structural approach to this problem by building a quantitative spatial trade framework with endogenous infrastructure investments. In the model, a government decides how to allocate investments across regions to maximise aggregate welfare. This framework allows me to characterise the optimal infrastructure network, calibrate the model using historical data and estimate the key structural parameter, the returns to highway investments, taking into account the endogeneity of infrastructure.

I use the calibrated model to study the economic impact of highway construction in the context of the division of Germany. The division of Germany provides an ideal set-up to test the ability of the model to predict highway construction after the division and to estimate the gains from reshaping the network in response to the shock. Using newly digitised data, I document that half of the highway kilometres built after the division, between 1950 and 1974, deviated from the initial prewar Highway Plan. I find that the reallocation of these investments (one third of the network) led to increases of 1.08% of welfare and 0.64% of real income annually, keeping the budget fixed. In the extended model with international trade, the gains are even larger: the new 1974 highway network increased welfare by 1.8% and real income by 2%. Finally, I measure the cost of path-dependence and show that reshaping the full network could have increased welfare by 1.62% and real income by 1.5%.

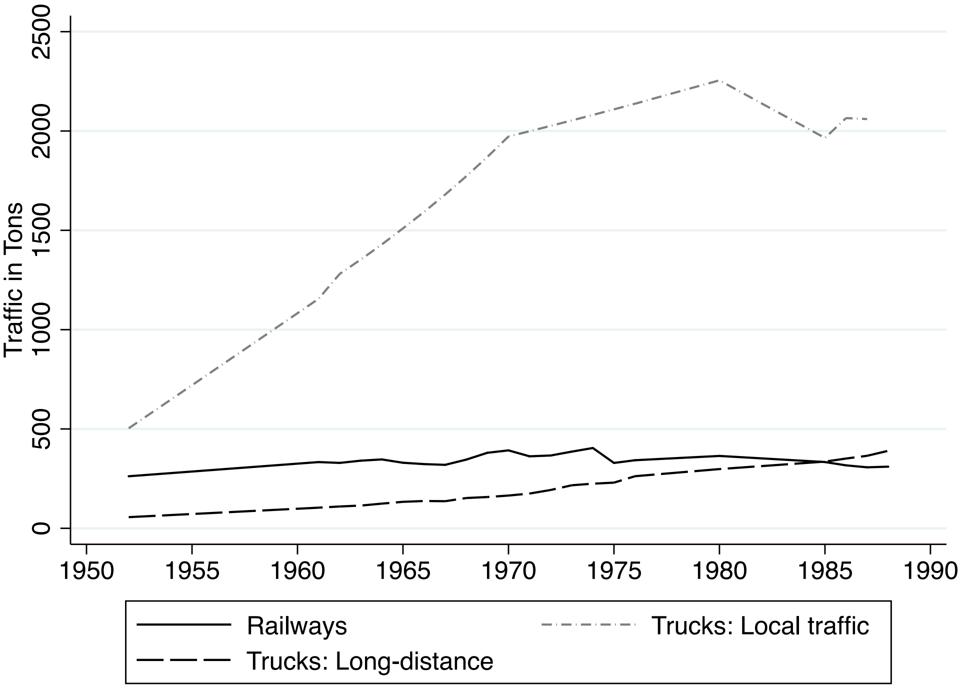
The magnitude of these reshaping gains is large relative to current estimates in the literature. For example Asturias et al. (2014) find gains of 2.7% of real income from the construction of the Indian highway network of almost six thousand kilometres and

Fajgelbaum and Schaal (2017) find welfare gains of between 0.9% and 1.5% from reshaping the totality of nowadays Germany's highway network. There are two main differences between my paper and other studies. First, I examine the gains from reacting to a change in economic fundamentals. My findings suggest that the economic impact of infrastructure is larger in this case than when evaluating the impact of infrastructure investments in the presence of stable economic fundamentals. Second, my results are estimated from the construction of the initial part of the highway network. It is likely that returns to investments decrease as infrastructure is accumulated. However, these differences make my results particularly relevant for countries that are going through structural reforms or large policy changes and for countries building the first stages of the infrastructure network. Making use of a quantitative framework like the one developed in this paper can help governments quantify the expected gains across different investment allocations.

There are several related questions that need to be addressed in relation with infrastructure choice more generally. First, what other factors shape the investment decisions of governments? The importance of political factors can be estimated with a framework of endogenous infrastructure that includes economic fundamentals and political incentives. Second, how does the optimal network decision change with the introduction of additional mechanisms such as intermediate input usage, heterogeneous agents or international trade? Finally, what is the optimal infrastructure policy to address different spatial settings such as fast-urbanising countries or deeply integrated free-trade areas? Expanding this framework in the mentioned directions is left for future research.

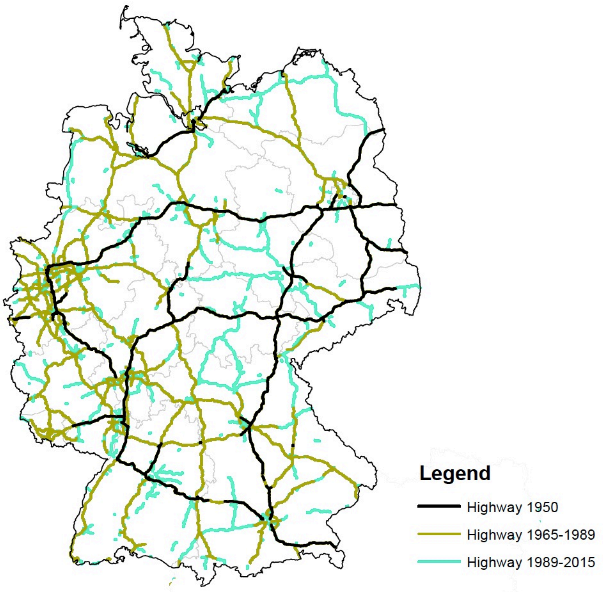
1.8. Appendix of Figures and Tables

Figure 1.A1: Goods traffic by transport mode



Notes: Values are for West Germany, collected by author from Statistical Yearbook of the Bundesrepublik, multple years. Values for missing years are interpolated.

Figure 1.A2: Evolution of the German highway network



Notes: German highway data collected from Michelin Atlases of the years 1950, 1964, 1975, 1980 and 1989 digitised by the author.

Figure 1.A3: Construction and Planning of the Highway Network

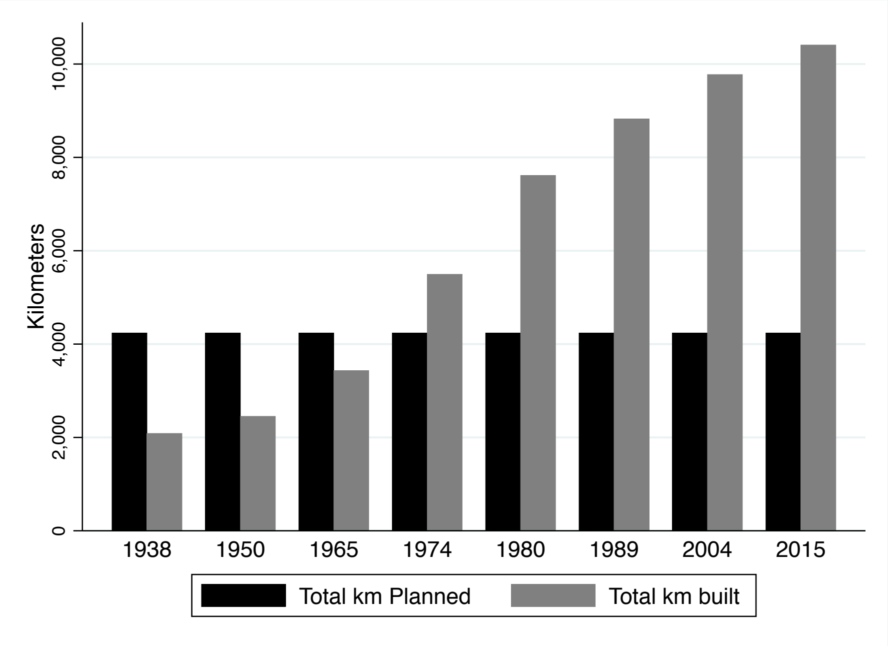
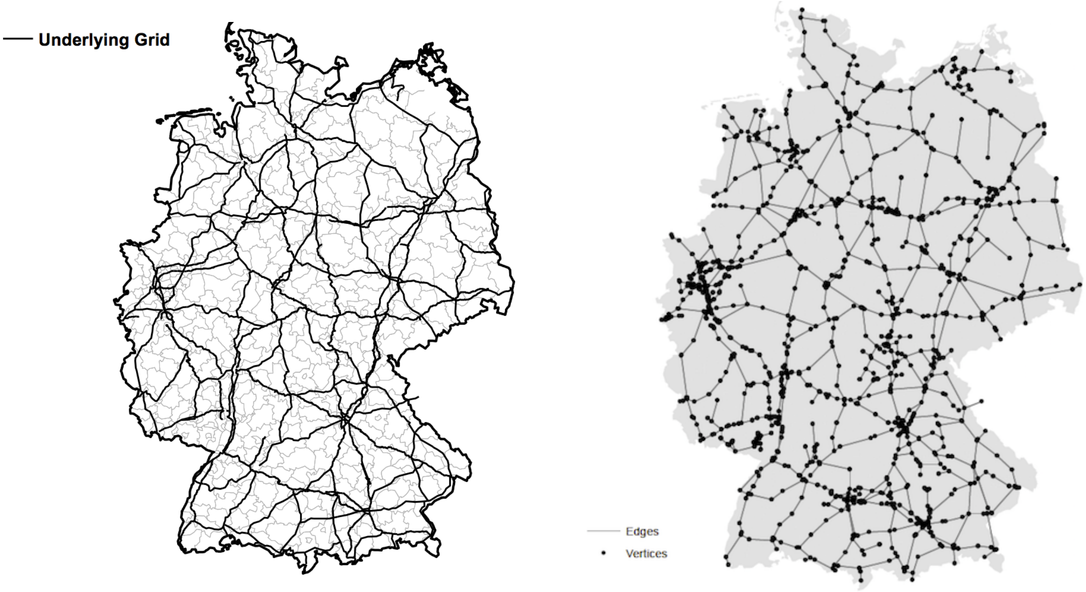


Figure 1.A4: Representative transport network and corresponding graph

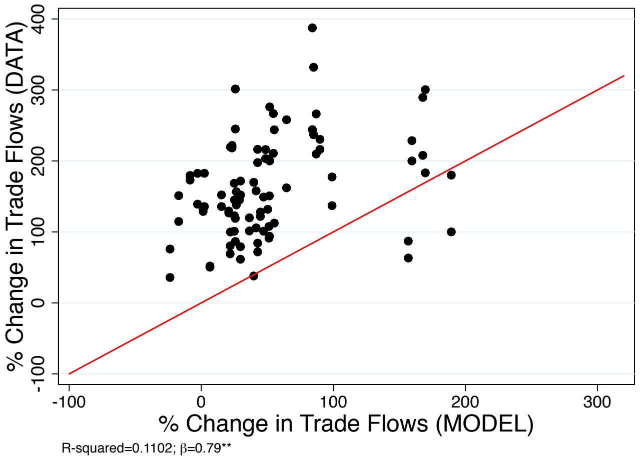


A) Chosen network

B) Graph

Notes: Panel A shows the roads that I choose to build the grid. This network connects all districts while the number of links remains small. Highways are the darkest lines, federal highways are the intermediate lines, and local roads are the thinnest lines. Panel B shows the discretisation of the network in panel A. Each dot represents a vertex, and each line represents an edge of the network.

Figure 1.A5: Model fit - Change trade flows over change in Infrastructure



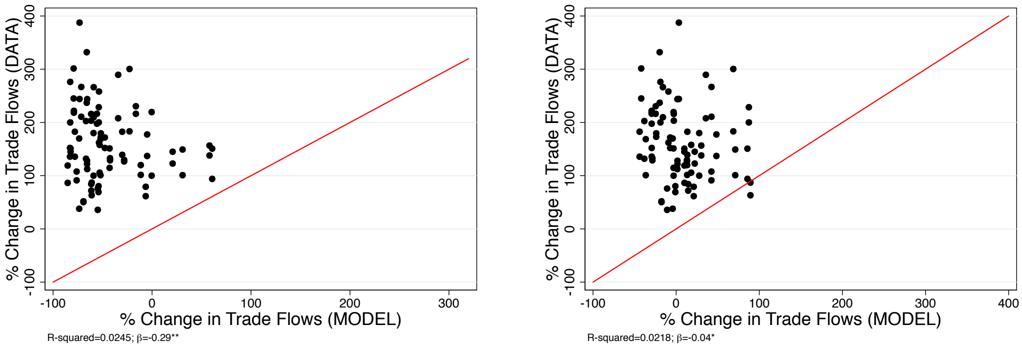
Effects of Change in Infrastructure

1966-1989 (in percentage)

R-sq=11%, β = 0.79**

Notes: The figure plots the observed change in shipments by road (in percentage) between each West German state pairs. Within state shipments are also included. Each dot is one state-pair.

Figure 1.A6: Model fit - Changes in trade flows (Alternative explanatory variable)



Panel A: No change in highways

1966-1989 (in percentage)

R-sq=2.45%, $\beta = -0.29^{**}$

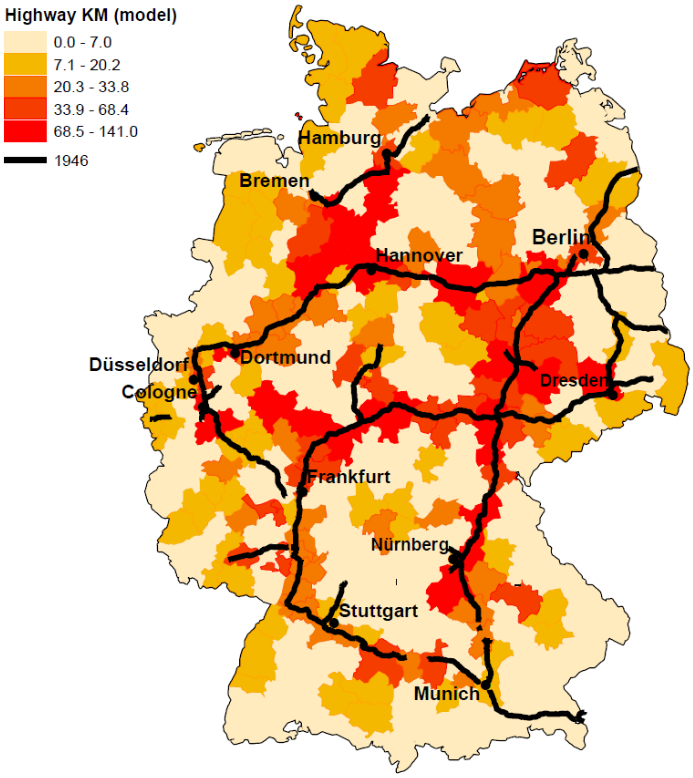
Panel B: Change in Productivities

1966-1989 (in percentage)

R-sq=2.18%, $\beta = -0.04^{*}$

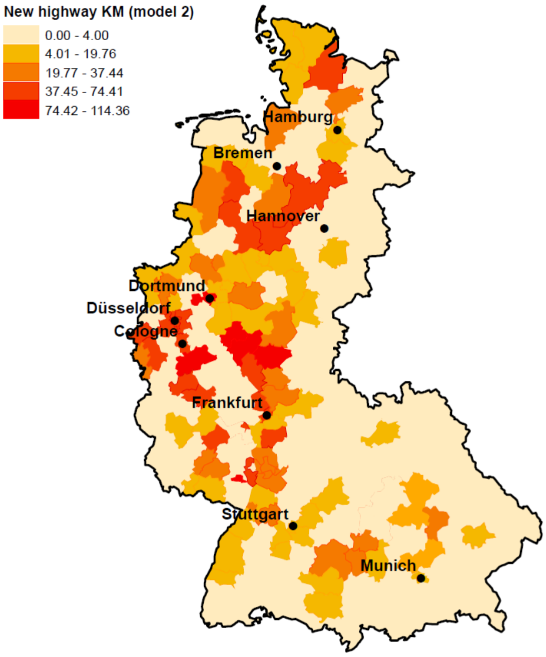
Notes: Both panels plot the observed change in shipments by road (in percentage) between each West German state pairs. Within state shipments are also included. Each dot is one state-pair.

Figure 1.A7: Simulated Infrastructure before the Division shock - Timing of Construction

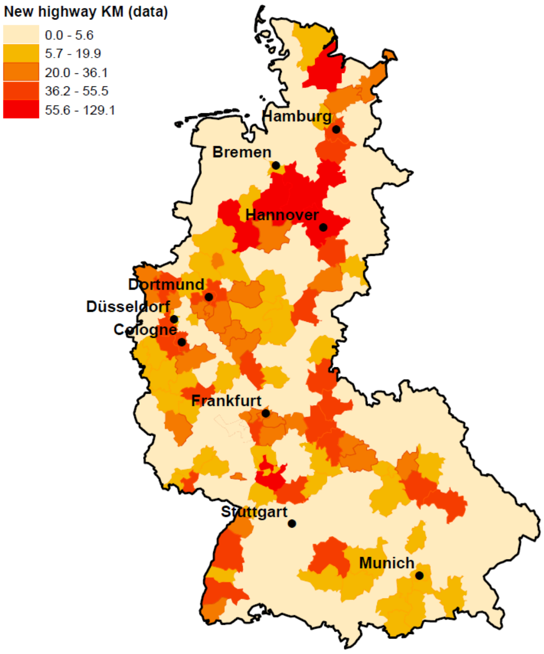


Notes: The shading represents the investment allocation by district, in terms of kilometres. The model predicts the optimal allocation of the investment budget to each district, as a share of the budget. I convert the share of investment into highway kilometres by assuming that the total number of kilometres built in the model is the same as in the 1934 Highway plan. The black lines represent the highways that had been built by 1946.

Figure 1.A8: Simulated Infrastructure investments after Division - International trade



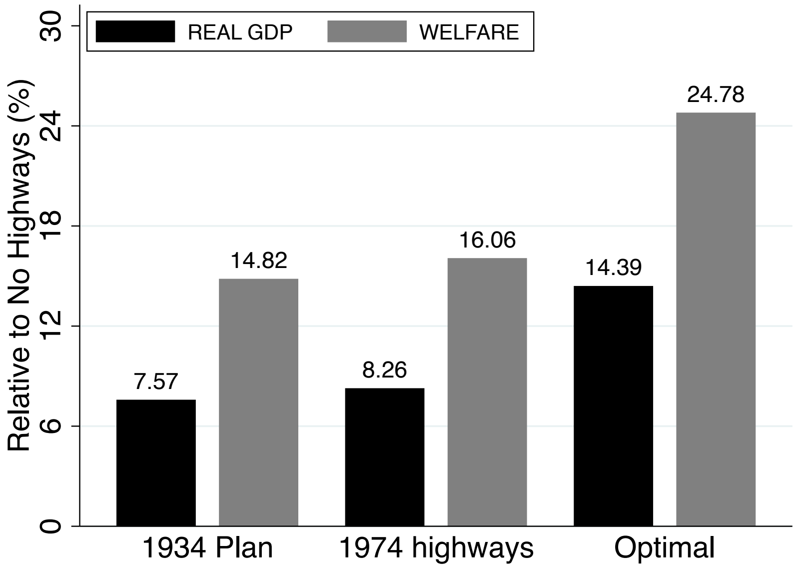
A) Model (International trade)



B) Data (1974-1950)

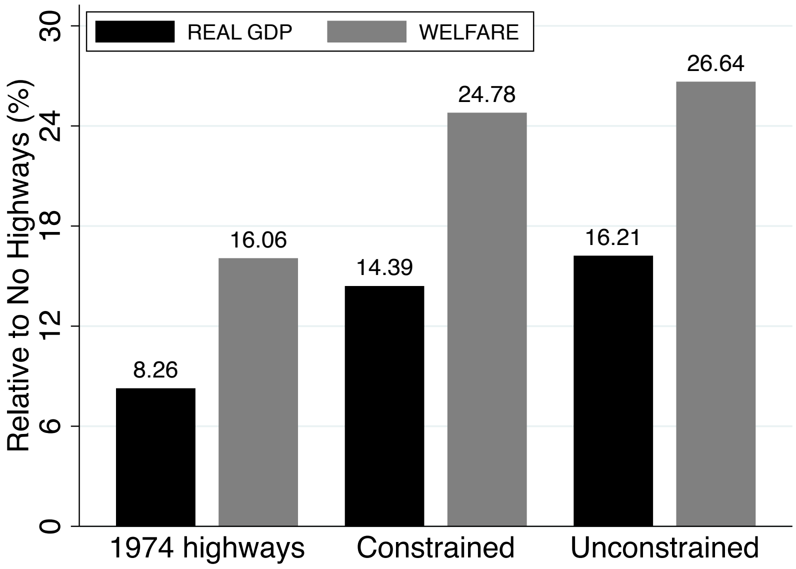
Notes: The shading represents the change in investment allocation by district. The upper panel displays the changes predicted by the model while the lower panel represents the highway changes observed in the data (new highway construction between 1950 and 1974). Darker shades indicate higher highway construction.

Figure 1.A9: Gains from Infrastructure Investments



Gains from reshaping

Figure 1.A10: Gains from Infrastructure Investments



Cost of Path-dependance

Table 1.A1: Elasticity of trade flows to distance

Outcome: Log (Road shipments in tons)	(1) DATA 1938	(2) MODEL ($\sigma = 7$)
Log(Distance)	-2.8674***	-2.7808***
	(0.2381)	(0.3762)
Constant	28.7349***	35.9057***
	(1.4651)	(2.3151)
Observations	13	13
R^2	0.929	0.832

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Standard errors, are in parentheses. Regression run using total tons shipped by truck by manufacturing firms over 13 distance brackets (from less than 50km to more than 1000km). Model regression using simulated trade data given parameter values and infrastructure in 1938 aggregated over the same distance brackets

Table 1.A2: Choice of Parameters

Parameter	Description	Source/Target	Value
From Literature			
ε	Shape parameter of Fréchet	Monte et al. (2015)	3
α	Share of tradables	Redding, Sturm (2008)	0.7
1938 Germany			
$\{A_i\}$	Productivity parameter	Match population 1938	
$\{H_i\}$	Land supply	Area in sqkm	
σ	Elasticity of substitution	Trade elasticity 1938	7
Infrastructure			
γ	Returns to highway investments	Concentration in 1934 Plan	0.84
Z	Budget of Government	Average km in Plan	1.5*regions
c	Marginal cost	-	1
Division Shock			
$B_{i,j}$	Cost of Border East Ger.-West Ger.	Prohibitive trade costs	∞

Notes: Further details about the calibration and estimation of the parameters can be found in section 1.4 in the main text and in section 1.10 of this Appendix.

Table 1.A3: 1934 Highway Plan : Cross-sectional variation

OUTCOME:	Highway km Plan (1)	Highway km Plan (2)	Highway km Plan (3)
Optimal highway km	0.4496*** (0.0675)	0.3213*** (0.0566)	0.3285*** (0.0583)
Pop 1938		3.8656** (1.6137)	4.6497* (2.3317)
Distance to the Border		0.0576** (0.0244)	0.0578** (0.0241)
Elevation		-0.0033 (0.0080)	-0.0169 (0.0104)
Area (sqkm)		0.0115*** (0.0027)	0.0117*** (0.0021)
Constant	10.2302*** (1.7620)	-48.3552** (20.3512)	-44.0867 (34.0212)
State FE	No	No	Yes
Observations	331	323	323
R^2	0.265	0.381	0.451
Mean Highway km	20		
SD Highway km	24		

Notes: Standard errors clustered at Government Region level, are in parentheses.* significant at 10%, ** significant at 5%, *** significant at 1%. Results for the 331 German districts. The dependent variable is the number of highway kilometres in the 1934 Plan for each district. *Optimal highway km* is the highway kilometres predicted by the quantitative model. Distance to West Border measures the distance from the district centroid to the German border with a western European country.

Table 1.A4: 1934 Highway Plan : Change in Transport Costs

OUTCOME	Transport cost 1934 Plan (1)	Transport cost 1934 Plan (2)	Transport cost (1934 Plan) (3)	Δ TC (Plan-1938) (4)
1938 Transport costs	0.5272***		0.2954***	
	(0.0012)		(0.0022)	
Optimal Transport costs		2.2899***	1.1600***	
		(0.0052)	(0.0097)	
Δ TC Optimal				0.6016***
				(0.0014)
Constant	0.2592***	-1.4559***	-0.7011***	-0.2166***
	(0.0032)	(0.0071)	(0.0085)	(0.0019)
Observations	77246	77246	77246	77246
R^2	0.724	0.713	0.767	0.716
Mean dep. var	1.62			
SD dep. var	0.512			

Notes: Standard errors, are in parentheses.* significant at 10%, ** significant at 5%, *** significant at 1%. . Results for the 395 German districts aggregated at the micro-region level. The dependent variable is the level of bilateral Transport costs under the counterfactual construction of the 1934 Highway Plan or the change in bilateral transport costs between 1938 and the construction of the 1834 Plan. $OptimalTransportCosts$ is the ad-valorem transport cost computed under the optimal investment predicted by the model. $1938TransportCosts$ measures ad-valorem transport costs under the 1938 network. $\Delta TCOptimal$ is the difference between Optimal and 1938 transport costs.

Table 1.A5: Change in highway construction: Highway kilometres

	Δ Kilometres (1974 – 1950) (1)	Δ Kilometres (1974 – 1950) (2)	Δ Kilometres (1974 – 1950) (3)
Optimal change in Kilometres	0.4101***	0.3628***	0.3295***
	(0.1180)	(0.0896)	(0.1019)
Distance to West Border		-0.0184	-0.0196
		(0.0132)	(0.0199)
Area (sqkm)		0.0087***	0.0096***
		(0.0028)	(0.0028)
Constant	7.1523***	1.4961	
	(1.7108)	(3.7236)	
State FE	No	No	Yes
Observations	258	258	258
R^2	0.188	0.243	0.475

Notes: Standard errors clustered at Government Region level, are in parentheses.* significant at 10%, ** significant at 5%, *** significant at 1%. Results for the 258 West German districts. The dependent variable is the change in highway kilometres between 1950, the end of the Division and 1974. *Optimal change km* is the predicted increase in highway kilometres simulated in the quantitative model. highways built. Distance to West Border measures the distance from the district centroid to the German border with a western European country.

Table 1.A6: Change in Highway construction: Transport Costs

OUTCOME	Δ TC (1974-1950) (1)	Δ TC (1974-1950) (2)
Δ Transport Cost Model	0.7368***	
	(0.0004)	
Δ Transport Cost 1934 Plan		0.9768***
		(0.0009)
Constant	-0.0384***	-0.0746***
	(0.0006)	(0.0010)
Observations	96721	96721
R^2	0.970	0.923
Mean dep. var	1.75	
SD dep. var	0.45	

Notes: Standard errors, are in parentheses.* significant at 10%, ** significant at 5%, *** significant at 1%. . Results for the 312 West German districts aggregated at the micro-region level. The dependent variable is the change in bilateral ad-valorem transport cost between 1950 and 1974. $TransportCostsOptimal$ is the ad-valorem transport cost computed under the optimal investment predicted by the model. $\Delta TCOptimal$ is the difference between Optimal and 1950 transport costs while $\Delta TCPlan$ is the difference between Transport costs in the 1934 Plan and 1950.

Table 1.A7: Change in highway construction: Model with International Trade

OUTCOME	Δ Kilometres (1974 – 1950) (1)	Δ Kilometres (1974 – 1950) (2)	Δ Kilometres (1974 – 1950) (3)
Optimal change in km	0.3352***	0.2026***	0.2065***
	(0.0875)	(0.0551)	(0.0595)
Log Pop 1950		4.6325***	4.3295**
		(1.7236)	(1.9214)
Distance to West Border		-0.0263**	-0.0266**
		(0.0112)	(0.0135)
Area (sqkm)		0.0038	0.0043
		(0.0026)	(0.0026)
Elevation		-0.0022	-0.0053
		(0.0037)	(0.0064)
Plan 1934 km		0.4118***	0.4077***
		(0.0707)	(0.0743)
Constant	7.9858***	-52.3630***	
	(1.0721)	(20.1133)	
State FE	No	No	Yes
Observations	258	257	257
R^2	0.139	0.436	0.600

Notes: Robust HAC standard errors, are in parentheses.* significant at 10%, ** significant at 5%, *** significant at 1%. Results for the 258 West German districts. The dependent variable is the change in highway kilometres between 1950, the end of the Division and 1974. *Optimal change km* is the predicted increase in highway kilometres simulated in the quantitative model. highways built. Distance to West Border measures the distance from the district centroid to the German border with a western European country.

1.9. Theoretical Appendix

1.9.1. A model of trade with endogenous infrastructure investments: No housing

In this section I describe in detail the model presented in section 1.3. The model features costly trade across many domestic districts, $i = 1...N$, endowed with an exogenous productivity, A_i. There is a measure L of workers that move across districts according to their own heterogenous preferences. This model builds on the family of widely used quantitative spatial models reviewed by Redding and Rossi-Hansberg (2017) and is specially close to Redding (2016). The main contribution of my theoretical framework is to embed the spatial equilibrium framework in the decision of a Government whose goal is to choose the optimal infrastructure investments to maximise aggregate welfare. The consumption, production and location decisions of consumers and workers are standard. The model, however, features a realistic geography where transport costs are determined as the solution to the problem of shipping across a grid of districts at the minimum cost.

Preferences The preferences of each worker are given by two components. First, a heterogenous preference taste (b), that represents how much a given worker values a given location (Redding (2016)). Second, a consumption component (C) that can be represented by a canonical CES demand system, with every agent choosing the level of consumption of each of the varieties available with a constant elasticity of substitution across varieties of σ. Specifically, the utility function of an agent ω living in district n is given by:

$$U_n = b_n(\omega)\left(\frac{C_n}{\alpha}\right) \tag{1.27}$$

where $C_n = \left[\sum_j^N \sum_v^M c_{jn}(v)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{(\sigma-1)}}$, is the consumption basket chosen by workers living in district n, $c_{jn}(v)$ is the consumption of a worker that lives in district n of variety v produced in district j[47]. M is the number of available varieties in the economy. The price index associated with the tradable varieties aggregator C_n is $P_n = \left[\sum_j^N \sum_v^M p_{jn}(v)^{1-\sigma}\right]^{\frac{1}{(1-\sigma)}}$, where $p_{jn}(v)$ is the price in district n of variety v produced in district j. The taste

---

[47] Notice that workers living in n face the same consumption prices and earn the same wage so they make the same consumption choices

component $b_n(\omega)$ is an idiosyncratic taste preference. Each worker draws a vector of N realisations $\{b_n(\omega)\}_{n=1...N}$ from a Frechet distribution that governs the individual preferences for each district:

$$G_n(b) = [Pr(b_n(\omega) \leq b)] = e^{-b^{-\epsilon}} \tag{1.28}$$

where ϵ is the shape parameter governing the dispersion of tastes across workers for each location. A large ϵ implies a low dispersion (low standard deviation). Less dispersion means that the idiosyncratic preferences are more equal across districts for all workers. In this case a small difference across districts will trigger big movements in population. In the limit, $\epsilon \rightarrow$ all workers behave identically. They become indifferent between locations and the model collapses to the *perfectly mobile labour* case because all districts are perceived as equally desirable and tiny changes in wages trigger large population reallocations. When $\epsilon \rightarrow 1$, highest dispersion, workers are very heterogeneous in their taste. This means that large differences in district-level outcomes are needed to make workers move from their preferred choices. In this type of models the labour supply in a district is upward slopping in the wage.

To keep the number of parameters of the model low and due to data limitations I abstract from consumption of housing in the model. The existence of heterogenous preferences will act as a dispersion force through wages and, under certain parameter values, the existence of the equilibrium will be unique and stable.

Production Production uses labour as only factor of production, happens within firms and takes the form of monopolistic competition. There is a fixed cost to pay to start production, F, but once a firm enters the market it will produce a differentiated variety. The existence of the fixed cost and free entry ensures that each variety will only be produced by one firm. This means that each district will produce a specific and unique set of varieties that will equal the sum of varieties produced by the firms in that district. All varieties are produced with the same technology that is district-specific, A_i. From the firm's profit maximisation we know that a firm producing a variety v in location i will set a price $p_i = \mu \frac{w_i}{A_i}$ where μ is the mark-up charged over the price $\mu = \frac{\sigma}{\sigma - 1}$. Notice that the price is constant across varieties produced in the same district. As we can see, each agent in the economy is endowed with one unit of labor that is supplied inelastically to produce A_i units of the district-specific varieties. I assume there is only one sector in the whole economy. The existence of free entry in each location drives down profits to zero and will pin-down the size of a firm in each district:

$$q_i^F(\nu) = \frac{F * A_i}{(\mu - 1)} = A_i(\sigma - 1)F \tag{1.29}$$

As we can see more productive districts will have larger firms because they will be able to cover the fixed cost more easily. Given the scale of each firm and the local labour supply, the labour market clearing condition pins down the number of varieties (equal the number of firms) in each district:

$$M_i = \frac{L_i}{\sigma F} \tag{1.30}$$

Again, we see that larger districts will produce a larger number of varieties. Therefore, the productivity of a district determines the scale of its firms and the size of a district determines the number of varieties locally produced. Finally, we can re-write the optimal price index of tradables P_i in terms of the local price and the number of varieties produced in each district, taking into account that all varieties in a given location have the same price and substituting the number of varieties in each district:

$$P_n = \frac{1}{\sigma F}\left[\sum_j^N L_j p_{jn}^{1-\sigma}\right]^{\frac{1}{(1-\sigma)}} \tag{1.31}$$

The price index in a district will depend on the local prices of imported varieties with larger regions exporting a larger share and therefore having a higher weight on the price index.

Location and Consumption choices Given the specification of preferences, we can write the indirect utility function of worker ω in district n as

$$v_n(\omega) = \frac{b_n(\omega)w_n}{P_n} \tag{1.32}$$

Since indirect utility is a monotonic function of the idiosyncratic preference draw, it has a Frechet distribution too:

$$G_n(u) = Pr[U_n \leq u] = e^{-\Psi_n u^{\epsilon}} \tag{1.33}$$

where $\Psi_n = (\frac{v_n}{P_n})^{\epsilon}$. $G_n(u)$ is the distribution of indirect utility realisations in district n.

Each worker chooses the location that maximises her indirect utility. Using the properties of the Frechet distribution we find that the probability that a worker chooses to live in district n:

$$\pi_n = Pr[U_n \geq max\{U_s; s \neq n\}] = \int_0^1 \underbrace{\prod_{s \neq n} [1 - G_n(U)]}_{Pr(U_s<u), s\neq i} \underbrace{dG_n(U)}_{\text{cdf } U_n} = \frac{(\frac{w_n}{P_n})^{\epsilon}}{\sum_k (\frac{w_k}{P_k})^{\epsilon}} \tag{1.34}$$

The fraction of workers that choose to live in district n coincides with the probability that any given worker chooses n:

$$L_n = \frac{(\frac{w_n}{P_n})^{\epsilon}}{\sum_{k=1}^{N} (w_k/(P_k))^{\epsilon}} L \tag{1.35}$$

As we can see ϵ is the elasticity of the labour share in any district to changes in real income income in that district w_n/P_n. Workers are more likely to choose districts with a relatively high real income.

Consumption is determined by the CES preference structure over varieties. The demand for variety v produced in district i and consumed in district n is:

$$x_{i,n}(v) = \frac{p_{i,n}^{1-\sigma}}{P_n^{1-\sigma}} w_n L_n \tag{1.36}$$

where $P_n = \frac{1}{\sigma F} (\sum_k L_k p_{k,n}^{1-\sigma})^{1/(1-\sigma)}$ is the price index of consumption goods in district n and $w_n L_n$ is the total expenditure in district n. Because each district produces a different set of varieties, the demand in district n for goods produced in district i (import share) will be:

$$X_{i,n} = \frac{L_i}{\sigma F} \frac{p_{i,n}^{1-\sigma}}{P_n^{1-\sigma}} w_n L_n \tag{1.37}$$

A district will import less goods from more expensive destinations (high p_i) and will import more goods from other districts, relative to domestic consumption, if it is more expensive (has a high Price index P_n). The above expression displays the Home market effect from the Krugman (1980) model: A larger district (high population) will produced a larger share of varieties and therefore export larger shares to other districts (notice that trade share X_{in} is increasing in L_n).

Finally, applying the same steps as before, we can compute the expected utility for each worker ex ante that is equal to the utility of the economy as a whole, ex post:

$$E(U) = \Gamma\left(\frac{\epsilon}{\epsilon - 1}\right)\left(\sum_{1}^{N}(w_n/(P_n))^{\epsilon}\right)^{1/\epsilon} \tag{1.38}$$

Geography The set of N districts can be thought of as located in a two-dimensional grid where each square represents a different district and we let the population be concentrated in the centre of the square. A more complex geography will be introduced in the calibration exercise. I define the size of district i, D_i, as the length of side of the square representing the region. We assume, as mentioned, that all the population is located in the centroid of the district and all the production happens there.

Given the dispersion of production and population across the grid, there are ad-valorem type of transport costs associated with the consumption of non-locally produced varieties. The price of variety i consumed in district n is given by:

$$p_{i,n} = p_i B_{i,n} T_{i,n}(D, \Phi) \tag{1.39}$$

The transport friction $T_{i,n}(D, \Phi)$ depends on two variables: $\{D\}$ and $\{\Phi\}$. The vector $\{D\}$ represents the size of each district in geographical terms while $\{\Phi\}$ represents the level of infrastructure in each district. Finally, the vector $\{B\}_{i,n}$ denotes the border friction between districts i and n that will be defined by:

$$B_{i,n} = \begin{cases} 1 & \text{if } i, n \text{ are within the same border} \\ > 1, & \text{if } i, n \text{ are separated by a border} \end{cases} \tag{1.40}$$

Transport Costs The transport cost that has to be paid to move a good produced in district i to be consumed in district n will be given by the minimum possible transport cost that can be realized when shipping between i and n. Therefore, the transport cost matrix will be the collection of the transport costs along the least cost path between each district pair. This set-up is similar to assuming the existence of some shadow transport sector that is driven by profit-maximization, and thus, ships goods at the minimum costs (see Fajgelbaum and Schaal (2017)). The transport cost minimization problem will consider both geography, D, that determines the distance that has to be covered in a given trip, and the quality of infrastructure, Φ, that will determine how costly (fast) can we cover this distance. While geographic distances are exogenous, infrastructure is *endogenous*, and can be upgraded to reduce transport frictions. A further simplifying assumption I make is that the quality of infrastructure is homogeneous within a district[48]. These two variables, distance and infrastructure quality, define the cost of shipping some good across district i as follows:

$$\text{Cost of transiting}^i = \frac{D_i}{\phi_i^\gamma} \tag{1.41}$$

where I use ϕ_i to denote the district-level infrastructure and Φ to denote the vector of infrastructure allocations. I assume that $\phi_j \geq 1$, so that the transport cost cannot be larger than the geographical cost but will be always bounded above by the physical geography[49]. The elasticity of the ad-valorem transit cost to the infrastructure investments is denoted by γ, that is model-specific and I assume it to be $\in (0, 1)$. This restriction ensures that there are decreasing returns to infrastructure and that the marginal benefit of infrastructure investments is concave.

As we can see, a higher quality of infrastructure will reduce the size of a district in terms of the ad valorem cost that is incurred in the transit. The elasticity of the shipping cost across district i to a marginal change in the level of quality of infrastructure Φ_i is $-\gamma$. In the quantitative exercise $\{\Phi\}$ is the quality of the road (highway, state highway or local road) and will affect shipping cost because of the different maximum speeds at which each road can be transited.

---

[48] In the real world a district may have one very high quality highway and one very low quality road. Therefore, we may think of ϕ_i as the average quality of the infrastructure stock in district i. Fajgelbaum and Schaal (2017) adopt the decision of defining the quality of infrastructure for every link but this implies solving the optimal investment for every link rather than keeping the district as the unit of analysis

[49] This is a very reasonable assumption in the context of land-shipping because there is always some residual way to go from one region to another even in the absence of roads (i.e. across the fields)

Optimal path matrix The solution to the least cost path problem to connect each district-pair will depend on distance, D, and the quality of infrastructure, Φ. For some given geography and a given level of infrastructure investment in each district this solution can be expressed using an optimal path matrix. This matrix indicates whether district i is a transit district for all the other district-pairs along their optimal cost path. It is similar to the *transition matrix* in the network literature as it indicates how to transition from one node of the network to any other. The element $\Pi^i_{o,d} \in \Pi^i$ indicates whether district i is relevant for the determination of the cost of shipping from district o to district d and is defined as:

$$\Pi^i_{o,d} = \begin{cases} 1, & \text{if } i \text{ is a transit district in } \textit{the path between o and d} \\ 0, & \text{if } i \text{ is not a transit district in } \textit{the path between o and d} \end{cases} \tag{1.42}$$

I can now define the transport cost $T_{o,d}$ between any two districts o and d as

$$T_{o,d} = \left[\sum_i \Pi^i_{o,d} \frac{D_i}{\phi_i^\gamma} \right] \tag{1.43}$$

where ϕ_i is the infrastructure level in district i and it reduces the travelling time across district i with an elasticity of $-\gamma$. This expression means that the cost of shipping a good from o to d will be the sum of the geographical distance scaled by the infrastructure quality of all the districts that have to be transited to reach destination d from origin o. Given that we have defined $\{ \Pi^i \}$ as the optimal path matrix this implies that we can also express the transport friction between o and d as $T_{o,d} = min_k(T(p^k_{o,p}))$ where $T(p^k_{o,p})$ is the transport cost of shipping a good from o to d along path k.

Finally, I adopt a normalisation common to all trade models by assuming $T_{i,i} = 1$, equivalent to assuming free intra-district trade and normalising the cost of trading out of the district by the internal shipping cost.

1.9.2. Spatial equilibrium

The spatial equilibrium is given by the clearing of the goods market at the equilibrium wages and the clearing of the labour market so that the expected utility for all workers is equalized. The goods market clearing implies that the income of each district has to equalise the total exports of that district (assuming trade is balanced):

$$w_i L_i = \sum_j \frac{L_i}{\sigma F} \left(\frac{\sigma}{\sigma - 1} \frac{w_i}{A_i} b_{i,j} T_{i,j} \right)^{1-\sigma} (P_j)^{\sigma-1} w_j L_j, \forall i \in N \tag{1.44}$$

where the Price index in each region will be:

$$P_i^{1-\sigma} = \sum_j \frac{L_j}{\sigma F} \left(\frac{\sigma}{\sigma - 1} \frac{w_j}{A_j} b_{j,i} T_{j,i} \right)^{1-\sigma}, \forall i \in N \tag{1.45}$$

The utility equalisation condition will hold when the labour share in each district n is given by:

$$\frac{L_n}{L} = \frac{(\frac{w_n}{P_n})^{\epsilon}}{\sum_{n=1}^{N} (w_n/(P_n))^{\epsilon}} \tag{1.46}$$

The labour market clearing implies that the sum of labour in each district will add up to the total number of workers: $\sum_n \frac{L_n}{L} = 1$. These conditions determine the equilibrium value of wages, labour shares and the equilibrium welfare given by expected utility E(U), equation 1.38. As shown in Allen and Arkolakis (2014) the condition for the existence and uniqueness of the spatial equilibrium will hold if the elasticity of expected utility to the labour share in a district is negative, this is, if the dispersion forces are stronger than the agglomeration forces of the model. Specifically, the condition for existence and uniqueness of the equilibrium with imperfect mobility of labour is

$$\sigma \left(1 - \frac{1}{1 + \frac{1}{\epsilon}} \right) > 1 \tag{1.47}$$

I calibrate the parameters so that the equilibrium is unique and stable.

1.10. Quantification details

1.10.1. Quantification of the model before Division

Initial transport network We have 3 types of roads in Germany in 1938: Highways, Federal Highways and Local roads. To construct the initial transport grid I choose the smallest set of edges and vertices that allows me to represent the underlying geography of Germany to transport goods. First I select the set of vertices to represent the 412

German districts I observe in the data. I choose as the vertex of the district the centroid of the path of any highways that transits the district. If there is no highway in the district I use the centroid of the federal highway inside the district. If there are no highways or federal highways I use the centroid of the local road. Second I build the set of edges that connects the vertices that represent the population centres. To do this I select all highways and federal highways that existed in 1938. I add the set of local roads needed to connect the remaining vertices that do not have highway access.[50]

Finally, I export the network to use in my quantitative analysis. The resulting network is composed of 1071 junctions (nodes), collected into 412 districts) and around 1200 edges (links) that can be exported as two vectors: one containing the links and one containing the cost of transiting each link (called weight in the networks literature).

Initial transport costs To compute the initial transport cost matrix I follow (Combes and Lafourcade, 2005) transport cost function. The function is derived to account for the cost of shipping one truck full of goods in France in the decade of 1978. The transport cost specification for shipping a truck between i and n is:

$$t_{i,n} = \text{Distance cost} \times \frac{\text{speed}_{i,n}}{\text{length}_{i,n}} + \text{Time cost} \times \frac{\text{speed}_{i,n}}{\text{length}_{i,n}}. \tag{1.48}$$

The table below represents the speed and costs assumed by (Combes and Lafourcade, 2005) that I use in the computation. I use this function to compute the cost of shipping one truck worth of goods for each link in the network using the length in kilometres along the underlying German network in 1938. I assume links with local roads can be transited at 40 kilometres per hour, link with federal highways at 60 kilometres per hour and links with autobahns at 80 kilometres per hour. Even if there was no speed limit in Germany for highways trucks can rarely go faster than 80 km/h. I use the Network Analyst toolbox in ArcGIS to construct the network that is then exported to MATLAB.

I compute cost of transit in each link using the above function and the actual kilome-

---

[50] To select this edges I choose the least cost path to connect each of the 57 districts that are not transited by a highway or a federal road to the closest district with federal highway using the ”Closest facility Tool” in ArcGIS that allows you to extract the path chosen to connect facilities (federal highway points) to incidents (district centroids). Local roads are used as the default way to more around to prevent any transport costs to be zero. Instead of manually recovering the local road network in 1938 I use the 2004 digitised map of local roads and enabled a truck to move through these links at 40 km per hour.

Table 1.A8: Cost of time and distance of truck shipping (France, 1978)

	Highway (Autobahn)	Federal Highway	Local road
Speed (km per hour)	80	60	40
Cost of distance (Euros)	85.8	89.8	97.18
Cost of time (Euros)	5.4	5.4	5.4

tres. I use the least-cost path algorithm to compute the matrix of initial transport costs in euros. To convert this measure to ad-valorem quantities I normalise the computed cost in euros by 28,000 euros that is the average cost of a truck full of German goods in the year 1995. This computation uses the average export price per ton from Germany to France. This normalisation ensures that the transport cost matrix expresses the cost of shipping one unit of the average German good across any district pair in ad-valorem terms.

1.10.2. Optimal infrastructure network

Estimation of parameter γ I estimate γ to match the skewness of investments in the 1934 highway plan. Skewness is a measure of the concentration of a distribution. Matching this moment ensures that the concentration of highway investments in the model is aligned with the data. I use the Simulated Method of Moments for the estimation. I simulate 100 times the optimal choice of infrastructure in a representative 50-district economy with 100 different random draws of the vector of district-specific productivities. I compute the skewness of these investments and estimate the value of γ that minimises the sum of squared differences between average skewness in the model and skewness in the data. Skewness in the data is 1.5. For the simulation I specify the productivity distribution as a Pareto distribution with shape parameter $\alpha_p = 1.6$, estimated from the calibrated productivity distribution for Germany, scale parameter $\sigma_p = 1$ and location parameter $\theta_p = 0$.This procedure yields an estimate of $\gamma = 0.84$.

1.10.3. Optimal infrastructure network with International Trade

To introduce international trade post-Division I consider trade with Belgium, France and Netherlands. I assume that trade with the rest of the world is only possible through the West German districts located in the border with these countries for which some highway had been designed in the prewar Highway Plan or for which some local road existed.

To model the new trading opportunities I choose to increase the population of the bordering regions with a share of the foreign population, so that access to these bordering regions allows a firm to sell products to the domestic population and to the foreign population as well. I assume that trade is possible with the whole population in the foreign countries but I compute a cost of trading with these foreign population equal to the average distance between the German border and the main foreign cities/capital city (for Belgium and Netherlands). I reduce the accessible population by a share to account for this distance cost. This simplifying assumption of considering trade opportunities as an increase in the size of regions at the border allows me to follow the same calibration strategy as before: I re-calibrate the productivity vector to match the new population distribution where the bordering regions have been allocated extra population coming from the foreign countries. The rest of the calibration procedure is the same as described previously.

1.10.4. Counterfactual exercises

Taking the highway network to the model To construct the model counterpart of the 1974 highway allocation I follow two steps. First, I compute the district share of highway kilometres. This is obtained by dividing the total highway kilometres in a district by the total highway kilometres built in 1974 in West Germany. Then, I multiply the highway share by the total budget allocated in the model to the Post-Division network, as follows:

$$\phi_i = 1 + share_{74} * (Z - 312), \tag{1.49}$$

where I subtract 312 from Z because that is the lower bound imposed by the requirement that highway investments cannot increase the transport costs and

$$share_{74} = \frac{\text{Highway km}_{i,1974}}{\sum_i^N \text{Highway km}_{i,1974}}. \tag{1.50}$$

In the same way, I build the counterpart of the 1934 plan as

$$\phi_i = 1 + share_{Plan} * (Z - 312), \tag{1.51}$$

1.11. Data Appendix

Highway data The highway network data (Autobahns) collected for the empirical exercise is of two types. First, I digitise the highway network plan of 1934 from historical documents. From the digitised data I construct a district level measure of the number of kilometres that the 1934 highway plan allocated to each district. Besides, I collect data of the actual highway network (only Autobahn) in Germany (both East and West Germany) for the years 1938, 1950, 1965, 1974, 1980 and 1989. This information is obtained from different road atlases and historical maps and geo-referenced using the software ArcGIS. Once the maps and atlases are digitised I manually collect the data to construct the highway network in each period. Additionally, I collect and geo-reference the pattern of federal roads in 1950 and 1965. Federal roads (Bundesstrase) are decided by the central government but are not restricted-access roads like the Autobahns and the network was developed earlier than the highway network. Finally, the network of local roads is imported from the EuroGlobal map by *Eurographics* that provides harmonised European open geographical data covering 45 countries and territories in the European region and is freely available. The website address of Eurographics is https://eurogeographics.org.

Economic outcomes As economic outcomes, I use population data by decade at the district level from the historical census. I also collect employment level at the district level from the historical census, available for some aggregated sectors.

Additionally, I collect traffic of goods by road for 18 aggregated traffic districts in Germany. The traffic data is collected in tons and reported in an aggregated way (Total tons of goods sent to the rest of Germany and received from Germany). The traffic data is collected from the "Statistisches Jahrbuch fr das Deutsche Reich". I use data from the year 1938, the closest to the beginning of the construction of the highway network. The scanned photocopies of the annual editions of the "Statistisches Jahrbuch fr das Deutsche Reich" are available at http://www.digizeitschriften.de/dms/toc/?PID=PPN514401303.

Finally, I collect and digitise traffic of goods by road between West German states. This data is available only after 1960 (most recent data I found was 1966). The traffic data is collected in tons and reported in an aggregated way (Total tons of goods sent to each state and received from each state in West Germany). This traffic data is collected and digitised from the "West Germany Road freight transport 1945- Statistics Serials" (Der Fernverkehr mit Lastkraftfahrzeugen: Zusammengefasste bersichten zur Gterbewegung).

Geographic variables As controls, I collect a series of measures related to the geography of Germany such as area of districts and distance to relevant points such as the inner German border. First, I measure the distance from each district to the closest point of the inner German border, to the closest point to the external West German border and West Berlin. I calculate these distances from the centre of each district to the geographic feature of interest over a straight line. Furthermore, I compute the distance to West Berlin through the transport network in 1950. Finally, I collect the district area in square kilometres.

Chapter 2

Measuring Borders within Europe

Joint with Jaume Ventura (CREi, Universitat Pompeu Fabra and Barcelona Graduate School of Economics) and Ugur Yesilbayraktar (Universitat Pompeu Fabra and Barcelona Graduate School of Economics).

2.1. Introduction

The creation of a single common European market for goods, services and labor is often portrayed as the main European policy achievement of the second half of the 20th century. And there are good reasons for this. The main idea behind the European Union and the myriad of treaties that go with it, such as Schengen or the Eurozone, was to eliminate the deleterious effects of national borders on trade and factor mobility. Sharing a supra-national government that regulates markets, it was argued, would foster European trade, which in turn would promote growth and raise the welfare of European citizens. Alongside this idea, there was also the shared belief that creating a common market would deter war by making it more costly. Despite recent signs of discontent around Europe, it seems fair to say that these predictions turned out to be correct. The creation of the single European market has coincided with an unprecedented period of peace and prosperity that stands out in the long and conflictual history of Europe.

An interesting research question is to determine the extent to which the European Union has managed to erase the negative effect of national governments on trade. This effect is often referred to as the “border” effect, and it constitutes our main object of study. We ask here: Has the European Union managed to fully eliminate the border effect? If not, how large is it still?

One strategy to answer these questions would be to directly examine the policies of countries. For instance, tariffs have been completely removed within Europe and this definitely reduces the ability of national governments to interfere with trade across borders. But there are many other ways through which national governments obstruct trade, effectively creating a border effect. For instance, trading some goods and services may require legal enforcement of contracts. Domestic courts are biased in favor of domestic importers and tend to rule against foreign exporters. Thus, promoting trade in those goods/services requires unbiased union courts that can overrule biased domestic ones. Another example is the adoption by national regulators of technical standards that impede foreign access to the domestic market for some goods/services. Trade in those goods/services becomes possible only if union regulators impose common standards to all member countries. A central goal of the European Union is to eliminate these non-tariff barriers to trade through the creation of union courts, the adoption of common standards, the harmonization of economic regulation, the creation of a common currency and other similar measures. A quantitative assessment of this maze of regulations and bureaucratic structures is highly desirable, but it seems too difficult to carry out at this point.

We therefore adopt another strategy that has a long tradition in the field of international trade. We look at trade outcomes rather than policy measures. That is, we compare trade within countries with trade across countries to detect whether national borders still act as a deterrent to trade. To implement this approach, we need to overcome two challenges. The first one is to find appropriate data on intra-national trade. There is plenty of data on trade across national borders, but there is a somewhat surprising scarcity of reliable data on trade within national borders. A first contribution of the paper is the construction of a new matrix of bilateral trade that covers 269 regions from 24 European countries. We construct this trade matrix using the European Road Freight Transport survey collected by Eurostat. This survey records information about the shipment of goods by road across sub-national regions in all European countries. From these data we are able to observe the origin and destination of the shipment, the industry of the type of good shipped, the volume of the shipment and the distance covered. An important advantage of this micro-dataset is that it provides us with an estimate of flow of goods within regions as well as across domestic and foreign regions. We aggregate these data and impute export prices to build a matrix of trade flows between 269 regions in 9 industries for the period 2011 to 2017.

The second challenge is to design a strategy to identify the border effect. Once we focus on trade outcomes, we need to recognize that there are many other determinants of these outcomes besides policy measures. We also need to recognize that borders are not assigned randomly. That is, we need a sound identification strategy for the border effect. Here we use two approaches, both of which build on the classic gravity approach. The first approach follows standard practice in trade studies and uses the entire data set. We then try to deal with border selection bias by using standard controls for geography and culture. The second approach is novel, as we restrict the sample to border regions and compare trade between contiguous regions in the same country, and contiguous regions in different countries. This strategy, which is possible only because of our new rich dataset, allows us to control for geographical and cultural distances in a more direct and perhaps effective way.

The main result can now be stated as follows: the border effect is still sizable in Europe! Ceteris paribus, intranational trade is about 4 times larger than international trade. The "ceteris paribus" clause here is important. Our statement is not unconditional, there are many reasons why intranational trade should be larger than international trade. What we are estimating here is the border effect *after* controlling for economic size, geographical and cultural distances. As usual, to convert this finding on quantities into a measure of costs, we need to make assumptions about the trade elasticity. Using a standard elasticity of five, our finding suggest that the border effect is equivalent to a tariff of about 30%. Both empirical strategies lead to similar estimates. These estimates vary somewhat across industries, but not much over time.

The paper is organized as follows. Section 2 describes the dataset. Section 3 describes our first identification strategy, presents its results and performs a variety of robustness tests. Section 4 expands on these results by examining how the border effect varies across industries and time. Section 5 describes our second identification strategy, presents its results and discusses differences and similarities with those obtained in section 3. Finally, section 6 concludes.

Related Literature In his pioneering study, McCallum (1995) revealed that Canadian inter-provincial trade was on average 22 times larger than cross border state-province trade with United States. This astounding result was confirmed by subsequent studies for North America (Helliwell, 2002), OECD (Wei, 1996) and Europe (Nitsch, 2000) and even extended to sub-national borders by Wolf (2000). Although the simple fact that borders hinder trade was not in itself very surprising, the magnitude of the effect came as a shock and explanations based on conventional national trade barriers such as tariffs, quotas and regulatory differences in standards and customs were not able to reconcile the size of the border coefficient (Head and Mayer, 2000).

As a consequence, initial efforts focused predominantly on exposing methodological shortcomings that could potentially lead to systemic overstatement of the border coefficient. Most prominently, Anderson and Van Wincoop (2003) derive a gravity equation consistent with the theory and argue that previous estimates suffered from omitted variable bias as they failed to account for "marginal resistance terms". Building on this huge theoretical leap, controlling for multilateral resistance terms either by employing non-linear least squares (Anderson and Van Wincoop, 2003) or using fixed effects (Feenstra, 2002) lead to a significant decline in estimated border effects. Furthermore, use of log-linear OLS came under scrutiny due to concerns regarding its performance in the presence of heteroskedasticity and its inability to incorporate zero trade flows (Helpman et al. (2008), Silva and Tenreyro (2006)). As a consequence, inclusion of more flexible estimation methods such as Poisson-Pseudo Maximum Likelihood (PPML) and Gamma-Pseudo Maximum Likelihood (GPML) became customary (Magerman et al., 2016).

In parallel to the advances in gravity methodology, economists explored the possibility that border effects may have been mis-measured in a way that leads to overestimation. In particular, there are three ways in which data quality matters. First, border dummy is introduced into the gravity specification as a part of the trade cost function. Hence accurate measurement of distance and other bilateral variables that proxy for trade frictions is critical to having a precise estimate for the border dummy (Head and Mayer, 2000). While the literature heavily relied on great-circle distances for inter-regional trade, there has been little to no consensus on how to measure within region distances. In general, these methods fail to account for the dispersion of economic activity and can lead to significant upward bias in measured border effects (Head and Mayer, 2009). Second, using geographically aggregated data causes one to overestimate within region distances relative to between region distances to a greater extent (Head and Mayer, 2009). Moreover, it has been shown that elasticity of trade with respect to distance is very non-linear and trade volumes decline rapidly in the first 300-350 kilometers and using large geographical units will fail to capture this effect (Hillberry and Hummels, 2008). Evidently both effects will push the border coefficient up as it picks up the residual effect of distance (Bemrose et al., 2017). Third, using total trade flows overlooks the fact that industries have varying trade cost elasticities (Chen and Novy, 2011) and select into geographies taking into account border related costs. Therefore estimates that employ aggregated data at the industry level suffer from compositional bias (Hillberry, 1999). Finally, Coughlin and Novy (2012) show that due to the aggregation of small geographical units into countries, larger countries (in terms of geographical extension) are associated to smaller border effects. This is because there are domestic barriers to trade such as distance. Bigger countries suffer larger domestic barriers to trade that lead to smaller domestic trade and to a smaller border effect.

Our first contribution to this literature is to use a dataset of trade flows across regions, that allows us to improve on many of these challenges. Namely, we are able to disentangle local trade (home bias) from regional trade and we can measure distances at the regional level taking into account the within-country trade frictions.

More recently, literature has moved from exploration of methodological explanations of border effects to investigating the likely causes behind it. The small but growing literature on domestic border effects (Novy (2013), Gil-Pareja et al. (2005)), or "home bias", hints at the existence of additional reasons for locality of trade. For instance, in the presence of intermediate goods trade and agglomeration forces, firms would endogenously locate to concentrate economic activity and minimize trade costs (Rossi-Hansberg (2005)). Since on average, trade in intermediate goods entail a much shorter distance, existence of such firm location effects can in part explain the observed border effects (Chen (2004), Hillberry and Hummels (2008)). Another possibility is that, borders proxy for information related barriers. This hypothesis is supported by several studies who report significant report up to 60% reduction in the border coefficient once the trade facilitating effects of social and business network are accounted for (Combes et al. (2003), Kyoji et al. (2004)). Finally, it has been shown that sharing a common judiciary can help mitigate border effects (Turrini and van Ypersele, 2010).

Understanding which factors are behind the border effect, however, requires that we give the gravity equation a causal interpretation. But, as we have pointed out before, this is difficult since borders are endogenous. Our second contribution to the literature is to move closer a causal identification of the border effect by proposing a new identification strategy that exploits our rich region-level dataset. We focus our estimation of the border effect on the regions that are located at the physical border of contiguous countries. This allows us to compare trade across two border regions that are contiguous and belong to the same country with trade across two regions that are also at the border and contiguous but that belong to different countries. Our assumption is that this specification allows us to control for variables such as culture, language, tradition and geography that are smooth over space and exploit the discontinuity created at the border.

2.2. European regional trade: a new dataset

In order to estimate whether borders deter trade across European regions, we use a new micro-dataset of shipment of goods by road to build a new region-level dataset of trade of goods across 269 European regions. In this section we explain the features of the micro-data and we give further details about how we use it to build a region-level matrix of regional trade.

2.2.1. The shipment data

We use the European Road Freight Transport dataset (ERFT), a micro-level dataset of freight road shipments collected by Eurostat. This dataset contains survey information about road shipments of goods between European regions. European countries ask national transport companies to collect data on vehicles, journeys and goods transport operations with an annual survey and transmit them to Eurostat. All survey rounds are carried out by individual countries following some pre-established guidelines. The goal of this dataset is to reflect the transport of freight by road across European countries, that represents 76.7% of all inland trade and 50% of within-EU trade.

The dataset includes data from 22 EU countries and EFTA countries (except Iceland) who report road freight transport survey data to Eurostat. The dataset covers the years 2011 to 2017.[1] Variables are collected at the vehicle-journey-good level. For each

---

[1] We exclude Estonia, Letonia, Lithuania, Luxembourg, Malta and Cyprus since they are one-region countries

shipment surveyed the dataset reports the origin and destination of the shipment (at NUTS2 level), the weight of the shipment, the distance covered in the journey and the type of good classified following the 2 digit NACE classification (Revision 2). There are 20 broad good categories that appear in the dataset, of which we focus on 9 categories that include all manufacturing.

We clean the data in the following way. First, we reduce our sample to countries for which we can also find other important variables to complement this survey. We focus on the 27 EU countries and Norway and Switzerland. These countries contain 269 regions.[2] Figure 2.1 displays the countries and regions included in our data.[3] Second, we drop the shipments of some of the industries that cannot be classified as manufacturing goods.

We use this micro-data at the industry-region level to construct industry-region quantity matrices:

$$W^{i,t} = \left[W^{it}_{nm}\right]_{N\times N}$$

where W^{it}_{nm} is the weight (in kg) from industry i shipped from n to m in year t.

2.2.2. The price estimates

Imputation of missing data: constructing region-industry prices

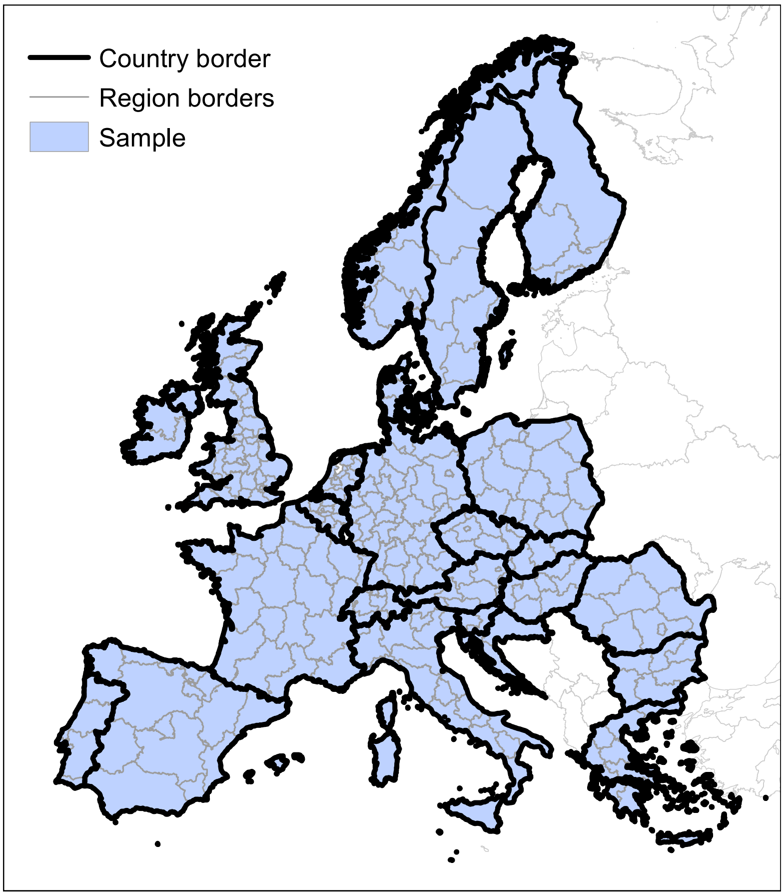
Using the European Road Freight Transport (ERFT) database we can construct the matrix of road shipments in tons between 269 European regions. In order to estimate the border effects between European regions we would like to know the value of these shipments across locations. In addition, we can use the value of the survey flows of goods to aggregate shipments at the country level and compare our data with international trade data covering all modes of transportation. Unfortunately, the value of the goods transported is not collected in the survey. We overcome this lack of data from the survey by borrowing from the imputation of missing data techniques. We leverage the fact that the statistical agencies of four European countries provide data of exports from domestic regions to foreign countries in value and volume.

---

[2] The region classification available at the level of the Survey is NUTS2.

[3] We drop the smaller islands because shipments by road are not representative of trade flows across these insular regions and other mainland regions or other insular regions located near the Eurasian continent.

Figure 2.1: Regions included in the sample



This allows us to construct a dataset of export flows from regions in France, Germany, Spain and the United Kingdom, overall 65 regions, to the countries in our sample. For these export flows we observe the value of the exports as well as the quantity, allowing us to compute the price per kilo of exports. Unfortunately, similar data could not be collected for the remaining countries in our sample. Given that the reasons to not make public the regional level data on exports are unknown to us and, hopefully, are not related to bilateral trade flows across regions, we think of our data as incomplete data. We observe the export prices for a subsample of european regions while the remaining prices are missing.

We use these regional prices to impute the missing values in our price matrix, borrowing techniques from the missing data literature. A plausible technique to deal with missing values is the conditional mean imputation, where an estimate of the mean of the conditional distribution of the variable with missing values given observables is imputed. We choose to impute values using regression imputation (Little, Rubin 1990). This procedure is recommended when the researcher is interested in estimating aggregate values.[4]

We assume that the price of exports between region *n* and region *m* in industry *i* is a log-linear function of different origin, destination and industry specific variables:

$$lnP^{it}_{nm} = \phi_n + \phi_m + \phi_i + \phi_t \tag{2.1}$$

We also include year fixed-effects since there could be macroeconomic trends affecting the level of prices. Origin-specific variables that may affect the export price can be wages or the skill level of the population. Similarly, destination-specific factors can be consumer price index or the level of competition in the destination. Finally, we add industry-level fixed effect since we expect the mean of the export price to vary by industry. We collect a large series of variables to proxy for origin and destination relevant characteristics. If the observables explain a large share of the variation of the dependent variable in the subsample with no missing values, we can use the estimated coefficients from the linear regression to impute the missing values. Table 2.A1 in the appendix contains the full list of variables included in the price regressions.

In our baseline specification we do not include any bilateral variables since we want to avoid the introduction of a bias in the estimation of the border effect. However, a potential source of differences in export prices may be distance. Since our export prices do not include transport costs we do not include distance as a control. This choice implies that all the border effects that we estimate in the data will come from the extensive margin of the weight of shipments but not through changes in the value of shipments. However, we know that countries seem to export more expensive goods to more remote locations (Hummels and Skiba, 2004). For this reason we also estimate a set of prices using the bilateral distance as an observable characteristic of the pair and we use both measures in our analysis to understand the possible bias of both approaches. Our preferred

---

[4] The main criticism against this data imputation technique has been that it assumes that no error is made in the data imputation and leads to a lower than actual variance in the imputed missing values. However, since our goal is to adjust the value of the flow of goods observed by aggregate factors such as productivity or wealth of a region, the conditional mean imputation is a seems to be a good approach.

specification is to pool all time periods and industries in the following regression:

$$lnp^{it}_{nm} = \phi^{it} + \eta' X^t_n + \pi' Z^t_m + d_{n,m} + e_{inmt},$$

where $p^{it}_{n,m}$ is the logarithm of the unit price of exports of industry i shipped from region n to region m in year t. The price of exports is calculated as the ratio between the value of exports and the weight of exports for each industry, origin and destination. This price is Free On Board (F.O.B) meaning that it does not include any transport or insurance costs that the shipment may incur in. However, recall that in the subsample of countries for which we observe region to country prices we observe exports of a region shipped by any means of transportation.

In this preferred specification we include industry-time dummies, ϕ^{it} to capture different time trends in prices across industry dummies. The vectors X^t_n and Z^t_m contain time-varying as well as time-invariant characteristics of the origin region (n) and of the destination region (m) related to economic activity (such as GDP per capita, population density, education...) and to geography (hours of sun, distance to the coast,...). The R squared in both specifications is higher than 50% meaning that, as we expected, our chosen explanatory variables capture a large part of the variation in the observed export prices.

We use this regression to build a matrix of imputed prices:

$$P^{it} = \left[P^{it}_{nm}\right]_N$$

where P^{it}_{nm} is the price of goods of industry i shipped from region n to region m in year t.

Test of the accuracy of imputed prices

In order to assess the accuracy of our imputed prices we perform two sets of checks. First, we perform a series of out-of-sample estimations where we drop each of the four countries at a time for which we observe regional export prices and we predict export prices. We then compare our out-of-sample estimates with the actual regional prices. Figure 2.A2 shows the fit between the actual regional prices and the out-of-sample predicted prices.

Second, we check the accuracy of our predicted regional prices by comparing them to actual country-level export prices. We collect export value and weights (kilos) from Eurostat for all European countries and compute unit export prices for every industry and country pair. We aggregate our region-pair estimated prices to the country-level and we compare them to the country-level export data.[5] Figure 2.A3 in the appendix shows the fit between the actual regional prices and the predicted prices. Both tests seem to suggest that our imputed prices are reasonable.

2.2.3. A first look at the data

For each industry i and year t, we construct the trade value data as follows:

$$V^{it} = \left[V^{it}_{nm}\right]_{N\times N} \qquad \text{where } V^{it}_{nm} = P^{it}_{nm} \cdot W^{it}_{nm}$$

This is the matrix of bilateral trade for industry i, in year t.

Comparison with Aggregate trade data

Before exploring the newly built regional dataset, we check that this data is consistent with country trade data from standard sources. Our sample contains $k = 1, ..., K$ countries. We construct country data as follows:

$$V^{it} = \left[V^{it}_{ks}\right]_{K\times K} \qquad \text{where } V^{it}_{ks} = \sum\nolimits_{n\in k} \sum\nolimits_{m\in s} V^{it}_{nm}$$

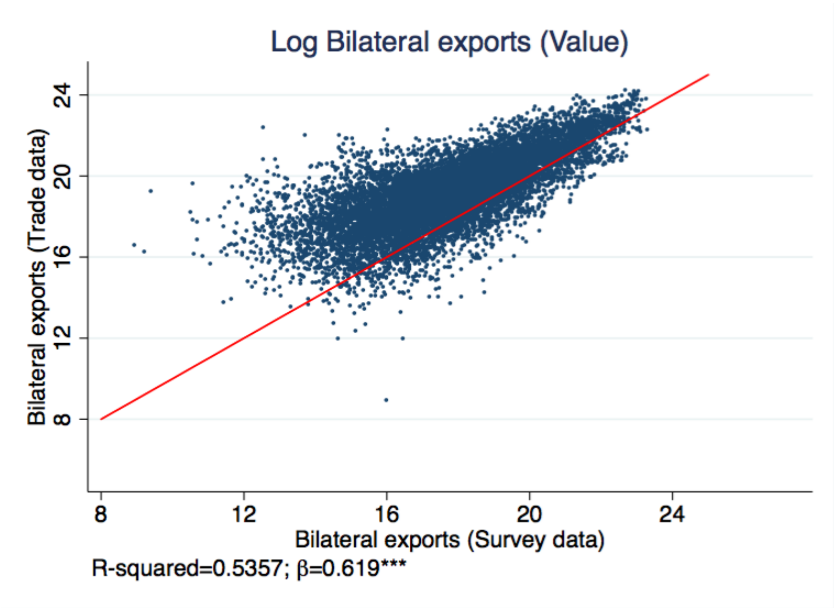
Thus, V^{it}_{ks} is the value of trade in industry i between countries k and s, in value and where i indicates one of the 9 aggregated manufacturing industries collected in the survey.

Let us compare this matrix with the matrix of bilateral trade among the countries in our sample, as obtained from standard sources (Eurostat). Figure 2.2 plots the country to country bilateral trade flows by industry over the bilateral country-level flows in our new dataset. Panel A displays the correlation using the value of shipments while panel B displays the correlation using the weight. Table 2.A2 in the Appendix reports the summary statistics of these two datasets. As we can see, the correlation between both measures is very high (0.72 for value and 0.77 for quantity) confirming that newly constructed data explains a large part of the country-level international trade flows.

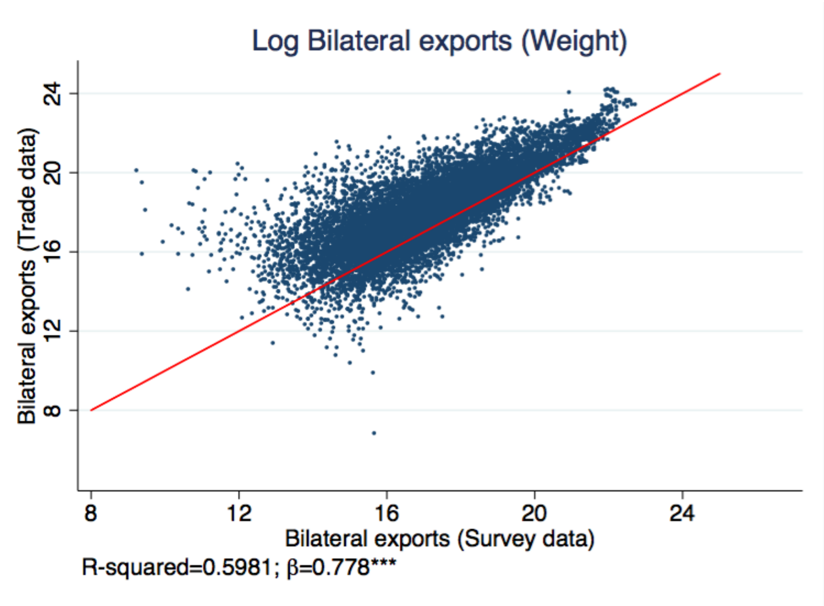
---

[5] We aggregate regional price by weighting by the GDP of each region but the simple un-weighted aggregate performs very similarly

Figure 2.2: Correlation with aggregate international trade data



A) Value



B) Weight

Notes: Bilateral exports by industry across all 27 countries in the sample. Aggregate trade data from IMF statistics.

However, the aggregate trade statistics report larger average values and quantities than our constructed regional flows dataset, pointing to a possible downward bias towards slightly cheaper and smaller trade flows in the shipments by road. Figure 2.2 shows that the correlation between aggregate international trade shipments and road shipments is very high for large values and quantities but slightly weaker for the lowest values and quantities. This means that shipments by road may be smaller in value and quantities than shipments using other means of transportation.

We have shown that both the imputed prices provide a sensible measure of the value of exports and that our data, in aggregate terms, explains a very large part of aggregate international data. Now we proceed to explore the newly built regional-level trade dataset.

European regional trade: Summary statistics

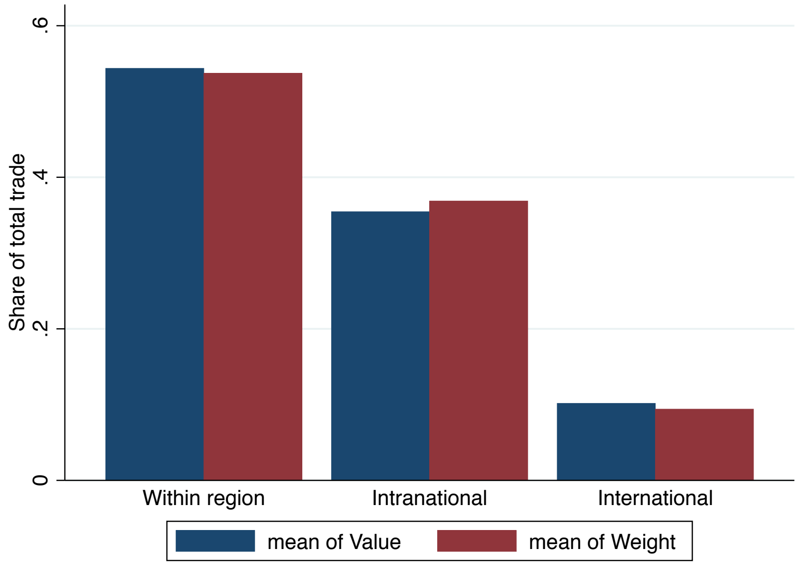
Our unit of observation will be the bilateral flows between region n and region m in industry i and year t, defined previously as:

$$V_{nm}^{it} = P_{nm}^{it} \cdot W_{nm}^{it}$$

Our dataset contains three types of region-pairs: First, the pair composed by one region and itself, that denotes the flow of goods within the region. We will refer to this group as *within region* trade. Second, the pair composed by two regions, n and m, that belong to the same country. We will call this group *Intranational* trade. Finally, we find a third type of region-pair formed by two regions, when they belong to different countries. The flow of goods that belong to this type of pairs will be denoted as *International* trade.

The first feature of our data that we will explore is how relevant are each of these group in overall value. Figure 2.3 shows the share of value and weight of each of the groups on the overall value and weight in the dataset. The most important type of flows in our dataset are *within region* flows of goods that account for more than 50% of the total value and total weight. Compared to this type of transactions, domestic flows account for slightly less than 40% while international flows account for around 10% of the total value and weight. Notice that, as we expected, the share of value of international good flows is higher than the share in weight. This figure shows that shipments by road are more common for shorter distances. This will be taken into account when we define our empirical strategy to identify the border effect. However, notice that, even if shipments by road could be over-estimating the importance of *within region* trade, our strategy will be focused on comparing *Intranational* trade with *International* trade.

Figure 2.3: Composition of trade flows



Next, we compare the summary statistics of these three groups of trade flows. It is important to notice that there are several pairs in our dataset, built from the European Road Freight Transport Survey, for which we do not observe positive trade flows. Specifically, only 35% of all the possible industry-region-pairs perform some transaction of goods in our data. Conditional on observing a positive flow of goods, Table 2.1 reports the most relevant summary statistics of the trade flows in our dataset.

Table 2.1: Summary Statistics, conditional on positive trade flows

	Own Region Flows		Inter-regional Domestic		International	
	Mean	Sd.	Mean	Sd.	Mean	Sd.
Value (Mill)	22631.27	31058.16	895.67	1745.29	54.10	157.14
Quantity (Mill kg)	10859.09	11217.99	452.31	731.13	24.29	65.72
Distance (km)	66.49	32.14	373.98	216.35	1032.16	570.86
Observations	1883		31013		147269	

As we can see, *within region* bilateral flows are more expensive and larger, and cover much smaller distances (on average 64 km). There are smaller differences between *Intranational* trade and *International* trade. However, domestic trade flows are also larger and more expensive than international bilateral flows and happen over a distance of 373 km on average, compared to international flows that happen over a distance of 1000 km on average. In the next section we explain our identification strategy and how we deal with these initial differences in order to estimate the border effect in a reliable way.

2.3. Measuring the Border Effect

2.3.1. Estimation Framework

Our goal in this section is to estimate the average border effect across industries and country-pairs. The literature in trade and international economics has used for more than fifty years the gravity equation framework as the main tool to study the determinants of bilateral trade (Head and Mayer (2014)). We borrow our empirical framework from this literature and apply it to understand the trade-related costs of borders following previous studies (McCallum (1995), Anderson and Van Wincoop (2003)). A wide variety of theoretical trade models give rise to a gravity equation structure as shown by Allen et al. (2015), allowing us to remain agnostic about the micro-foundations that determine trade flows. According to this framework, the value of trade flows between two regions n and m in industry i and time t is a function of the existence of a border between these two regions, a set of bilateral trade costs and origin and destination specific factors. Our estimating equation is the following:

$$logV^{it}_{nm} = \phi_t + \phi_i + \phi_n + \phi_m + \beta C_{nm} + \gamma h_{nm} + \tau' X + u_{nmkt}, \tag{2.2}$$

where n,m denote region of origin and region of destination, k denotes industry and t denotes time. The dependent variable, $logV^{it}_{nm}$, denotes the logarithm of trade flows, that we will study both in values and volumes. The dummy $C_{n,m}$ captures the existence of an international border between regions n and m:

$$C_{n,m} = \begin{cases} 1 & \text{if country}_n \neq \text{country}_m \\ 0, & \text{otherwise} \end{cases} \tag{2.3}$$

The dummy $h_{n,m}$ captures whether the origin region is the same as the destination region, this is, whether the trade flow happens *within* the region. We use $h_{n,m}$ as a proxy for home bias, and we define it as:

$$h_{n,m} = \begin{cases} 1 & \text{if } n = m \\ 0, & \text{otherwise} \end{cases} \tag{2.4}$$

The vector X contains other determinants of trade costs such as geographic distance, contiguity and common language. We add time fixed effects, α_t, industry fixed effects α_i, and origin and destination fixed effects, ϕ_m and ϕ_n, respectively. Finally $u^{it}_{nm} \equiv v^{it}_{nm} - Ev^{it}_{nm}$ is a mean zero residual.

The fixed effects ϕ_n and ϕ_m capture factors that make origin n export more to all destinations and, similarly, factors that make destination m to import more from all origins. It is well known that the main such factor is economic size. Indeed, it is common to use measures of production and spending as regressors instead of these fixed effects. But the fixed-effects specification is more flexible. It allows us, for instance, to capture systematic tendencies of an origin to export more than what its economic size would suggest due to being located in a strategic position in trade routes or having a natural resource such as oil. Finally, as shown by Anderson and Van Wincoop (2003), the bilateral trade between two economic units will also be affected by the relative remoteness of these units, that is believed to be captured by the degree of competition that an average exporter would face when trading with a given destination. The introduction of exporter and importer fixed effects controls also for these factors, referred to as multilateral resistance terms. The standard approach in the literature is to include region-industry fixed effects to allow these resistance terms to vary across industries. However, due to computational constraints, we decide to include the smaller set of fixed effects described above. In any case, we are confident that the introduction of region, rather than country level, fixed effects will be rich enough to capture these differences in trading patterns that are region-specific.

The empirical challenge that we face in order to estimate the effect of an international border on bilateral trade is that borders are not allocated exogenously. For example, two regions that are two thousand kilometres away from each other are more likely to belong to different countries than two regions separated by two hundred kilometres. Thus, it is necessary to compare pairs of regions conditional on some important determinants of trade costs in addition to borders.

We follow the literature and add different controls to condition on some factors that would make us overestimate the border effect. Our vector of controls, X, includes three variables. First, we control for transport costs as proxied by the log of the geographic distance between every pair of regions.[6] Second, we add a dummy to capture whether these regions are contiguous to each other. The introduction of this control does not come from a theoretical foundation but is introduced to capture non-linearities in distance. Finally, we control for the existence of a common language between the two regions by adding a common language dummy. The underlying assumption of this specification is that we can think of borders as quasi-random once we control for transport costs, as proxied by distance and contiguity, and cultural differences, as proxied by language.

So far we have described the standard gravity framework, that has been used to estimate the effects of different bilateral variables on trade. Our focus in this paper is to examine the role of governments in facilitating or deterring trade. Most theories suggest that sharing a government reduces bilateral trade costs and fosters bilateral trade, but the identification of the border effect is not trivial. Since measuring this effect is the main focus of our study, it is important to be very explicit about our approach and how it differs from existing studies.

We note first that, unlike us, most gravity studies use country data and, thus, origin/destination pairs do not share the same national government by definition. Some origin/destination pairs share however some of the supra-national governments that have proliferated around the world since the end of World War II. These include the European Union, the Eurozone and many other regional trade and currency agreements, as well as international organizations such as the WTO. The goal of many of these supra-national

---

[6] We use the geodesic distance between the centre point of the origin and the destination region. For intra-region shipments, we follow Nitsch et al 2000, and compute internal distances using a function of the country's surface

governments is to reduce trade costs among their members. Thus, gravity studies that use country data typically include dummy variables indicating whether origin/destination pairs share these supra-national levels of government or not. The results tend to suggest that trade between regions that share these supra-national governments is larger than among origin/destination pairs that do not share them.

However, it is unclear whether we can interpret these findings as evidence that sharing supra-national governments *reduce* bilateral trade costs and *cause* bilateral trade. To reach this conclusion, we must assume that, once we condition on the geographical and cultural distance of an origin/destination pair, the sharing of supra-national governments is randomly assigned. This does not seem too outlandish, as most supra-national governments are either shared by the vast majority of countries (for instance, the WTO or the UN) or have a regional nature (for instance, the European Union or most trade and currency agreements). This is why gravity studies often interpret the estimated coefficients on these dummy variables as the causal effect on trade of sharing the corresponding supra-national government.

Our study has the distinctive advantage of using regional data instead of country data. This allows us to measure not only the effects of sharing supra-national governments, but also the effects of sharing national governments. To do this, we compare trade between regions in the same country with trade between regions in different countries, holding constant geographical and cultural distances. That is, we can introduce the dummy variable C_{nm} that takes value one if an origin/destination pair are in the same country, and zero otherwise.

This immediately raises a question. Should we also add to our specification dummy variables indicating whether an origin/destination pair share a supra-national government such as the European Union, the Eurozone or the Schengen treaty? We think the answer is negative, as this would be a typical case of a bad control. Participation in supra-national governments is negotiated by national governments and, as a result, sharing a supra-national government is a function of whether an origin/destination pair shares a national government. This means that a comparison of bilateral trade conditional on whether origin/destination pairs share supra-national governments would not have a causal interpretation. The difference in bilateral trade between origin/destination pairs that share or do not share a national government conditional on sharing a supra-national government equals the causal effect of sharing a national government on bilateral trade, plus a selection bias term that reflects the fact that sharing a national government changes the composition of the pool of origin/destination pairs that share a supra-national government. This is why we drop from our baseline estimation all dummies related to supra-national governments.

Under the assumptions discussed above, the coefficient β captures the difference between the average trade flow of two regions divided by an international border and the average trade flow between two regions that are within the same border (country), controlling for transport costs and language. Since our estimation compares pairs of regions controlling for distance, contiguity and language, we interpret β as as the effects on trade of not sharing a national government. That is, we interpret β as the "border effect".

Finally, the coefficient on the home bias dummy, γ, will capture how much more does a region trade with itself compared to how much it trades with the average region. The introduction of the home bias dummy, h_{nm}, allows us to separate between the effect of trading $outside$ of your region (h_{nm}="home bias") and the effect of trading $outside$ of your country (C_{nm}="border effect").

We follow two alternative estimation methods based on the type of data we are studying. As is common in trade data, we observe a large number of zero trade flows in our sample. This prevalence of zeros makes it reasonable to run our main regression in a way that allows us to account for the zero trade flows. We estimate our baseline specification using Poisson Pseudo-maximum likelihood (Santos and Tenreyro) that allows the dependent variable to take value zero and that is considered more reliable when the number of zeros is very prevalent (Head and Mayer).[7] For comparison with the previous literature, we also estimate the border effect in the sub-sample of region-pairs with positive trade flows. We use Poisson Pseudo-maximum likelihood (PPML) as well as Ordinary Least Squares for the estimation in the sub-sample with no zeros.

2.3.2. Results

Table 2.2 reports the results from the least squares and poisson pseudo-maximum likelihood estimation. Columns 1 and 2 display the results for the positive-trade flows

---

[7] Helpman, Melitz and Rubinstein propose an alternative way to account for the zero trade flows in a two-step estimation procedure. However, their approach requires firm level data that is unavailable to us at the region level

sample and column 3 displays the results for the full sample.

Table 2.2: Border Effect - Baseline results

Outcome: Bilateral Trade (Value)	Positive flows		Full sample
	OLS	PPML	PPML
	(1)	(2)	(2)
Border effect	-1.007***	-1.073***	-1.408***
	(0.082)	(0.173)	(0.170)
log(Dist)	-0.619***	-0.694***	-1.469***
	(0.061)	(0.076)	(0.084)
Contiguity	0.643***	0.692***	0.238**
	(0.092)	(0.128)	(0.112)
Language	0.072	0.007	0.533***
	(0.074)	(0.098)	(0.139)
Home Bias	1.536***	1.703***	1.214***
	(0.122)	(0.221)	(0.217)
Constant	21.080***	20.621***	26.635***
	(0.226)	(0.505)	(0.692)
Origin FE	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Obs.	5.37e+05	5.37e+05	4.52e+06
R-Squared	0.583	0.618	0.613

Notes: Standard errors are in parentheses clustered at the country-pair level.* significant at 10%, ** significant at 5%, *** significant at 1%.

The border coefficient is negative and highly significant in all the specifications and across subsamples. As the literature has pointed out, the ordinary least squares results are very similar to the Poisson pseudo-maximum likelihood when we only consider positive trade flows (Columns 1 and 2), while the results are different when we use the complete sample. The coefficient on international border, β, is -1.073 in the sub-sample conditional on observing positive trade and -1.408 in the full sample. Thus, it seems that the existence of a border may deter trade both in the extensive and the intensive margins. In our sample 90% of the industry-year-region pairs feature no trade so we believe it is important to use this information rather than focusing on the positive trade flows. The positive-trade subsample is included in the regressions for comparison. Our results suggests that two regions that are separated by an international border trade 4 times less ($exp(1.408) = 4$) with each other than the average pair of domestic regions. Conditional on observing positive trade across two regions, these two regions still trade 2.9 times less with each other, compared to the average domestic region-pair, if they are divided by an international border.

The coefficient on the Home Border, γ, is positive and highly significant. According to our estimation, a pair of domestic regions trades 3.36 ($=exp(1.214)=3.36$) times less with each other than with themselves. Our dataset provides us with the opportunity to separate the home bias and the international border effect. We find that the border effect is slightly larger but of a similar magnitude as the home bias.

The inclusion of the controls seem to be relevant too. A 10% increase in bilateral distance reduces trade by 14.7% and contiguous regions trade 1.8 times more on average than non-contiguous regions. Finally, sharing a language increases trade: regions that have a common language trade 1.7 times more than regions that do not share a language. As we can see, sharing a language is about 1/3 of the magnitude of the international border effect.

We find a coefficient for the average border effect in Europe of -1.408 (using all observations) and of -1.073 (using only positive trade flows.) What do our results imply in terms of tariff equivalence? Under the assumption of a CES demand system, the coefficient on International border β would be the combination of the trade elasticity to trade costs $(1 - \sigma)$ and the border effect ($\beta = (1 - \sigma) *$ Border effect). This assumption allows us to transform our estimated coefficient into tariff-equivalent costs, as a function of the trade elasticity. Table 2.3 summarises our findings.

Table 2.3: Average border effect in Europe

Estimates			Tariff-equivalent effect		
Effect	Coefficient	Trade red.	$\sigma = 3$	$\sigma = 5$	$\sigma = 7$
Border Effect (all trade flows)	-1.408	4.08	59.9%	32.5%	22.3%
Border Effect (positive trade)	-1.073	2.92	43.0%	23.9%	16.6%

We conclude that regions separated by a border trade 4 times less between each other than the average pair of domestic regions. An international border is therefore estimated to be equivalent to a tariff of around 32.5% (assuming a trade elasticity of 5).

Once we have established our main results, we present a set of robustness checks to proof the stability of our findings.

2.3.3. Robustness checks

In this subsection we show that our results are robust to several tests regarding different potential limitations of our baseline estimation. First, we test that our findings are robust to including additional controls related to international trade agreements. Second, we confirm that our estimates are stable to using different methods of price imputation. Finally, we show that our results hold after we control for a different distance measure and when we address the existence of alternative means of transport.

Additional Controls

So far we have included as controls variables related to the geography and the the language of regions. However, it is common in the gravity equation literature to include other determinants of trade costs such as the membership of a trade agreement.

Unfortunately, only 2 of our 24 countries do not to the European Union, Switzerland and Norway, leaving us with very little variation across our sample. However, we can control for two additional international agreements that may be relevant for bilateral trade flows. The first one is the Schengen agreement, that abolished checks at the country borders between members of the agreement and has been effective since 1995. Belonging to this agreement would reduce the barrier for the movement of people, making the international border more comparable to a domestic border. 19 of the countries in our sample belong to this agreement. The second agreement is the Euro Zone, by which a monetary union was created between 19 of the European Union members (13 of them in our sample). Sharing a currency has been pointed out by the literature as an important trading since it allows trade exchanges to be independent of exchange rate variation. The list of countries that belong to the different agreements can be found in Tables 2.A3 and 2.A4 in the Appendix. Table 2.A5 in the appendix displays the results of our baseline specification after including two new controls: a dummy to capture membership of the Schengen agreement and a dummy to capture membership of the Euro Area.

The border coefficient is reduced to -1.287 (column 3) which is a bit smaller than the results in the baseline specification but the difference is not striking and the confidence intervals of both coefficients overlap. The rest of the coefficients are almost equal to the coefficients we estimated in table 2.2.

Price imputation method

We now turn our attention to the possible caveats of our price estimation method. To show that our results do not depend on how we value the trade flows that we observe in the road freight survey, we repeat our baseline estimation changing the price imputation method.

First, we show our results using the trade flows in tons, without making any price adjustment. Table 2.A6 in the appendix displays the results when the dependent variable is the quantity shipped (in kilos) rather than the value of the shipments. The results are very similar.

Second, we try to improve our price adjustment by considering adjusting the export price by the bilateral distance between regions. Our current price estimation considers only origin and destination characteristics. However, one important bilateral variable that has been shown to affect the price of exports is distance (Hummels, 2004)[8]. Not adjusting the quantities traded over long distances by a higher than average price could make us underestimate the border effect. Table 2.A7 in the appendix displays the results of using distance as a determinant of prices. As expected, the effect of distance as a barrier to trade is reduced considerably, but the border effect estimated remains very stable.

Tables 2.A6 and 2.A7 show that all the coefficients are very stable to the change in the price adjustment method. These results suggest that our estimate of the border effect is not driven in any way by the price imputation we perform to transform volumes of trade into values. Our results are also robust to including distance as a determinant of prices, showing that our initial price imputation method was not causing an overestimation of the border effect.

---

[8] Alchian and Allen (1964) were the first to develop this theory that countries trade more expensive goods with distant trade partners.

Distance measure and modes of transportation

Finally, we run a series of robustness checks to confirm that our results are not driven by our use of a survey of shipments by road. First, we want to mention that road freight is the main model of transport of inland trade in the European Union. 76.7% of inland transport is made by roads compared to 17.3% by railway and 6% by inland waterways. Thus, it is clear that we are capturing a very large share of inland trade. However, if we look at total intra-EU trade, transport by road represents 51.5% while maritime transport represents a 32.4% and railway transport an 11.6%. Given this information, we try to address the possible biases that the use of road freight data could introduce in our estimates.

First, we address the fact that we only observe shipments by road. Therefore, pairs of regions that are very well connected by sea could use always the maritime transport mode and showing up in our data as trading less than the actual total trade by all means of transportation. Second, a shipment in our data could be part of a longer shipment chain. For example, our data could include shipments that come from China, arrive at the port of Rotterdam and are then distributed by road to different European regions. These distributional chains could bias the representation of the shipments in the survey. To control for this we repeat our main estimation adding controls for the existence of large ports. We introduce two dummies. The first dummy, “Origin & Destination top Port”, captures whether both the origin and the destination have a port that is within the top 20 ports in Europe. The second dummy, “Both contiguous to top Port”, captures whether both regions in the region-pair are contiguous to a region with a top port.

Table 2.4 reports the results from the baseline specification in column 1 and the results after the introduction of the controls in column 2. The coefficient on the international border, β, -1.393, is remarkably similar to the coefficient we obtained in our preferred specification. The coefficient of the Home Bias is also very stable. Note that the Port controls we introduced are significant. Two regions that have a top port do trade less with each other than the average pair of regions, confirming our hypothesis that shipments by road and shipments by sea may be substitutes. The second port dummy, that captures the connection to large ports, is mildly significant and positive. These results support our second hypothesis, that road shipments may be part of a longer chain of distribution using different transport modes.

Table 2.4: Average Border Effect - Mode of Transport Controls

Outcome: Bilateral Trade (Value)	Port controls		
	Full Sample (1)	Port Controls (2)	No Islands + Greece (3)
Border effect	-1.408***	-1.393***	-1.331***
	(0.170)	(0.162)	(0.175)
Origin & Dest top Port		-0.828***	
		(0.168)	
Both contiguous to top Port		0.213*	
		(0.110)	
log(Dist)	-1.469***	-1.496***	-1.390***
	(0.084)	(0.077)	(0.085)
Contiguity	0.238**	0.213**	0.351***
	(0.112)	(0.105)	(0.100)
Language	0.533***	0.530***	0.431***
	(0.139)	(0.132)	(0.139)
Home bias	1.214***	1.251***	1.357***
	(0.217)	(0.192)	(0.203)
Origin FE	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Obs.	4524912	4524912	2939328
R-Squared	0.613	0.616	0.627

Notes: Standard errors are in parentheses clustered at the country-pair level.* significant at 10%, ** significant at 5%, *** significant at 1%.

Second, we address the fact that the use of road shipment data may be specially biased when we consider countries that are islands, or that contain islands. To check that our results are not driven by these countries, we repeat our baseline specification dropping the United Kingdom, Ireland and Greece. Since these regions are islands or have a large part of coastal or insular territory, we are concerned that their trade flows may happen mostly by ship, and if these regions are remote, this may introduce a bias in the border coefficient. The results are reported in column 3 of table 2.4. Our estimate of the border coefficient is now -1.331, smaller than the baseline coefficient, but the confidence intervals of both coefficients overlap. This confirms the robustness of our results.

2.3.4. Why does regional data matter?

In this section, we compare our results to the literature on border effect that has mostly focused on country-level trade. Our goal is to find out how aggregating our data to the country-level could bias the border effect estimation.

Remember that we find an average border effect that is smaller than the one estimated in the literature: we find a coefficient of -1.408 versus -1.65 for US and Canada (Anderson and Van Wincoop (2003)), -2.75 for EU members (Head and Mayer, 2000).[9] There are two important differences between our dataset and the standard approach used in previous studies (i.e. Chen (2004), Wei (1996)) that can lead to differences in the estimates. First, we do not have to proxy for domestic trade using total GDP minus exports. This could present a problem if a lot of the production in a country is locally consumed rather than locally traded, overestimating the border effect. On the contrary, in our data we observe domestic shipments across different regions as well as within region shipments (*Own Region* trade flows). Second, these studies normally have to estimate the transport costs of domestic trade. The literature has struggled to find a convincing measure of domestic distance but the best possible alternative would be to create a weighted average of domestic distances, where the weights are given by the observed trade flows. In general this cannot be done. However, since we have regional good flows we can measure domestic distances region to region, overcoming part of this aggregation bias.[10]

We use our data to create a country-level dataset in order to compare our results with the findings in the literature. Our methodology is the following: first, we aggregate all domestic flows at the country level including shipments within region and shipments across different regions. This is the traditional approach that considers all non-exported output as domestic trade. Second, we use average bilateral distance across all pairs of

---

[9] Chen (2004) finds a slightly lower border effect of -1.32 but her work considers the six largest countries in the EU. In addition she treats the zero trade flows differently, by replacing the $log(X_{ij} = log(1 + X_{ij})$ for zero trade flows and running a Tobiit model.

[10] Some common approaches are to compute internal distances between different large domestic cities or use a measure of internal distance that is a function of the area of the country (Nitsch, 2000).

regions to compute the average distance across countries and we follow the literature in order to compute the internal distance within a country using the area-based measure described before (Nitsch, 2000). Thus, we are introducing two biases compared to our baseline approach, one coming from the composition of trade flows and the other from the aggregation of distance.

We run our baseline regression using the aggregated dataset, formed by country-country pairs:

$$logV_{nm}^{it} = \phi_t + \phi_k + \phi_n + \phi_m + \beta C_{nm} + \gamma h_{nm} + \tau' X + u_{nmkt}, \tag{2.5}$$

where n,m denote country of origin and country of destination, i denotes industry and t denotes time. The dummy $C_{n,m}$ captures the existence of an international border

$$C_{n,m} = \begin{cases} 1 & \text{if country}_n \neq \text{country}_m \\ 0, & \text{otherwise} \end{cases} \tag{2.6}$$

The international border is now identified from comparing the average cross-country trade flow with the average within country trade flows. Since we do not have intra-country trade we cannot estimate the Home Bias effect. Table 2.5 shows that the aggregated data yields a larger estimate of the border effect.

These results show that if we had not been able to observe regional trade, we would have estimated a border effect of -2.938 (according to the OLS estimate) and of -3.66 (according to the PPML estimate). These estimates are larger than the estimated effect in our baseline estimation that is reproduced in column 3. On the contrary, we would be under-estimating the effect of distance and we would over estimate the effect of Contiguity, that now captures contiguity at the country level.

This methodology yields coefficients that are at least 200% larger when we aggregate the data at the country level compared using the same data at the regional level, highlighting the importance of using sub-country level data.

Table 2.5: Average Border Effect: Country-level

Outcome: Bilateral Trade (Value)	Country level data		Region level data
	(OLS) (1)	PPML (2)	PPML (3)
Border effect	-2.938***	-3.669***	-1.408***
	(0.189)	(0.202)	(0.170)
log(Dist)	-1.152***	0.071	-1.469***
	(0.086)	(0.226)	(0.084)
Contiguity	1.038***	1.584***	0.238***
	(0.124)	(0.288)	(0.112)
Language	0.497***	0.425**	0.533**
	(0.101)	(0.193)	(0.139)
Home Bias			1.214***
			(0.217)
Origin FE	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Obs.	22724	22724	4524912
R-squared	0.775	0.851	0.613

Notes: Standard errors clustered at the pair level are in parentheses.* significant at 10%, ** significant at 5%, *** significant at 1%. Columns (1) and (2) use our regional data aggregated at the country-level and column (3) displays the baseline specification, regional level aggregation.

2.4. Heterogeneous Border effects

2.4.1. Border effect across industries

In this section, we explore the heterogeneity in the border effect. We first exploit the industry-variation in our data. We observe shipments in nine different manufacturing industries: Food-Beverage-Tobacco, Textiles, Wood, Chemicals, Non-Metal, Metal, Machinery, Vehicles and other manufactured goods. We run nice different regressions, industry-by-industry:

$$logV_{nm}^{it} = \phi_t + \phi_n + \phi_m + \beta C_{nm} + \gamma h_{nm} + \tau' X + u_{nmt}, \tag{2.7}$$

where n,m denote country of origin and country of destination, k denotes industry and t denotes time. The dummy $C_{n,m}$ captures the existence of an international border. By running industry-specific regressions we do not restrict the border effect to be the same across industries.

Tables 2.6 and 2.7 report the results of this estimation using value as the dependent variable while the results for quantities as the dependent variable are reported in Table 2.A8 and 2.A9 in the Appendix. We find that there is variation in the border coefficients both in the intensive (Table 2.6) and in the extensive margin (Table 2.7). The results of both estimations are generally in line, only for a couple of industries the estimates of the OLS are significantly different than from the Poisson pseudo-maximum likelihood. Considering the full sample, the border effect ranges between -0.986 (Chemicals) and -1.837 (Food, Beverage and Tobacco). It seems that the border is higher for final/consumer goods (Food, Beverage, Tobacco; Textiles and Other manufactured goods) and lower for intermediate goods (Chemicals, Wood and Metal). The exception is Vehicles, that have a relatively low border effect as well.

Discussion of results

To understand what is the tariff-equivalent cost of the border effect across industries, we need to make some assumptions. As noted in the previous subsections, the coefficient on international border is composed of the border friction and the trade elasticity. In the case of industry-level regression, we estimate a industry-specific border coefficient. For a given industry k, the coefficient of the border is: $\beta_k = (1 - \sigma_k) \times Border$.

One potential approach is to assume that the trade elasticity is common across industries. Assuming a trade elasticity of $(1 - \sigma) = -4$, as is common in the literature, Table 2.8 displays the corresponding tariff-equivalent cost. The tariff-equivalent cost of the border effect ranges from 21.8% (for Chemicals) to 44.4% (for Food, Beverage and Tobacco). The range is of variation is 20%, which is as large as two thirds of the average estimated border effect. We believe this heterogeneity should be explored further in future research.

Table 2.6: Border Effect in Europe by Industry - OLS

Outcome: Trade (Value)	Food/Bev	Textiles	Wood	Chem	Non-Met	Metal	Machinery	Vehicles	Other
Border effect	-1.634***	-0.619***	-1.081***	-0.845***	-1.166***	-0.924***	-0.988***	-0.886***	-0.964***
	(0.096)	(0.102)	(0.108)	(0.090)	(0.106)	(0.083)	(0.107)	(0.096)	(0.136)
log(Dist)	-0.827***	-0.299***	-0.793***	-0.838***	-0.951***	-0.673***	-0.408***	-0.434***	-0.382***
	(0.074)	(0.052)	(0.075)	(0.063)	(0.116)	(0.069)	(0.062)	(0.073)	(0.047)
Contiguity	0.840***	0.296***	0.709***	0.496***	1.097***	0.635***	0.778***	0.493***	0.489***
	(0.073)	(0.083)	(0.097)	(0.109)	(0.116)	(0.095)	(0.119)	(0.104)	(0.076)
Language	0.194***	-0.045	-0.068	0.187**	0.016	0.130*	0.143*	0.050	-0.144
	(0.071)	(0.099)	(0.106)	(0.085)	(0.087)	(0.078)	(0.079)	(0.107)	(0.114)
Home Bias	1.202***	1.335***	1.346***	1.375***	2.248***	1.537***	2.133***	1.707***	1.354***
	(0.115)	(0.181)	(0.130)	(0.129)	(0.190)	(0.113)	(0.192)	(0.146)	(0.104)
Constant	21.043***	16.730***	20.546***	20.200***	22.933***	20.446***	21.083***	23.225	17.644***
	(0.602)	(0.299)	(0.763)	(0.266)	(0.540)	(0.461)	(0.721)	.	(0.242)
Origin FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	88957.000	26283.000	66336.000	73401.000	51555.000	68844.000	61223.000	56983.000	43565.000
R-Squared	0.712	0.510	0.631	0.602	0.722	0.613	0.592	0.551	0.501

Notes: Standard errors are in parentheses clustered at the country-pair level.* significant at 10%, ** significant at 5%, *** significant at 1%.

Table 2.7: Border Effect in Europe by Industry - PPML (ALL TRADE FLOWS)

Outcome: Trade (Value)	Food/Bev	Textiles	Wood	Chem	Non-Met	Metal	Machinery	Vehicles	Other
Border effect	-1.837***	-1.594***	-1.227***	-0.986***	-1.327***	-1.268***	-1.361***	-1.166***	-1.568***
	(0.193)	(0.200)	(0.124)	(0.109)	(0.252)	(0.128)	(0.245)	(0.187)	(0.142)
log(Dist)	-1.585***	-1.405***	-1.717***	-1.729***	-1.914***	-1.580***	-1.234***	-1.292***	-1.349***
	(0.064)	(0.093)	(0.090)	(0.088)	(0.075)	(0.082)	(0.108)	(0.098)	(0.078)
Contiguity	0.130	-0.006	-0.017	-0.014	0.592***	0.037	0.388***	0.101	-0.018
	(0.100)	(0.125)	(0.096)	(0.119)	(0.122)	(0.104)	(0.143)	(0.107)	(0.070)
Language	0.639***	0.401**	0.605***	0.446***	0.824***	0.561***	0.761***	0.382**	0.416**
	(0.185)	(0.173)	(0.136)	(0.168)	(0.162)	(0.149)	(0.164)	(0.191)	(0.181)
Home Bias	0.571***	0.717***	0.647***	0.684***	1.696***	0.688***	1.508***	1.121***	0.664***
	(0.131)	(0.224)	(0.092)	(0.135)	(0.286)	(0.133)	(0.314)	(0.174)	(0.125)
Constant	27.693***	22.828***	26.295***	28.158***	30.010***	27.612***	27.542***	26.061***	21.876***
	(0.460)	(0.564)	(0.669)	(0.596)	(0.591)	(0.608)	(0.767)	(0.837)	(0.586)
Origin FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	5.03e+05	4.84e+05	5.03e+05	5.03e+05	5.03e+05	4.99e+05	4.95e+05	4.97e+05	4.88e+05
R-Squared	0.909	0.371	0.844	0.890	0.939	0.858	0.852	0.842	0.721

Notes: Standard errors are in parentheses clustered at the country-pair level.* significant at 10%, ** significant at 5%, *** significant at 1%.

Table 2.8: Tariff-equivalent effect of border across Industries

	Food, B, T	Text.	Wood	Chem.	Non-metal	Metals	Machin.	Vehicles	Other
Assumption	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
σ=-4	44.4	35.7	27.8	21.8	30.4	28.9	31.3	26.3	36.8
Broda & Weinstein (2006)	33.9	28.8	21.5	17	23.5	22.3	24.2	20.4	28.3

The assumption that the trade elasticity is constant across industries seems like a strong assumption. A second approach is to allow for the trade elasticity to change across industries. To improve our estimates, we use the elasticities we aggregate the elasticities in Broda and Weinstein (2006) at the level of our nine industries and use industry-specific elasticities to compute the tariff equivalent. The Broda and Weinstein (2006) elasticities are lower than 5, on average, but the ranking of industries is almost unchanged.

It seems that the border is higher for final/consumer goods (Food, Beverage, Tobacco=33.9%; Textiles =28.8% and Other manufactured goods= 36.8%). The border effect is lower for intermediate goods (Chemicals=17%, Wood=21.5% and Metals=22.3%). The exception is still Vehicles, for which we estimate a border effect of 20.4%.

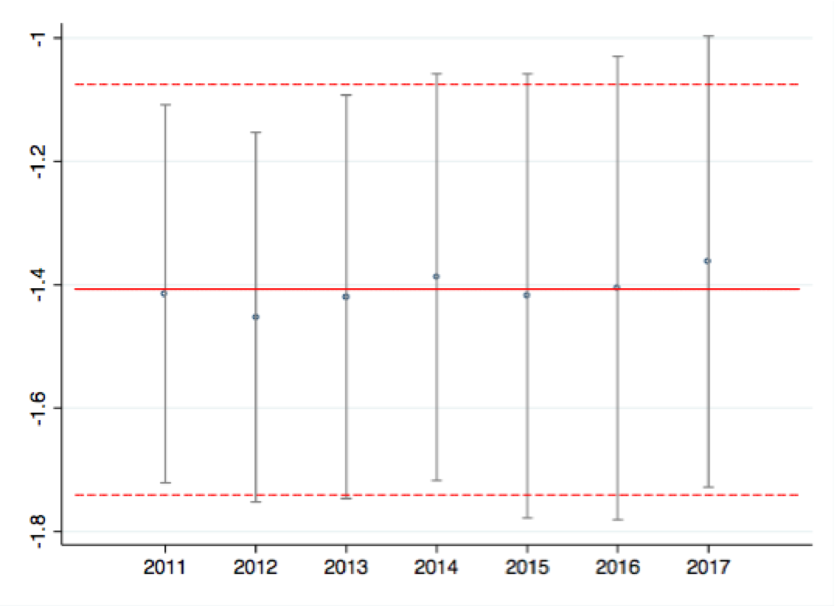
2.4.2. Border effect over time: 2011-2017

Finally, we explore whether the average border effect has remained constant throughout our sample. To test this, we allow the border effect to vary over time:

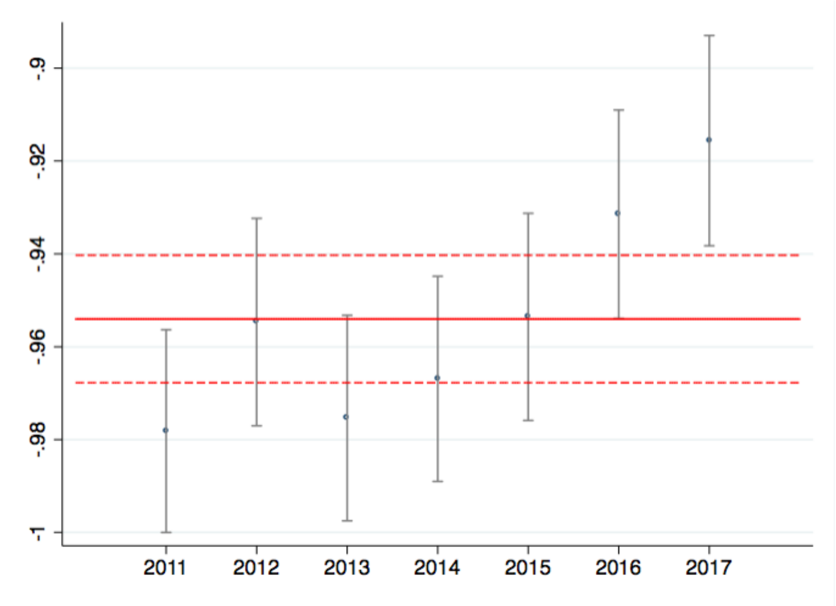
$$logV_{nmt}^{it} = \phi_t + \phi_k + \phi_n + \phi_m + \beta_t C_{nm} + \gamma_t h_{nm} + \phi' X + u_{nmkt}, \tag{2.8}$$

The estimated coefficients β_t are plotted in Figure 2.4. Over all, the effect of the border seems very stable over time. When we restrict the sample to the pairs of regions that are trading, then we see that the intensive-margin effect of the border has been falling during this period.

Figure 2.4: Border effect over time



A) All trade flows



B) Positive trade flows

Notes: The figure displays the coefficient of the year dummies (and 95% confidence interval bands). The first panel shows the coefficient of the year dummies introduced in the PPML estimation including the full sample (panel A) while the second panel reports the coefficient of the year dummies in the OLS estimation in the positive-trade sample (panel B). The continuous red line plots the estimated average effect estimated in the baseline specification and the discontinuous lines display the 95% confidence interval around this average effect.

2.5. “Random” borders: Identification of the border effect

In the previous section we presented our baseline estimation strategy. The identification of the border effect comes from comparing the value of trade between two regions that belong to different countries to the average trade flow between two regions that belong to the same country. The key assumption is that, controlling for other determinants of trade costs such as distance, contiguity and different languages, two regions that belong to different countries trade less than the average pair or regions *because* they do not share the same government. If our assumptions are correct, not sharing a government within Europe reduces trade flows as much as a 32.2% tariff. This result suggests that there is still considerable role for policy, since governments could put effort in reducing the obstacles to international trade created by the existence of the border. On the contrary, if our assumptions are wrong, the trade-reducing effect of the border could be caused by other factors and there may be not be a role for policy.

In this section we examine the threats to our identification strategy and propose an alternative estimation strategy to show that, indeed, the current institutions seem to be behind a large part of the border effect. The main challenge to identification comes from the fact that borders are not exogenous, but have been forming and evolving endogenously over time. Therefore, to establish the causality of our findings we have to consider a well-known threat to identification: selection bias.

2.5.1. Selection Bias

As we have already mentioned, borders divide regions into countries in a non-random way. In general, it is a reasonable assumption to consider that groups of regions have conformed the countries we observe today for historical, economic, political or cultural reasons. Maybe one of the key factors for a group of regions to form a country is that they like to trade with each other. If this was the main driver of country formation, we would see that regions within a country trade more with each other because of preferences or tastes that are previous (and exogenous) to the border. If regions “selected” into countries according to the level of trade between them, our estimated border effect would also be capturing this selection.

Using the concept of potential outcomes from the causal model (Angrist and Pischke, 2009), the causal effect of the border can be expressed as the difference between the expected bilateral trade of two regions that belong to different countries minus the expected bilateral trade of those two regions, had they been part of the same country (the potential outcome if the "treatment" had not happened). Following Angrist and Pischke (2009) let us denote the potential bilateral trade between two regions by V_{mn}. The bilateral trade that would be observed if m and n belong to different countries is $V_{nm,1}$ (this is, $Border_{n,m}$=1) and the bilateral trade that would be observed if m and n belong to the same country is $V_{nm,0}$, (meaning $Border_{nm}$=0). Given that we cannot make regions switch countries, we can only observe the difference in average bilateral trade conditional on being part of the same country or not:

$$\begin{aligned} E[V_{nm}|B_{nm}=1]-E[V_{nm}|B_{nm}=0] &= \underbrace{E[V_{nm,1}|B_{nm}=1]-E[V_{nm,0}|B_{nm}=1]}_{\text{Average causal effect}} \\ &\quad + \underbrace{E[V_{nm,0}|B_{nm}=1]-E[V_{nm,0}|B_{nm}=0]}_{\text{Selection Bias}} \end{aligned} \tag{2.9}$$

The average difference in trade flows between regions in different countries with respect to domestic regions can be decomposed into the average effect of the border and the selection bias.

The *average causal effect* is the difference in trade between regions divided by a border and how would they have traded, had they been sharing a government. According to our assumptions borders create some frictions to trade, since crossing a border implies entering into a different institutional framework with different rules, ways of doing business, etc. Therefore, we believe the causal effect of the border to be large and negative.

The *selection bias*, in this example, is the difference in the bilateral trade that we would observe between two regions that do not share a government *if a border did not exist* between them. If the reason why two regions divided by a border trade less is that they have a stronger preference for goods produced in their own countries, eliminating the border between these regions would not increase trade. We expect that the selection bias in this case to be negative and to increase the average difference observed in trade flows between foreign regions and domestic regions.

A way to neutralize the selection bias is the random assignment of "treatment" , making sure that both groups would be comparable in the case of not being treated. Thus, the random assignment of borders would eliminate the selection bias, allowing us to estimate the causal effect of the border by comparing the trade flows between regions separated by a border with the average domestic region-pair. In order to get closer to a random assignment of borders, we exploit the fact that culture, preferences and tastes generally change smoothly over space, while borders create discontinuities.

We focus on examining only trade flows that happen across contiguous regions located at international borders. By focusing on the sample of regions that are contiguous, we are making the assumption that border assignment is quasi-random in the proximity of the border. Indeed, conditional on a specific geographical location, the actual delineation of the border can be thought of as quasi-random. In this specification, the border effect would be identified by comparing bilateral trade across two regions that are contiguous but divided by a border with the average bilateral trade across a pair of contiguous regions that belong to the same country and are also located at the border. We run the following regression:

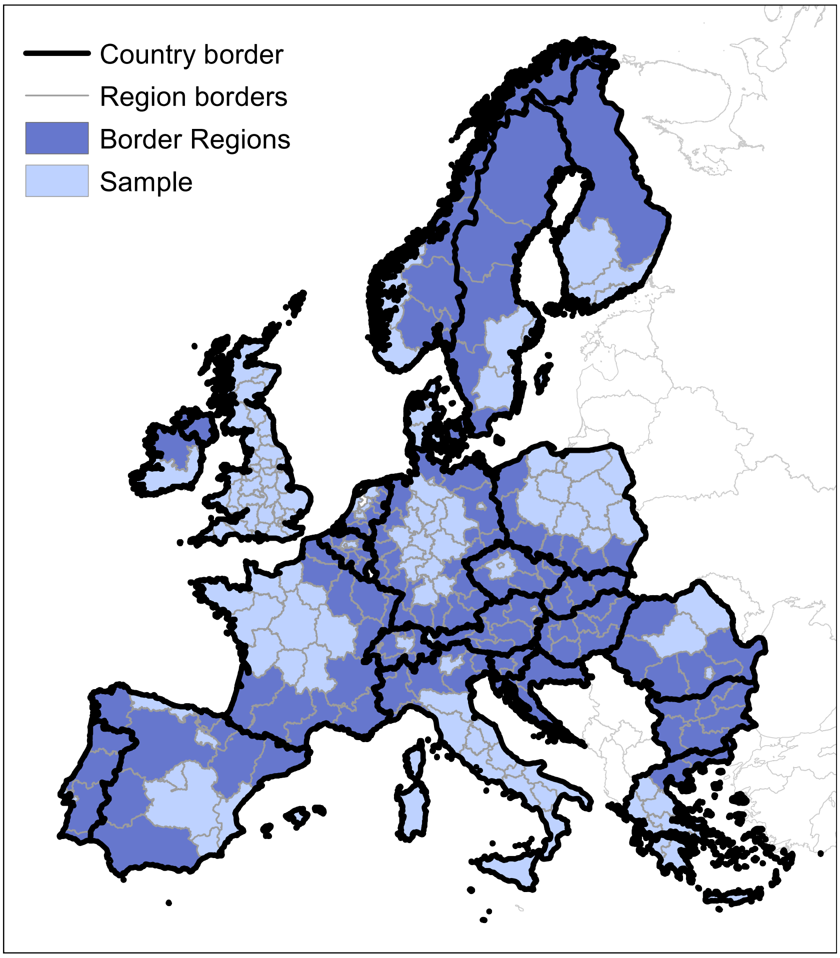
$$logV_{nm}^{it} = \phi_t + \phi_k + \phi_n + \phi_m + \beta C_{nm} + \gamma h_{nm} + \tau' X + u_{nmkt}, \tag{2.10}$$

where n,m denote region of origin and region of destination, k denotes industry and t denotes time. The dependent variable, $logV_{nm}^{it}$, denotes the logarithm of trade flows, C_{nm} indicates the existence of a border, h_{nm} measures the home bias and X is the vector of controls included earlier but without Contiguity.

For this empirical exercise we select pairs of regions that are contiguous to an international border. The identification of the border effect coefficient will come now from the difference between trade flows across two regions that are contiguous and divided by an international border with the average trade flows observed across two regions that are contiguous, located at the border but in the same of the international border. We select the sample of border regions as all the regions located at a political border and being contiguous to other border regions. Figure 2.5 displays our sample of border regions.

We run this specification in the complete sample as well as in the positive-trade flows sample. Table 2.9 reports the results from this estimation. Table 2.A10 reports the estimation results conditional on observing positive trade flows (OLS regression). Column (1) reports the baseline regression including all regions while column (2) shows the results on the sample of border regions. As we can see the border effect is still negative and significant, with an estimated coefficient of -1.351, slightly smaller than the average border effect estimated in column (1), -1.408.

Figure 2.5: Border Regions



This result suggests that the selection bias is not very important in our sample, since the result we get using only contiguous regions at the border is strikingly similar to the average effect. In addition, the coefficient is a bit smaller confirming, to some extent, our hypothesis that the selection bias would tend to increase the border effect. These alternative identification strategy confirms our initial findings, namely, that country borders are still an important obstacle to trade for European countries. Further work should be directed into exploring whether the border effects that we estimated in this paper come purely from current borders or are influenced, to some extent, by past borders.

Table 2.9: Border Effect - All vs Border regions

Outcome: Bilateral trade (Value)	Baseline	Border Regions
	PPML	PPML
	(1)	(2)
Border effect	-1.408***	-1.351***
	(0.170)	(0.164)
log(Dist)	-1.469***	-0.588*
	(0.084)	(0.316)
Contiguity	0.238**	
	(0.112)	
Language	0.533***	0.649***
	(0.139)	(0.153)
Home Bias	1.214***	1.805***
	(0.217)	(0.359)
Origin FE	Yes	Yes
Destination FE	Yes	Yes
Year FE	Yes	Yes
Industry FE	Yes	Yes
Obs.	4524912	40068
R-squared	0.613	0.659

Notes: Standard errors are in parentheses clustered at the country-pair level.* significant at 10%, ** significant at 5%, *** significant at 1%. Column 1 reports the results including all region-pairs in our sample. Column 2 restricts the sample to a set of origin/destination pairs that are contiguous and are located at an international border.

2.6. Conclusion

Our findings uncover that national borders explain a large reduction in bilateral trade flows across European regions. These findings are confirmed in a specification where the identification of the border effect comes from comparing border, neighbouring regions.

We find that not sharing a government is equivalent to 32.5% tariff. Even if we have done our best to document the presence of the border effect, we still cannot provide one unique explanation for the size of the borders within Europe. If there is a role for policy, it is important to continue exploring which factors are responsible for the large difference in how two regions trade when they belong to the same government vis-a-vis when they do not share a government.

A number of possible explanations are clear candidates. The existence of different legal systems and institutional settings may create non-tariff barriers for firms. The behaviour of governments would also result in a large border effect. In addition, the internal structure of firms may be most usually enclosed within the borders of a country contributing to the border effect. We hope to keep exploring this rich dataset to give answers to some of these questions in future research.

There are, at least, three avenues for further research aimed at providing a better characterization of the border effect and a sharper description of how its size relates to specific policies and institutions. The first one would be to explore how the border effect varies across industries and time. We do present some results here, but our dataset covers only 9 very broad industries and for only 7 years. Extending the analysis to a finer industry classification and more years could reveal useful patterns that allow us to relate the size of the border effect to some aspects of the legal framework or the use of some local business norms. It would also allow us to determine impact on the border effect of specific policy changes, such as the creation of the Euro.

The second one refers to the duration of the border effect. Is the border effect estimated here attributable to current borders? Or is it the result of past borders and inertia? Answering this question is important to establish the policy implications of our results. Current borders are highly correlated to past borders, making it difficult to determine the source of our estimated border effect. If a present border is removed, or a new border is created, would this affect the border effect? To answer this question, we need to devise an identification strategy that uses past and present borders, and tries to establish the extent to which the border effect we estimate is the result of present or past borders.

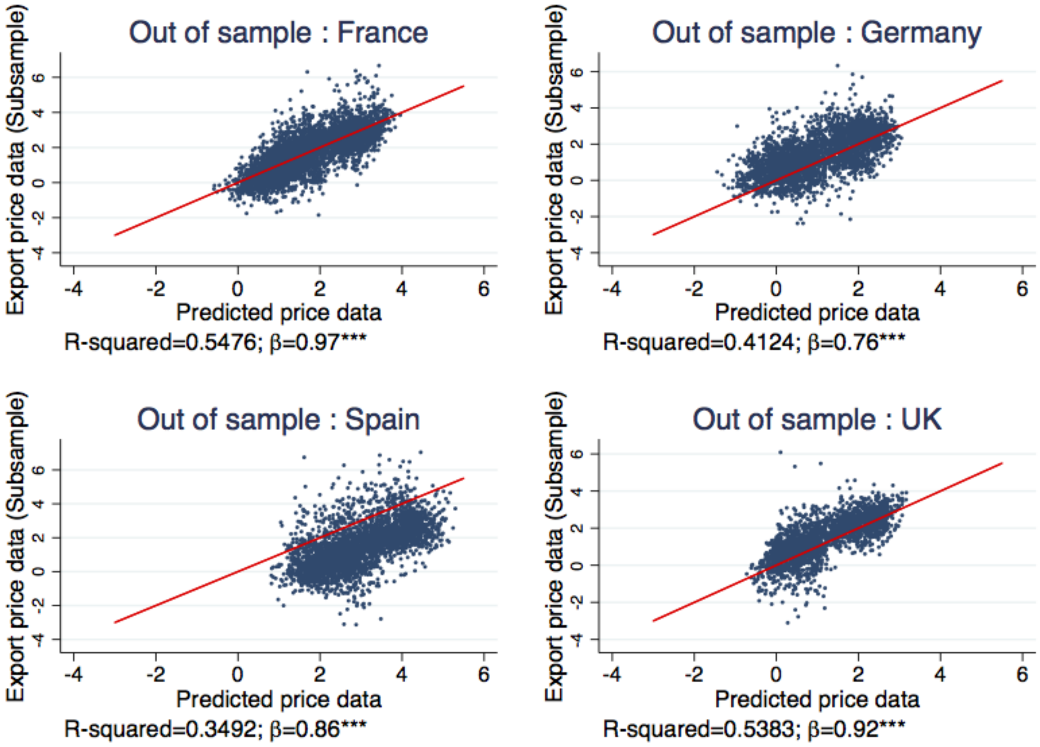
The third avenue for further research would recognize that all borders are not the same, and it would try to characterize heterogeneity across borders. Is the border effect between Sweden and Norway the same as that between Poland and Germany, or Spain and Portugal? Throughout the paper, we have estimated an average border effect. But this average effect is likely to mask substantial heterogeneity. If this is the case, this heterogeneity might allow us to determine what aspects of government are conducive to a large border effects and which aspects are not.

2.7. Additional Figures

Figure 2.A1: Explanatory variables in the price regressions

Variable	**Source**
Log GDP pc	Eurostat
Log Population density	Eurostat
Empoyment % (>25 years)	Eurostat
Employment % in hightech	Eurostat
Employment % in manufacturing	Eurostat
Employment % in low tech	Eurostat
Life expectancy	Eurostat
Share of pop with primary educ	Eurostat Census 2011
Share of pop with tertiary educ	Eurostat Census 2011
Share of pop with no education	Eurostat Census 2011
Share of migrants	Eurostat Census 2011
share of migrants out of EU	Eurostat Census 2011
Share of agriculture	Eurostat Census 2011
Share of manufacturing	Eurostat Census 2011
Share of information sector	Eurostat Census 2011
Share of finance sector	Eurostat Census 2011
Share of science and research	Eurostat Census 2011
Share of public sector	Eurostat Census 2011
Average temperature	Eurostat
Distance to closest river	Own collection
Distance to clostest coast	Own collection

Figure 2.A2: Out-of-sample Estimates



R-squared=0.5476; β=0.97***

R-squared=0.4124; β=0.76***

R-squared=0.3492; β=0.86***

R-squared=0.5383; β=0.92***

Figure 2.A3: Country-to-Country Estimates

R-squared=0.4470; β=0.837***

2.8. Additional Tables

Table 2.A1: Price regressions

	DEP.VAR: Log Price	DEP.VAR: Log Price
	(1)	(2)
log(Temperature)$_o$	-0.955***	-0.937***
	(0.039)	(0.038)
log(Temperature)$_d$	-0.418***	-0.293***
	(0.062)	(0.061)
log(Dist to River)$_o$	0.057***	0.053***
	(0.004)	(0.004)
log(Dist to Coast)$_o$	0.061***	0.073***
	(0.007)	(0.007)
log(Dist to River)$_d$	-0.081***	-0.078***
	(0.012)	(0.012)
log(Dist to Coast)$_d$	-0.149***	-0.006
	(0.018)	(0.018)
log(Dist)$_{o,d}$		0.451***
		(0.012)
Constant	10.550***	0.728
	(1.184)	(1.194)
Industry-Year FE	Yes	Yes
Origin Variables	Yes	Yes
Destination Variables	Yes	Yes
Obs.	48995	48995
R-squared	0.525	0.539

Notes: Standard errors are in parentheses.* significant at 10%, ** significant at 5%, *** significant at 1%. First column displays the results including only origin and destination level variables. The second column reports the results when adding the bilateral distance between origin and destination as a determinant of export prices.

Table 2.A2: Country-level summary statistics

	Agg. Exports data				Road Freight Data			
	Mean	Sd.	Min	Max	Mean	Sd.	Min	Max
Bilateral exports (Mill.)	458.14	1550.73	0.00	32000.00	186.10	754.33	0.00	13225.15
Bilateral exports (Mill. Kg)	220.92	1029.70	0.00	31497.11	83.33	347.98	0.00	7438.19
Export Price (per Kg)	8.27	25.19	0.02	1461.67	3.90	4.46	0.13	44.51
Observations	23400				23400			

Table 2.A3: Members of the Schengen Agreement

Agreement	Country members
Schengen Area	Austria
	Belgium
	Czech Republic
	Denmark
	Finland
	France
	Germany
	Greece
	Hungary
	Italy
	Netherlands
	Norway
	Poland
	Portugal
	Slovakia
	Slovenia
	Spain
	Sweden
	Switzerland
Total	19

Table 2.A4: Members of the Euro Area

Agreement	Country members
Euro Zone	Austria
	Belgium
	Finland
	France
	Germany
	Greece
	Ireland
	Italy
	Netherlands
	Portugal
	Slovakia
	Slovenia
	Spain
Total	13

Table 2.A5: Border Effect - Additional Controls

Outcome: Bilateral Trade (Value)	Positive flows		Full sample
	OLS	PPML	PPML
	(1)	(2)	(2)
Border effect	-0.992***	-1.124***	-1.287***
	(0.078)	(0.178)	(0.159)
log(Dist)	-0.566***	-0.685***	-1.417***
	(0.067)	(0.079)	(0.074)
Contiguity	0.695***	0.700***	0.285***
	(0.082)	(0.128)	(0.108)
Language	0.015	-0.023	0.416***
	(0.061)	(0.103)	(0.120)
Schengen	1.148***	0.772**	1.488***
	(0.299)	(0.309)	(0.315)
Eurozone	-0.027	-0.180*	0.374**
	(0.086)	(0.093)	(0.149)
Home Bias	1.569***	1.709***	1.247***
	(0.117)	(0.222)	(0.217)
Constant	19.751***	20.015***	25.837***
	(0.407)	(0.607)	(0.657)
Origin FE	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Obs.	5.37e+05	5.37e+05	4.52e+06
R-Squared	0.587	0.618	0.613

Notes: Standard errors are in parentheses clustered at the country-pair level.* significant at 10%, ** significant at 5%, *** significant at 1%.

Table 2.A6: Average Border Effect - No price adjustment

Outcome: Bilateral Trade (Weight)	Positive flows		Full sample
	OLS	PPML	PPML
	(1)	(2)	(2)
Border effect	-1.006***	-1.124****	-1.369***
	(0.082)	(0.165)	(0.162)
log(Dist)	-0.619***	-0.863***	-1.595***
	(0.061)	(0.082)	(0.069)
Contiguity	0.643***	0.623***	0.182
	(0.092)	(0.110)	(0.111)
Language	0.072	0.018	0.601***
	(0.074)	(0.118)	(0.141)
Home Bias	1.536***	1.554***	1.088***
	(0.122)	(0.171)	(0.162)
Origin FE	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Obs.	5.37e+05	5.37e+05	4.52e+06
R-Squared	0.515	0.680	0.671

Notes: Standard errors are in parentheses clustered at the country-pair level.* significant at 10%, ** significant at 5%, *** significant at 1%.

Table 2.A7: Average Border Effect - Price adjustment including Distance

Outcome: Bilateral Trade (Value 2)	Positive flows		Full sample
	OLS	PPML	PPML
	(1)	(2)	(2)
Border effect	-1.007***	-1.025***	-1.423***
	(0.082)	(0.148)	(0.153)
log(Dist)	-0.168***	-0.182**	-1.021***
	(0.061)	(0.083)	(0.099)
Contiguity	0.643***	0.805***	0.254**
	(0.092)	(0.119)	(0.118)
Language	0.072	-0.017	0.516***
	(0.074)	(0.089)	(0.138)
Home Bias	1.536***	1.800***	1.232***
	(0.122)	(0.229)	(0.229)
Constant	17.604***	17.378***	23.249***
	(0.226)	(0.488)	(0.837)
Origin FE	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Obs.	5.37e+05	5.37e+05	4.52e+06
R-Squared	0.498	0.554	0.535

Notes: Standard errors are in parentheses clustered at the country-pair level.* significant at 10%, ** significant at 5%, *** significant at 1%.

Table 2.A8: Border Effect in Europe by Industry - OLS

Dep Variable: Weight	Food/Bev	Textiles	Wood	Chem	Non-Met	Metal	Machinery	Vehicles	Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Border effect	-1.635***	-0.619***	-1.081***	-0.844***	-1.165***	-0.922***	-0.987***	-0.885***	-0.963***
	(0.096)	(0.101)	(0.109)	(0.090)	(0.106)	(0.083)	(0.106)	(0.095)	(0.136)
log(Dist)	-0.827***	-0.299***	-0.793***	-0.839***	-0.950***	-0.673***	-0.408***	-0.433***	-0.382***
	(0.074)	(0.052)	(0.075)	(0.063)	(0.116)	(0.069)	(0.062)	(0.073)	(0.047)
Contiguity	0.840***	0.296***	0.709***	0.495***	1.097***	0.635***	0.776***	0.494***	0.491***
	(0.073)	(0.083)	(0.097)	(0.109)	(0.117)	(0.095)	(0.119)	(0.104)	(0.076)
Language	0.194***	-0.045	-0.068	0.188**	0.017	0.132*	0.144*	0.050	-0.145
	(0.071)	(0.099)	(0.106)	(0.085)	(0.087)	(0.078)	(0.079)	(0.107)	(0.114)
Home Bias	1.202***	1.335***	1.345***	1.374***	2.249***	1.537***	2.132***	1.707***	1.355***
	(0.115)	(0.182)	(0.130)	(0.129)	(0.190)	(0.113)	(0.192)	(0.146)	(0.104)
Origin FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	88957	26283	66336	73401	51555	68844	61223	56983	43565
R-squared	0.678	0.377	0.573	0.536	0.689	0.540	0.500	0.456	0.390

Notes: Standard errors are in parentheses clustered at the country-pair level.* significant at 10%, ** significant at 5%, *** significant at 1%.

Table 2.A9: Border Effect in Europe by Industry - PPML (ALL TRADE FLOWS)

Dep Variable: Weight	Food/Bev	Textiles	Wood	Chem	Non-Met	Metal	Machinery	Vehicles	Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Border effect	-1.734***	-1.409***	-1.188***	-0.973***	-1.296***	-1.192***	-1.328***	-1.129***	-1.510***
	(0.161)	(0.159)	(0.123)	(0.120)	(0.234)	(0.110)	(0.209)	(0.177)	(0.119)
log(Dist)	-1.593***	-1.350***	-1.716***	-1.717***	-1.931***	-1.540***	-1.204***	-1.241***	-1.307***
	(0.055)	(0.077)	(0.096)	(0.080)	(0.065)	(0.072)	(0.094)	(0.082)	(0.072)
Contiguity	0.100	-0.000	-0.008	-0.064	0.489***	0.008	0.346***	0.132	-0.031
	(0.101)	(0.100)	(0.096)	(0.132)	(0.116)	(0.109)	(0.134)	(0.104)	(0.064)
Language	0.746***	0.466***	0.613***	0.475***	0.826***	0.574***	0.689***	0.381**	0.445***
	(0.158)	(0.154)	(0.145)	(0.161)	(0.154)	(0.133)	(0.154)	(0.183)	(0.164)
Home Bias	0.536***	0.707***	0.671***	0.678***	1.605***	0.705***	1.481***	1.191***	0.741***
	(0.114)	(0.186)	(0.082)	(0.125)	(0.223)	(0.112)	(0.249)	(0.151)	(0.110)
Origin FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	502768	484183	502768	502768	502768	499023	495292	497154	487872
R-squared	0.897	0.478	0.846	0.866	0.939	0.887	0.823	0.836	0.678

Notes: Standard errors are in parentheses clustered at the country-pair level.* significant at 10%, ** significant at 5%, *** significant at 1%.

Table 2.A10: Border Effect - All vs Border regions

Outcome: Bilateral trade (Value)	Baseline	Border Regions
	OLS	OLS
	(1)	(2)
Border effect	-1.007***	-1.351***
	(0.082)	(0.085)
log(Dist)	-0.619***	-0.722***
	(0.061)	(0.139)
Contiguity	0.643***	
	(0.092)	
Language	0.072	0.325***
	(0.074)	(0.098)
Home Bias	1.536***	1.557***
	(0.122)	(0.123)
Origin FE	Yes	Yes
Destination FE	Yes	Yes
Year FE	Yes	Yes
Obs.	537147	32774
R-squared	0.583	0.710

Notes: Standard errors are in parentheses clustered at the country-pair level.* significant at 10%, ** significant at 5%, *** significant at 1%.

2.9. Data Appendix

2.9.1. Price imputation

Our goal is to estimate an average bilateral price for each origin/destination pair across the different industries and years. To do this, we collect a wide set of variables at the region level as well as at the country level. To test the ability of our price estimate to predict actual prices we collect export prices. Finally, to show that our dataset of regional trade captures a large part of total trade across Europe, we collect aggregate trade values. The list of variables collected is displayed in table 2.A1. We use this appendix explains the sources of data used.

Regional price data The region to country trade flows, in values and quantities, are collected individually for each country. The price data from Spain comes from C-Interreg project.[11] The data from France comes from the National direction of International trade statistics (Direction Nationale des Statistiques du Commerce Extérieur). The data from Germany come from the Federal Office of Statistics (Statistiches Bundesamt). Finally, the data from the United Kingdom is collected from the statistical department of the United Kingdom government (https://www.gov.uk/search/research-and-statistics).

Region and Country level variables Most aggregate variables come from Eurostat. A set of this variables come from the 2011 census and are only available for this year. The temperature variable comes from Eurostat. Geographical distances such as closeness to a River or Coast are collected from the EuroRegionalMap dataset using ArcGIS.[12]

International Trade We collect aggregate trade flows across European countries in our sample. We construct export unit prices to assess our price imputation method by dividing the value of trade over the volume in kilos.

2.9.2. Gravity equation

We collect some variables to use as controls in the gravity equation.

Distance The distance is computed from the latitude and longitude of the centroid of the origin region and the centroid of the destination region. We use the geodesic distance in kilometres. Distance within a region is computed using the formula proposed by Nitsch, 2000:

$$\text{Internal distance} = 0.5 * \sqrt{area} \tag{2.11}$$

Contiguity Region-level contiguity is computed using a shapefile of European regions (form Eurostat) using ArcGIS. Contiguity at the country level is takn from the CEPII dataset of gravity variables.

---

[11] http://212.227.102.53/explotacion$_{m}ultidimensional_{c}omercio_{i}nterregional/estadisticas.aspx$

[12] EuroRegional Map: https://eurogeographics.org/products-and-services/euroregionalmap/

Language Region-level languages are collected from language maps in Encyclopedia Britannica. We manually collect language by region and create a dummy to indicate a common language. The common language at the country level comes from the CEPII dataset of gravity variables.

Additional controls We collect data on the currency in each country from CEPII dataset and from membership in the Schengen Agreement from the European Commission.

Bibliography

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2018). The economics of density: Evidence from the Berlin Wall. *Econometrica 83*(6), 2127–2189.

Alder, S. (2014). Chinese roads in India: The effect of transport infrastructure on economic development. *Work. Pap., Univ. North Carolina, Chapel Hill*.

Alder, S. and I. Kondo (2018). Political Distortions and Infrastructure Networks in China: A Quantitative Spatial Equilibrium Analysis. In *2018 Meeting Papers*. Society for Economic Dynamics.

Allen, T. and C. Arkolakis (2014, August). Trade and the Topography of the Spatial Economy. *The Quarterly Journal of Economics 129*(3), 1085–1140.

Allen, T. and C. Arkolakis (2017). The welfare effects of transportation infrastructure improvements. *Working Paper*.

Allen, T., C. Arkolakis, and Y. Takahashi (2014). Universal gravity. *Working Paper*.

Allen, T., C. Arkolakis, and Y. Takahashi (2015). Universal gravity. *Working Paper*.

Anderson, J. E. and E. Van Wincoop (2003). Gravity with gravitas: a solution to the border puzzle. *American economic review 93*(1), 170–192.

Asturias, J., M. García-Santana, and R. Ramos (2014). Competition and the welfare gains from transportation infrastructure: Evidence from the Golden Quadrilateral of India. *Work. Pap., Georgetown Univ., Washington, DC*.

Atack, J., M. R. Haines, and R. A. Margo (2008, October). Railroads and the Rise of the Factory: Evidence for the United States, 1850-70. Working Paper 14410, National Bureau of Economic Research.

Bai, Y. and R. Jia (2018). The Oriental City: Political Hierarchy and Regional Development in China, AD1000-2000. *Working Paper*.

Balboni, C. (2017). *Living on the Edge: Infrastructure Investments and the Persistence of Coastal Cities*. manuscript.

Banerjee, A., E. Duflo, and N. Qian (2012). On the road: Access to transportation infrastructure and economic growth in China. Technical report, National Bureau of Economic Research.

Baum-Snow, N. (2007). Did highways cause suburbanization? *The Quarterly Journal of Economics 122*(2), 775–805.

Baum-Snow, N., L. Brandt, J. V. Henderson, M. A. Turner, and Q. Zhang (2017). Roads, railroads, and decentralization of Chinese cities. *Review of Economics and Statistics 99*(3), 435–448.

Baum-Snow, N., J. V. Henderson, M. Turner, L. Brandt, and Q. Zhang (2015). Transport Infrastructure, Urban Growth and Market Access in China.

Bemrose, R., W. M. Brown, and J. Tweedle (2017). Going the distance: estimating the effect of provincial borders on trade when geography matters. Technical report, Statistics Canada, Analytical Studies Branch.

Broda, C., N. Limao, and D. E. Weinstein (2008, December). Optimal Tariffs and Market Power: The Evidence. *American Economic Review 98*(5), 2032–2065.

Broda, C. and D. E. Weinstein (2006). Globalization and the Gains from Variety. *The Quarterly journal of economics 121*(2), 541–585.

Brulhart, M., C. Carrere, and F. Trionfetti (2012). How wages and employment adjust to trade liberalization: Quasi-experimental evidence from Austria. *Journal of International Economics 86*(1), 68–81.

Chen, N. (2004). Intra-national versus international trade in the European Union: why do national borders matter? *Journal of International Economics 63*(1), 93–118.

Chen, N. and D. Novy (2011). Gravity, trade integration, and heterogeneity across industries. *Journal of international Economics 85*(2), 206–221.

Combes, P.-P. and M. Lafourcade (2005). Transport costs: measures, determinants, and regional policy implications for France. *Journal of Economic Geography 5*(3), 319–349.

Combes, P.-P., M. Lafourcade, and T. Mayer (2003). Can business and social networks explain the border effect puzzle?

Davis, D. R. and D. E. Weinstein (2002). Bones, bombs, and break points: the geography of economic activity. *American Economic Review 92*(5), 1269–1289.

Disdier, A.-C. and K. Head (2008, February). The Puzzling Persistence of the Distance Effect on Bilateral Trade. *Review of Economics and Statistics 90*(1), 37–48.

Donaldson, D. (2018). Railroads of the Raj: Estimating the impact of transportation infrastructure. *American Economic Review 108*(4-5), 899–934.

Donaldson, D. and R. Hornbeck (2016, May). Railroads and American Economic Growth: A âMarket Accessâ Approach. *The Quarterly Journal of Economics 131*(2), 799–858.

Duranton, G., P. M. Morrow, and M. A. Turner (2014, April). Roads and Trade: Evidence from the US. *Review of Economic Studies 81*(2), 681–724.

Duranton, G. and M. A. Turner (2012, October). Urban Growth and Transportation. *Review of Economic Studies 79*(4), 1407–1440.

Eaton, J. and S. Kortum (2002, September). Technology, Geography, and Trade. *Econometrica 70*(5), 1741–1779.

Faber, B. (2014, July). Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System. *The Review of Economic Studies 81*(3), 1046–1070.

Fajgelbaum, P. D. and E. Schaal (2017). Optimal transport networks in spatial equilibrium. Technical report, National Bureau of Economic Research.

Feenstra, R. C. (2002). Border effects and the gravity equation: Consistent methods for estimation. *Scottish Journal of Political Economy 49*(5), 491–506.

Felbermayr, G. J. and A. Tarasov (2015, November). Trade and the Spatial Distribution of Transport Infrastructure. SSRN Scholarly Paper ID 2706533, Social Science Research Network, Rochester, NY.

Fogel, R. W. (1962). A Quantitative Approach to the Study of Railroads in American Economic Growth: A Report of Some Preliminary Findings. *The Journal of Economic History 22*(2), 163–197.

Fretz, S., R. Parchet, and F. Robert-Nicoud (2017). Highways, Market Access, and Spatial Sorting.

Gallen, T. and C. Winston (2018). Transportation Capital and its Effects on the US Economy: A General Equilibrium Approach.

Garcia-Lopez, M.-a., A. Holl, and E. Viladecans-Marsal (2015). Suburbanization and highways in Spain when the Romans and the Bourbons still shape its cities. *Journal of Urban Economics 85*, 52–67.

Ghani, E., A. G. Goswami, and W. R. Kerr (2016, March). Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing. *The Economic Journal 126*(591), 317–357.

Gil-Pareja, S., R. Llorca-Vivero, J. A. Martínez-Serrano, and J. Oliver-Alonso (2005). The border effect in Spain. *World Economy 28*(11), 1617–1631.

Head, K. and T. Mayer (2000). Non-Europe: the magnitude and causes of market fragmentation in the EU. *Review of World Economics 136*(2), 284–314.

Head, K. and T. Mayer (2009). Illusory border effects: distance mismeasurement inflates estimates of home bias in trade. In *In: The Gravity Model in International Trade: Advances and Applications. Editors: Bergeijk and Brakman*. Citeseer.

Head, K. and T. Mayer (2014). Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of international economics*, Volume 4, pp. 131–195. Elsevier.

Heblich, S., S. J. Redding, and D. M. Sturm (2018). The Making of the Modern Metropolis: Evidence from London. Technical report, National Bureau of Economic Research.

Helliwell, J. F. (2002). Measuring the width of national borders. *Review of International Economics 10*(3), 517–524.

Helpman, E., M. Melitz, and Y. Rubinstein (2008). Estimating trade flows: Trading partners and trading volumes. *The quarterly journal of economics 123*(2), 441–487.

Hillberry, R. (1999). Explaining the ?border effect?: what can we learn from disaggregated commodity flow data. *Indiana University Graduate Student Economics Working Paper Series 9802*.

Hillberry, R. and D. Hummels (2008). Trade responses to geographic frictions: A decomposition using micro-data. *European Economic Review 52*(3), 527–550.

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *The American Economic Review 70*(5), 950–959.

Kyoji, F., O. Toshihiro, K. Fukao, and T. Okubo (2004). Why has the Border effect in the Japanese Market Declined? The Role of Business Networks in East Asia.

Magerman, G., Z. Studnicka, and J. Van Hove (2016). Distance and border effects in international trade: A comparison of estimation methods. *Economics: The Open-Access, Open-Assessment E-Journal 10*(2016-18), 1–31.

McCallum, J. (1995). National borders matter: Canada-US regional trade patterns. *The American Economic Review 85*(3), 615–623.

Michaels, G. (2008). The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System. *The Review of Economics and Statistics 90*(4), 683–701.

Moller, J. and M. Zierer (2018, January). Autobahns and jobs: A regional study using historical instrumental variables. *Journal of Urban Economics 103*, 18–33.

Monte, F., S. J. Redding, and E. Rossi-Hansberg (2015). Commuting, migration and local employment elasticities. Technical report, National Bureau of Economic Research.

Morten, M. and J. Oliveira (2018). The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City. *NBER Working Paper 22158*.

Nagy, D. K. (2016). City location and economic development. *Princeton University, mimeograph.*

Nitsch, V. (2000). National Borders and International Trade: Evidence from the European Union. *The Canadian Journal of Economics / Revue canadienne d'Economique 33*(4), 1091–1105.

Novy, D. (2013). International trade without CES: Estimating translog gravity. *Journal of International Economics 89*(2), 271–282.

Redding, S. and D. Sturm (2016). Estimating Neighborhood Effects: Evidence from War-time Destruction in London, draft. Technical report.

Redding, S. J. (2016). Goods trade, factor mobility and welfare. *Journal of International Economics 101*, 148–167.

Redding, S. J. and E. Rossi-Hansberg (2017). Quantitative spatial economics. *Annual Review of Economics 9*, 21–58.

Redding, S. J. and D. M. Sturm (2008, December). The Costs of Remoteness: Evidence from German Division and Reunification. *American Economic Review 98*(5), 1766–1797.

Redding, S. J. and M. A. Turner (2015). Transportation costs and the spatial organization of economic activity. In *Handbook of regional and urban economics*, Volume 5, pp. 1339–1398. Elsevier.

Rossi-Hansberg, E. (2005). A spatial theory of trade. *American Economic Review 95*(5), 1464–1491.

Silva, J. S. and S. Tenreyro (2006). The log of gravity. *The Review of Economics and statistics 88*(4), 641–658.

Tsivanidis, N. (2018). The Aggregate and Distributional Effects of Urban Transit Infrastructure: Evidence from BogotÃ¡âs TransMilenioâ. Technical report, Working Paper, University of Chicago Booth School of Business.

Turrini, A. and T. van Ypersele (2010). Traders, courts, and the border effect puzzle. *Regional Science and Urban Economics 40*(2-3), 81–91.

Voigtländer, N. and H.-J. Voth (2015, October). Highway to Hitler. SSRN Scholarly Paper ID 2684404, Social Science Research Network, Rochester, NY.

Wei, S.-J. (1996). Intra-national versus international trade: how stubborn are nations in global integration? Technical report, National Bureau of Economic Research.

Wolf, H. C. (2000). Intranational home bias in trade. *Review of economics and statistics 82*(4), 555–563.

Wolf, N. (2009). Was Germany ever united? Evidence from intra-and international trade, 1885â1933. *The Journal of Economic History 69*(3), 846–881.

Zeller, T. and T. Dunlap (2010). *Driving Germany: The Landscape of the German Autobahn, 1930-1970* (1 ed.). Berghahn Books.