

# The Spoils of War: Trade Shocks during WWI and Spain's Regional Development \*

Simon Fuchs <sup>†</sup>

[Click here for the latest version](#)

July 30, 2018

## Abstract

This paper analyzes to what extent labor market frictions limit the gains from market integration. I use an external demand shock to the Spanish economy as a natural experiment to identify and quantify the effect of labor mobility costs on Spain's development. Using newly digitized trade and labor market data, I show that during WWI (1914-1918) a large, temporary and sectorally heterogeneous demand shock emanated from belligerent countries, as a result of which Spain expanded its manufacturing employment and exports, while income growth between the north and south in Spain diverged. To quantify and analyse the role of mobility costs I build and estimate a multi-sector economic geography model that allows for sectoral and spatial mobility costs. Spatial mobility costs dominated with an estimated 80% of reallocation of labor taking place *within* rather than *between* provinces. I use the estimated model to calculate counterfactuals to examine the effects of and interaction between output and input market integration: Comparing to the non-shock counterfactual I find that the WWI-shock increased manufacturing employment by 10%, and induced highly uneven spatial development with the north growing 27% faster. The shock constituted a 6% increase in market size and increased aggregate real incomes by 20%. Lowering mobility costs by 10% increases real income gains from the WWI-shock by an additional 3%, and exceeds gains in the non-shock scenario, suggesting that labor market integration and output market integration are complements.

*JEL classification:* D5, F11, F12, F15, F16, N9, N14, R12, R13

---

\*I am indebted to my advisor Thomas Chaney for his invaluable guidance. For crucial help and advice, I also want to thank Treb Allen, Nina Pavcnik and the research community at Dartmouth College - much of the paper was written while visiting the college. For their comments, I am grateful to Konrad Adler, Andy Bernard, Emily Blanchard, Klaus Desmet, Chris Edmond, Patrick Feve, Jim Feyrer, Sharat Ganapati, Christian Hellwig, Rob Johnson, Tim Lee, Thierry Mayer, Rory McGee, Marti Mestieri, Franck Portier, Vincent Rebeyrol, Mohammed Saleh, Chris Snyder, Francois de Soyres, Bob Staiger, Robert Ulbricht, Nikolaus Wolf as well as seminar participants at Dartmouth College, the ENTER Jamboree at UCL, the Interwar Economic history workshop at LSE, the joint TSE-Melbourne University workshop and the Macro and Applied Micro Workshop at TSE. In particular I want to thank Javier Silvestre for sharing his data on internal migration in Spain at the beginning of the 20th century, the administrative staff at Biblioteca Nacional de Catalunya for aiding me to find the right documents, and Begoña Alarcón for helping me translate and digitize historical Spanish records. Finally, I acknowledge financial support via ERC grant No 337272-FiNet. All remaining errors are my own.

<sup>†</sup>Toulouse School of Economics, [sfuchs-de.github.io](https://github.com/sfuchs), [sfuchs.de@gmail.com](mailto:sfuchs.de@gmail.com)

*“Spain is today a bundle of small  
bodies tied together by a rope of sand.”*

---

Ford (1845), p.2

# 1 Introduction

Why might an economy be trapped at a low level of economic development? Why is the adjustment to trade liberalization slow and often does not seem to effectively equilibrate local labor markets across space?<sup>1</sup> A common explanation to these questions is that high mobility costs and low initial gains to the worker from migration might prevent labor reallocation towards higher productivity sectors and regions. This can in turn limit the gains from market integration and undermine development, but to what extent this might be the case is difficult to determine. Understanding and quantifying these frictions is therefore of primary importance in understanding the obstacles to growth and structural change in developing countries as well as the welfare effects of trade liberalization episodes.

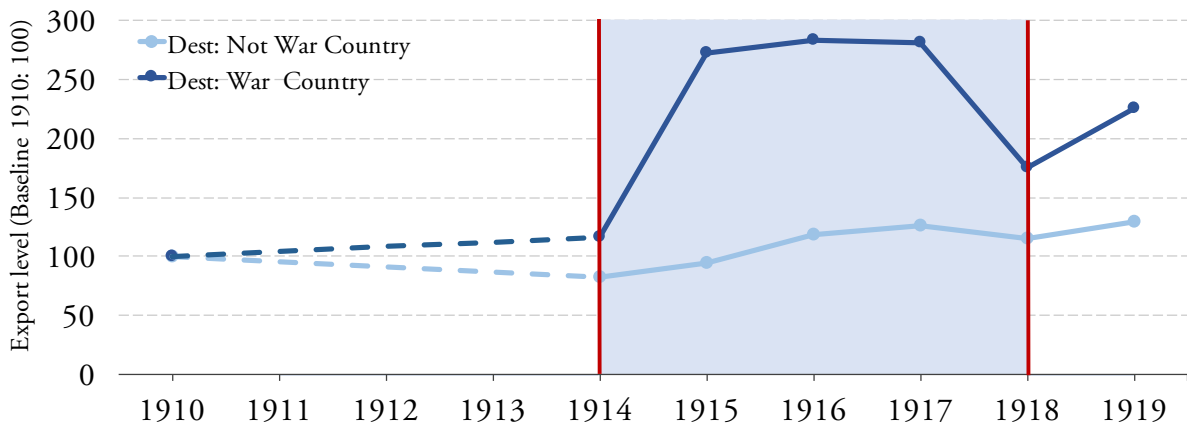
However, empirically verifying and quantifying these frictions is challenging, since neither labor market frictions nor the counterfactual gains from reallocation are directly observed. This paper overcomes this problem by using a natural experiment where a foreign demand shock reallocates labor across sectors and space. The reallocation patterns are informative about the sectoral and spatial mobility frictions that inhibit labor movements even in the absence of a shock. Using the shock in tandem with an economic geography model, I show how to estimate the gains from reallocation as well as the labor market frictions. The key point of this paper is to illustrate how and to what extent labor market frictions can limit gains from market integration and how this can be analyzed in a setting where a temporary foreign demand shock reallocates labor by creating temporary gains that offset adjustment costs. This analysis singles out mobility costs as a key factor in determining the size of welfare gains from market integration as well as the spatial distribution.

This study examines a trade shock to the Spanish economy that was caused by the participation of Spain’s key trading partners - in particular France - in the first World War (1914-1918) while Spain remained neutral. Prior to the shock, Spain had experienced a

---

<sup>1</sup>The empirical finding that local labor markets only adjust slowly to shocks goes back to [Blanchard and Katz \(1992\)](#). [Dix-Carneiro and Kovak \(2017\)](#) show that labor market frictions and slow adjustment processes can permanently prevent spatial arbitrage.

Figure 1: Aggregate Trade levels



*Notes:* This figure compares aggregate export levels in constant pre-war prices between destination countries that participated in WWI and those that did not. To adjust for additional spatial disruptions of the frontline the belligerent countries are made up of France, Italy and the United Kingdom. The non-belligerent countries exclude the United States and other later participants of WWI. Data is not available for the years between 1910 and 1914 therefore a trend line is imputed. The blue shaded area indicates the period of WWI. The source data are the digitized product-destination level trade statistics.

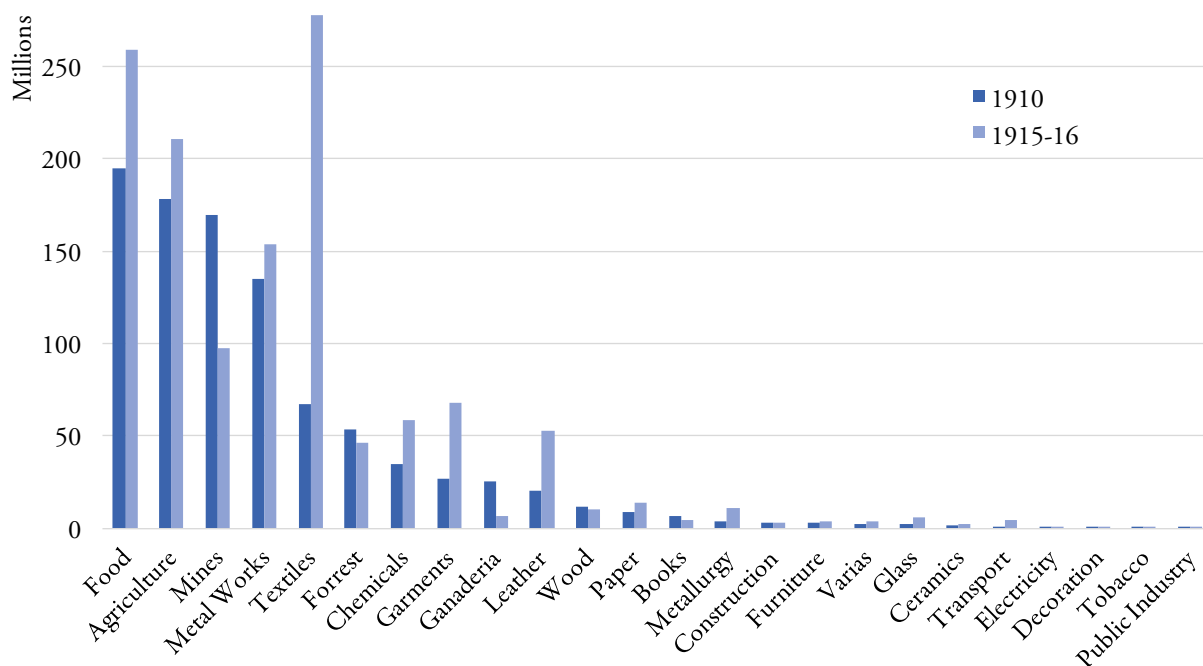
prolonged period of low GDP growth with little structural change (Prados de la Escosura; 2017). Using newly collected trade data on Spanish product level exports between 1910-1919, as well as labor market data on wages and employment across 48 different provinces and 24 different sectors before and after the war, I document five stylized facts about the shock and its impact. Firstly, the trade shock was large, increasing aggregate exports by 40% at constant prices, and additionally the shock was spatially biased with most of the aggregate increase being due to higher volumes of trade with belligerent countries with France being by far the most important destination. Secondly, the trade shock was asymmetric across sectors. Comparing the trade increase between belligerent and non-belligerent countries before and during the war, I find that exports to belligerent countries increased in particular for garments, textiles, paper and products from the heavy industry. Thirdly, sector-level income growth was spatially tilted towards the French border, with each additional 100km distance to the French border decreasing - on average - the growth rate by 4 percent. Fourthly, provinces with a higher specialization in industries favored by the shock experienced faster population growth compared to their pre-trend, with the opposite being true for provinces with less favorable industrial composition. Finally, regional industry dynamics depended on the tightness of the local labor market, indicating an important role for spatial frictions in segregating labor markets and thus preventing arbitrage between geographically segregated labor markets.

The general point that provinces with a prior specialization in sectors that benefited from the war shock had an opportunity to expand their production can be illustrated with an example: Already before the War the *Sociedad Minera y Metalúrgica de Peñarroya* operated a factory for fertilizer and other chemical goods in Cordoba. During the War the factory faced higher wages and input prices, but they also experienced a substantial increase in both domestic and foreign demand, allowing them to expand their output of superphosphate - a fertilization agent - by 30 percent while expanding their workforce by 20 percent between 1914 and 1917 ([Instituto de Reformas Sociales; 1916](#)). Companies like the *Sociedad Minera* make up the individual industries considered in this paper. With their industrial capacity in place they were well positioned to benefit from the shock, but had to attract labor from other provinces and sectors. In doing so, industries found themselves competing with each other to attract workers from the agricultural hinterland. The focus of this paper is to learn more about the labor market conditions and frictions that shaped the response to the shock.

I develop a quantitative economic geography model to understand the *aggregate* impact of that shock, accounting for the *disaggregated* geographical margins of adjustments. The model is consistent with the stylized facts and focuses on taking explicitly into account the spatial linkages in the labor market and the patterns of comparative advantage across provinces, as well as the sectoral switching costs within provinces. I build on the existing quantitative economic geography literature model - as recently surveyed by [Redding and Rossi-Hansberg \(2016\)](#) - and extend a baseline model into several directions. Labor demand is determined by a framework where multiple sectors conduct intra-national and international trade subject to geographical frictions. Differently to most of the commonly used models in the literature, I do not take a stance on the strength of industry level scale economies. Rather, the patterns of comparative advantage across space and sectors are partially endogenous, with higher labor densities translating into productivity gains, depending on the strength of a set of sector specific parameters that determine industry level scale economies. The adopted models - first introduced into the international trade literature by [Kucheryavyi et al. \(2016\)](#) - can be represented by a tractable log linear gravity formulation and is consistent with a Ricardian multi-sectoral trade model with external scale economies, but also nests multiple other canonical models currently used in the literature, depending on the interpretation and values of the parameters.

Labor supply is determined by a nested discrete choice framework where workers first make a decision about reallocating across space subject to incurring switching costs, and then upon arrival in the new province sort into sectors. A two-staged sequence of

Figure 2: Sectoral Export Composition (1910, 1915/1916)



*Notes:* This figure reports the aggregate export composition in sectoral terms. The product level trade has been aggregated to sector level trade data to match the level of aggregation of the labor market panel. The total value of exports for each section in 1910 as well as the mean exports for 1915/1916 is reported. The source data are the digitized product-destination level trade statistics.

preference shocks from a Fréchet distribution make the framework tractable. Two kinds of switching costs are introduced: Firstly, workers who leave a sector incur a switching cost that is specific to the sector and proportional to the expected utility of its destination, secondly, a worker who reallocates to a different province incurs a switching cost that scales with distance. This framework extends the commonly used economic geography models by allowing for stickiness in employment at the sectoral and provincial level - a key feature of the data. At the same time the number of parameters that is being introduced is limited.

I then show how the structure of the model and the exogenous variation due to the natural experiment can be combined to obtain credible estimates for structural parameters that pin down the gains from reallocation. The general intuition for my estimation strategy is that benchmark economic geography models can be inverted to obtain a unique set of province-sector specific market share shifters. This is related to inverting market shares to obtain mean utilities when estimating demand in differentiated product demand markets as recently applied in the trade literature in [Adao et al. \(2017\)](#). These market share shifters are structurally related to prices adjusted for the curvature of the demand function. In a large class of commonly used economic geography models the

responsiveness of this price measure to wages is directly informative about how trade patterns respond to wage changes, and the responsiveness with regard to industry scale is informative about scale effects.

More specifically, my approach can be described as follows: Conditional on specifying the strength of geographical frictions in input markets, the structure of the model together with income data can be used to solve for the origin-specific prices. The strategy behind this is that economic geography models allow to decompose total sectoral income into two parts, a first determinant of income that is due to proximity to lucrative destination markets, and a second part that describes how given market access lower marginal costs translates into a higher captured trade share across all locations, with this part being theoretically interpreted as an origin-specific price and is empirically related to the origin fixed effect in a gravity equation. These origin specific prices can be regressed on (log) wages and (log) employment sizes of sectors to obtain elasticities that describe how changes in wages and sectoral employment translate into higher trade shares and thus higher incomes. The elasticity with regard to wage changes is commonly referred to as *trade elasticity*, while the other elasticity determines scale effects. I will refer to it as *scale elasticity* in the remainder of the text.<sup>2</sup> An obvious problem in this estimation is the endogeneity of wages and labor densities. I utilize instruments that effectively exploit differential shock exposure across provinces interacted with differences in labor market tightness to estimate the parameters. A challenge is that wage and labor changes are correlated, thus *differential* variation is needed to distinguish the independent effect of each variable. Labor market tightness induces variation in the extent to which the shock is being absorbed into wages or employment levels, making it possible to identify the trade and scale elasticity.

The estimated parameters point to the presence of *decreasing* returns to scale in the medium run, effectively limiting the immediate gains from reallocating labor in the absence of the shock. Similar estimates for scale economies over a 10-year horizon indicate that decreasing returns vanish over the long(er) run. The estimation also gives insights into the performances of a broad class of economic geography models in capturing adjustment patterns. Specifically, the fit of the regression expresses how much of the observed variation in residual income shifters can be explained by the endogenous mechanism provided by the model. The model can explain half of that residual variation.

---

<sup>2</sup>Note that this elasticity does not correspond to an output scale elasticity, but rather combines how scale translates into productivity gains which translate into lower prices and thus into higher market shares across all trading partners where the responsiveness of the trade growth depends on the trade elasticity.

For the estimation of labor market frictions the structure of the model is combined with additional data. Usually such an estimation requires flows of workers across space and sectors. However, in a historical context this type of data is rarely available. I show how to estimate labor market frictions in the absence of such data. The structure of the model allows for a conveniently separable estimation of geographical and sectoral frictions. Geographical frictions are being estimated by fitting the model to additional data available in the censuses. The data decomposes the stock of residents along their place of birth and is available in 1920 and 1930. Following [Silvestre \(2005\)](#), comparing the stocks between 1920 and 1930, I can obtain net migration rates between provinces, thus providing implicitly geographical information to estimate the impact of distance on migration flows. The estimation itself is a minimum distance estimation that fits the geographical stage of the labor supply model to the data.

In order to estimate sectoral switching costs I fit the model to changes in labor market conditions at the province-sector level from before to after the war. A key concern is that migration decisions were made during the war based on wage dynamics that are not available. I overcome this data limitation by using the estimated labor demand model together with the trade shock to simulate unobserved wages during the war and estimate sectoral frictions consistent with those wages. As has been pointed out by [Silvestre \(2005\)](#), levels of internal migration during that period were markedly low, amounting to decennial net flows of less than 5 percent out of the population. Consistent with that, the estimated model indicates high frictions to labor mobility across sectors and in particular across space, implying similarly low levels of migration with less than 3 percent moving over the 6 year period that is being considered.

As a result of the estimation I obtain simulated reallocation patterns of labor that are consistent with the changes in labor market conditions due to the war. The implied reallocation patterns strongly suggest that spatial frictions dominate sectoral adjustment frictions, with 83 percent of the adjustment happening across sectors *within* provinces, rather than *between* provinces. Finally, I use the estimated model to obtain the counterfactual evolution of the Spanish economy in the absence of the World War I shock. This exercise shows that the War increased the overall size of the manufacturing sector by 13 percent, while shifting the national industry composition towards more advanced industrial sectors such as chemicals, metal works, textiles and garments. The model can also be used to calculate changes in nominal income in the counterfactual. As suggested above, during the War Spain experienced a differential growth pattern between northern provinces (defined as above Madrid in terms of proximity to France), and southern provinces. Northern provinces experienced around 30 percent larger (nominal) income

growth than southern provinces. The counterfactual without the War indicates only a minimal spatial gradient of 4 percent, residual productivity trends can explain a further 15 percent with the remaining 11 percent being attributed directly to the War. Since the model only allows a parametrically limited channel this can be understood as a lower bound for the effect of the War shock on spatial inequality.

The current study implies a tentative explanation for the lack of development prior to the shock. The presence of decreasing returns in the short run suggest that even if labor reallocation took place it would not generate a fortuitous circle of productivity gains, higher wages and further reallocation. Rather sectoral productivity would be on average decreased as a result of employment growth and only recover in the medium run. Such dynamics in productivity gains would inhibit structural change, especially when combined with a high level of labor market frictions. If furthermore workers reside in low productivity sectors in provinces that are distant from the provinces that feature highly productive sectors then the high level of *spatial* labor market frictions is particularly prohibitive. In the Spanish case, the analysis seems to suggest that the pre-shock wage differential between the industrializing North and the agricultural South was insufficient to surmount spatial labor market frictions. An additional obstacle to reallocation might be present if workers do not respond effectively to individual sectors' wage dynamics but rather make migration decisions based on the general appeal of provinces as a whole - an approach that is implicit in the two stage labor supply model formulated in the current study and is consistent with the data.

This interaction between decreasing returns and labor market frictions can actually be beneficial in the presence of the shock. Labor market frictions effectively lock in labor in new sectors until the decreasing returns vanish, inducing a delayed industrial dynamism as a response to the shock that can be related to the economic take off observed in Spain in the 1920s, long after the shock of the War had faded away.

The current setting has three distinct advantages that make the analysis possible. Firstly, the shock is large as well as spatially and sectorally asymmetric and plausibly exogenous towards prior industrialization patterns in Spain. This provides a large amount of independent variation that allows to identify the parameters. Secondly, there is prior substantial variation in sectoral specialization across cities, allowing for an uneven impact of the shock across space. Finally, the policy response was limited. During the War the central government in Madrid was dominated by the land-based oligarchy, who took little interest in the economic needs of the business community in Catalonia or the Basque country (Harrison; 1978). Policy remedies only came late and in a limited



form.<sup>3</sup>

**Related Literature** This paper contributes to a growing literature that looks at how industries and regions within countries might respond to an external shock. What sets the current paper apart is that it looks at a natural experiment that affected the whole economy while also accounting for the sectoral dynamics. As such it is a convenient setting to shift the focus towards the effect of labor market frictions and scale economies as well as their interaction. In doing so, this paper brings together different aspects that have been looked at separately before. One of these aspects is the endogenous productivity response to trade changes and how that in turn affects industry dynamics as in [Juhasz \(2017\)](#). She studies how temporary import protection can induce import substitution and productivity improvements in the textile industry during the Napoleonic blockade. In the current setting the model allows for endogenous productivity responses in a more abstract way as a function of the observable scale of employment.

Another aspect is how trade shocks can fuel differential population dynamics between cities and provinces creating persistent differences. This has been explored before by [Hanlon \(2014\)](#) in the case of a negative supply shock caused by the U.S. Civil War (1861-1865) which dramatically reduced the cotton imports to the English cotton textile industry. This differentially affected cities that were more specialized in that industry compared to those that were not. The same effect is present in the current setting, but crucially the data in combination with the structure allows us to examine the labor market interactions across multiple sectors and provinces, granting valuable insights on how labor market frictions shape regional dynamics as a response to the shock. Furthermore, as suggested above, the combination of labor market frictions and productivity dynamics suggest an interesting perspective on delayed and to some degree persistent effects of external shocks.

Finally, some of the findings in this paper are reminiscent of a study conducted by [Dix-Carneiro and Kovak \(2017\)](#). They examined Brazil's regional dynamics as a result of trade reforms and trade liberalization and find slow adjustment and steadily increasing divergent trade effects driven by a mechanism where high labor market frictions and slow capital accumulation drive the adjustment pattern. Their empirical and theoretical setting is very different: While they focus on a permanent change in the trade envi-

---

<sup>3</sup>For example only in 1917 did the Spanish government introduce a *law for the protection of new industries and the extension of existing ones* earmarking 10 million pesetas for the use of industries falling far short of the demand of the industry lobby to establish a foreign exchange bank and a commission house to facilitate the financing of exports instead.

ronment of a country, I focus on how a temporary shock can reallocate inputs across provinces and sectors and the analysis is more focused on counterfactuals employing an extended quantitative spatial equilibrium model. However, some of their results are reflected in the current paper, such as the prolonged effect of the shock as well as the limited labor mobility across space.

Secondly, the paper adds to the quantitative economic geography literature as recently surveyed by [Redding and Rossi-Hansberg \(2016\)](#). I contribute by showing how to adapt a baseline economic geography paper to examine reallocation patterns of labor by accounting for several key aspects of the data. Firstly, the model is more flexible with regard to the presence of productivity returns to scale, which are important in determining the gains from reallocating labor. Secondly, the proposed model manages to match the observed persistence of employment at the province and sector level despite large wage differentials, by introducing sectoral and spatial labor market frictions, making it possible to compute unobserved patterns of reallocation and allowing to distinguish within and between provincial adjustments. A key underlying theme of the current work is how to combine sufficient structure to augment the paucity of the historical data. A constraint is that a model with rich spatial interactions usually requires flow data to infer the structural parameters and to disentangle different labor market frictions. I demonstrate how to leverage the structure of the model to estimate its parameters, relying on a separate treatment of labor demand and labor supply as well as a convenient separation of spatial and sectoral labor market frictions.

Finally, the paper adds to the literature on Spanish economic history by showing that the WWI shock had an important impact on the Spanish economy, not necessarily by creating large output and productivity gains directly, but by reallocating factors across space and sectors to provide the preconditions for an economic take-off in the 1920s. As such it is a middle ground between the two opposing views in the literature. The established view, represented by [Roldan and Delgado \(1973\)](#), interprets the war as a large turning point for economic development. Using his own constructed GDP series, [Prados de la Escosura \(2016\)](#) emphasises that the World War shock actually decreased GDP per head and instead he points towards the 1920s as a much more important decade for Spain's development. My analysis implicitly connects the two events by pointing towards the reallocation of labor across sectors as a fertile ground for capital fueled growth in the 1920s.

The remainder of the paper is structured as follows. Section 2 discusses the historical background, describing both the situation in Spain before the War and during the War.

Section 3 describes the various data sources as well as the construction of the labor market panel that underlies most of the analysis. Section 4 gives reduced form evidence on the trade shock and its effect on regional population dynamics. Section 5 describes the theoretical model that guides the estimation and analysis. Section 6 then proceeds with describing the estimation procedure. In Section 7 I then use the quantitative model to simulate Spain in the absence of the War before discussing the results. Finally, in section 8 I conclude.

## 2 Historical Background

This section describes the historical circumstances. The first part gives an overview of the state of the Spanish economy towards the beginning of the war. The second section gives an overview of the historical circumstances of the World War itself and how Spain itself was connected to it.

### 2.1 Spanish Economy at the beginning of the 20th Century

After missing the first wave of the industrial revolution in the first half of the 19th century (Harrison; 1978), the Spanish economy underwent a period of rapid industrialization in the second half of the 19th century, fueled by market integration due to the expansion of the railroad network which in turn resulted in the devolution of industrial capacity to the peripheral provinces with the cotton industry in Catalonia and Metallurgy in the Basque country developing especially rapidly (Nadal; 1975). However, industrialization soon came to an early halt with the census data showing little increase in industrial employment from 1887 onwards as can be seen in figure 6. This is also mirrored by very low GDP per head growth rates averaging 0.6 percent between 1883-1913 (Prados de la Escosura; 2017). Some authors attribute the low levels of growth to limited demand for manufacturing goods domestically as well as little capacity to compete with goods from countries such as Germany, France and the UK that are more advanced in terms of their industrialization (Harrison; 1978).

As a result, at the beginning of the 20th century, the industrial sector barely continued to expand and Spain remained at a low level of industrial development. According to census data, in 1900 roughly 70% of the working population worked in agriculture and only 12.5% worked in industrial/manufacturing sectors. Industrialization only proceeded slowly, with the industrial sector only growing marginally in total employment

by 3%, adding a little bit less than 40,000 jobs nation-wide in the first decade of the century. At that time, the largest share of the industrial sector was made up by sectors associated with primary goods, such as the exploitation of mines or the production of construction material.

In terms of the spatial distribution of the population, most of the population was still concentrated in predominantly rural and agricultural areas such as Andalusia<sup>4</sup> or Castilla y Leon<sup>5</sup>. However, looking beyond the larger regional aggregation and looking at individual provinces, it is precisely such major urban centers such as Oviedo, Valencia, Bilbao, Madrid and Barcelona that increasingly attracted and concentrated the Spanish population. The provinces that contained these urban centers tended to concentrate most of the industrial activity as can be seen by the map in figure 7 indicating the spatial distribution of manufacturing employment. While internal migration was perennially low, with net migration amounting to less than 5% of the population before 1920, the two largest cities, Barcelona and Madrid, tended to nevertheless attract a large share of agricultural workers from other provinces, making them unique magnets for migrants around 1900 (Silvestre et al.; 2015).

The industrial structure of those urban centers was heterogenous. For example, Barcelona was highly specialized in the cotton textile industry, while Valencia specialized in garments. Because of natural endowments mining and associated downstream industries dominated in Oviedo and Jaen. The Basque country had an early advantage in the heavy metal industries, featuring numerous Martin-Siemens open hearth furnaces for steel production as well as other installations. This degree of agglomeration of specific industries even at this early stage of industrialisation suggests some degree of agglomeration externalities.

In terms of external markets, at the end of the 19th century, (former) colonies and other Latin American markets played a particularly important role, while after the loss of the colonies Spain's exports shifted more towards European countries with France and Great Britain taking up the biggest share of exports. Most of the exports were raw materials or agricultural products consistent with the low developmental status of Spain at the time as depicted in figure 2. In general Spain ran a trade deficit for most of the beginning of the 20th century except for the short period under consideration in this paper.

---

<sup>4</sup>Andalusia comprises eight provinces: Almería, Cádiz, Córdoba, Granada, Huelva, Jaén, Málaga and Seville, with major industrial activity located in Seville and Mining employment in Huelva

<sup>5</sup>Castilla y Leon comprises nine provinces: Avila, Burgos, Leon, Palencia, Salamanca, Segovia, Soria, Valladolid and Zamora with major industrial activity centered in Valladolid.

In summary, it can be stated that at the beginning of the 20th century Spain was a predominantly agricultural economy with a low level of industrial activity and while there was some rural urban migration, there was in general little dynamism towards further industrialization.

## 2.2 The Spanish Economy and World War I

The assassination of the Austrian Archduke Franz Ferdinand on 28 June 1914 by Yugoslav revolutionaries, triggered a series of declarations of Wars that set off the first World War on 28 July 1914, with the allied powers spearheaded by France, the British Empire, Russia, and later on the United States, fighting the central powers, composed of the German Empire, Austria-Hungary, the Ottoman Empire and other co-belligerents. The consensus is that a conflict limited in terms of duration and extent was expected, but instead it would become one of the largest wars in history, spreading across all major populated continents and lasting until 11 November 1918.

At the onset of the war Spanish society was divided into two opposing camps, with liberal fractions supporting the allied powers, and the remainder of the population supporting the central powers. However, a participation in the war itself was not considered feasible (Harrison; 1978), so Spain remained neutral throughout the war.

The effects of the first World War on the Spanish economy are well documented in the reports by the *Instituto de Reformas Sociales* (Instituto de Reformas Sociales; 1916). They can be broadly summarized into two categories. Firstly, the war brought about opportunities to provide war materials to the belligerent nations. This spawned increased demand for textiles, garments, and for the heavy metal industry. Secondly, a lack of British, French and German competition in the home market provided an opportunity for domestic producers to produce import substitutes. The report mentions new factories that produced goods as varied as supplies for cars, paper folders, perfumes, small machinery, lightbulbs and others. I will examine the effect in more detail in the reduced form section below.

## 3 Data

**Labor Market Panel Data** The main source for labor market data is an industry survey that covers the years 1914, 1920, 1925 (Ministerio de Trabajo; 1927). This industry

survey was published by the Ministry for Labor and Industry and is based on surveys conducted at all public firms and large private enterprises in cities that are larger than 20,000 inhabitants (Casanovas 2004). It covers 23 different industries<sup>6</sup> and 48 different provinces.<sup>7</sup>

While the industry survey covers a large range of the manufacturing sector, it does not give further information on the remaining economy. As mentioned before, a crucial feature of the Spanish economy was a large agricultural sector. In order to account for that, I digitized the occupation-province specific section of the census for 1900, 1910, 1920 and 1930. I use the 1920 data on agricultural employment<sup>8</sup> to augment the 1920 data.<sup>9</sup> For the 1914 data, I use the 1910 province specific agricultural employment data and extrapolate by calculating province specific fertility trends until 1914. Finally, I use data contained in the official Spanish statistical yearbooks on province specific agricultural mean wages for 1915 and 1920.

**Trade Data** The trade data is taken from annual trade records released by the Spanish custom agency. Using crowdsourcing services, I digitized the trade statistics for the years 1910 and 1914-1919. For those years, the quantity of exports in 383 product categories across 77 different destination countries are available. Furthermore, the border agency uses a system of product level prices to obtain total export values. These prices do not vary throughout the period of consideration and can be interpreted to give the relative pre-war prices across goods.

---

<sup>6</sup>The industries included are called: Books, Ceramics, Chemicals, Construction, Decoration, Electricity, Food, Forrest, Furniture, Garments, Glass, Leather, Metal Works, Metallurgy, Mines, Paper, Public, Public Industry, Textiles, Tobacco, Transport, Varias, Wood.

<sup>7</sup>The census for 1910 lists 49 different provinces. They mostly correspond to the modern administrative units called *provincias* - provinces - which are in turn roughly the NUTS3 level administrative units of Spain. There are some minor differences, e.g. in how different off-continental administrative units are being treated. For my analysis I drop the Canary islands from the sample since their distance from the mainland makes it hard to argue that they are similarly integrated as other provinces.

<sup>8</sup>More specifically I add the Agriculture (Owner) section and the workers in fishery, forest and agriculture together to obtain an aggregate size of the agricultural sector at the time in each province

<sup>9</sup>When merging the census data with the industry survey, I adjust for the fact that the survey does not cover the universe of workers, while the census does. In order to maintain the correct relative size of agriculture to manufacturing sector, I compare the total size of industry employment in the survey data with the census - with the census potentially accounting for informal employment as well as industries in smaller villages. On average, the manufacturing employment size of the survey data represents at least 44% of the manufacturing sector in the census data. I scale the agricultural employment accordingly when merging the census and survey data.

**Correspondence** In order to construct a correspondence between product-level trade data and industry-level labor market data, I used an additional publication that lists the official correspondence between industries and occupations ([Instituto Nacional de Pre-  
vision Social; 1930](#)), often explicitly stating the associated product as occupation name for an industry. From that I constructed a correspondence table that matches products to industries.<sup>10</sup> While some products can be uniquely associated to one industry, others can be at least matched with two industries. In matching exports to industry levels, I add the export values for those products to both relevant industries.

**Migration Data** In order to infer labor mobility costs, data on migration flows is necessary. I follow [Silvestre \(2005\)](#) and use the province level data on inhabitants that are *Born in Another Province* which is contained in the censuses. For 1920 and 1930 additional information is available listing not only the stock of migrants which were born in another province, but their origin province as well. The difference between 1930 and 1920 in the stock of migrants - adjusted for decennial survivability rates - is informative about net migration. In order to construct net migration, I follow [Silvestre \(2005\)](#) and use the decennial census survivability rate between 1921-1930,  $S \equiv 0.86$ . Net internal migration can be obtained by constructing the survivability adjusted change in stock of migrants, i.e.

$$\text{Internal migrations}_{1930,1920,i,j} = BAP_{i,j,1930} - S \times BAP_{i,j,1920}$$

where  $BAP_{i,j,1920}$  refers to the stock of residents in  $i$  who were born in province  $j$  in 1920.

**Distance** Using GIS software, I georeferenced the Spanish railroad network in 1920. Then, using MATLAB's internal shortest path function, I obtain bilateral distances between provincial capitals along the shortest path of the railroad network. In order to obtain distances to Paris, I augmented the graph with the French railroad network and further added maritime linkages between important ports in France and Spain. Again using the shortest path functionality of MATLAB I can obtain the shortest distance along this transportation network between provincial capitals in Spain and Paris.

---

<sup>10</sup>The correspondence table is available upon request.

**Housing** The housing expenditure share as well as stock and rental rates can be imputed combining different data sources. The statistical yearbooks make available the number of buildings available in a province as well as the inhabitants and thus the effective occupancy rate, the inverse of which is the share of a building that is rented by an average resident. Additionally, average yearly rental expenditure is selectively available across provinces in the Boletins of the *Instituto de Reformas Sociales*. This yearly rate can be adjusted towards an hourly rate in a province,  $r_i$ . Total expenditure on housing can be imputed by firstly multiplying the rental rate and the inverse of the occupancy rate - call this the unit rental rate - with the stock of housing. Calculating total expenditure on housing as a share of total labor income across all provinces defines the expenditure share on housing, which I will refer to as  $\delta$ .

## 4 Reduced Form Evidence

In this section I develop five stylized fact that characterize the nature of shock, as well as the impact it had on regional development within Spain. The stylized facts will guide the choice of the model and will inform the empirical estimation.

**Stylized Fact 1: The Trade Shock was large & spatially biased** The export shock was large from an aggregate point of view. In 1915 aggregate exports increased by 40% compared to 1914 and stayed at a high level for as long as the war lasted.<sup>11</sup> Most of the increase was due to differential increase of belligerent countries compared to non-belligerent countries as shown in figure 1: The trade to belligerent countries tripled, while trade with non-belligerent countries remained at a relatively low level and only grew in the later war years above pre-war levels. Most of the increase in trade with belligerent country stems solely from export increases to France. Since the trade shock originated mostly from France, provinces close to the French border had a more favorable position since they facing lower transport costs when shipping towards France. If transport costs matter, then the fact that most of the increase was due to France implies a spatial bias in the trade shock.

---

<sup>11</sup>This increase is probably underestimated since official statistics kept the price for the calculation of values of exported goods at a constant level during the decade under consideration, while it is plausible that increased demand has further increased the price.



**Stylized Fact 2: The Trade Shock was asymmetric across Sectors** Most of the increased demand can be associated with war needs, such as Textiles, Garments, Metal Works and Leather goods which is evident in the shift in the sectoral composition of exports from Spain to France. However, it is not clear whether these changes in sectoral trade flows are driven by plausibly exogenous demand side effects or by potentially endogenous domestic supply side trends. In order to obtain a sector specific measure of the foreign demand shock, I construct a theoretically consistent measure by leveraging a standard gravity trade equation,

$$X_{od,s,t} = \tau_{od}^{-\epsilon_s} w_{o,s,t}^{-\epsilon_s} A_{o,s,t}^{\epsilon_s} P_{d,s,t}^{\epsilon_s} E_{d,s,t}$$

where  $X_{od,s,t}$  denotes the export level from origin ( $o$ ) to destination ( $d$ ) in sector  $s$  which depends on bilateral resistance term,  $\tau_{od}$ , as well as the marginal cost of production in the origin country,  $w_{o,s,t}/A_{o,s,t}$ , positively on the sectoral expenditure in the destination country,  $E_{d,s,t}$  and the price index,  $P_{d,s,t}$ , measuring the competitiveness in the destination market, and where  $\epsilon_s$  denotes the sector specific trade elasticity. Constructing the growth of exports,  $\hat{X}_{od,s,t} \equiv \frac{X_{od,s,t}}{X_{od,s,t-1}}$ , and comparing the growth rate across destination countries, one can obtain the following expression,

$$\Delta_{o,s,t} \equiv \frac{\hat{X}_{od,s,t}}{\hat{X}_{od',s,t}} = \left( \frac{\hat{P}_{d,s,t}}{\hat{P}_{d',s,t}} \right)^{\epsilon_s} \times \left( \frac{\hat{E}_{d,s,t}}{\hat{E}_{d',s,t}} \right)$$

where hat variables refer to changes. In words, this double difference states that export growth from origin  $o$  to destination  $d$  compared to export growth from  $o$  to some other destination  $d'$ ,  $\hat{X}_{od,s,t}/\hat{X}_{od',s,t}$  is a function of relative changes in the price index in the two destination countries,  $\hat{P}_{d,s,t}/\hat{P}_{d',s,t}$ , as well as relative growth in their expenditure levels  $\hat{E}_{d,s,t}/\hat{E}_{d',s,t}$ . This double difference can be used to isolate destination specific effects, in particular, the relative changes in the expenditure and competitiveness of one destination market compared to some other, plausibly unaffected, comparison group.

When calculating this measure for the WWI shock, I compare sectoral export growth to belligerent countries to non-belligerent countries. However, some adjustments are necessary to account for secondary effects of the war. First of all, the war made trade across the frontline and maritime trade after 1917 difficult. Therefore the sample of belligerent countries that I focus on only includes France, Italy and the United Kingdom and I construct export growth by comparing the mean export levels for 1915/1916 with the baseline export in 1910, thus avoiding additional distortions after 1916 and the

partial-year war effect of 1914. For the non-belligerent comparison group I exclude belligerent countries as well as the United States, to avoid any war preparations to pollute the measure. The sectoral results can be seen in the appendix in figure 10. The sectors that benefited from particularly high levels of demand during the war are Garments, Glass, Metal Works, Mines, Paper and Textiles. These sectors experienced between 5-20 times more growth from belligerent countries than they did from non-belligerent countries.<sup>12</sup>

**Stylized Fact 3: Regional Dynamics exhibited a Spatial Gradient** The shock induced a demand shock that had spatial and sectoral characteristics, but how did the shock affect regional dynamics? I use the labor market data introduced in the previous section to construct income growth at the sector-province level. In order to examine whether the spatially biased shock induced regional development that was spatially tilted, I run the following regression,

$$\frac{Y_{i,s,1920}}{Y_{i,s,1914}} = \alpha + \beta_1 distance_{i,Paris} + \epsilon_{i,s}$$

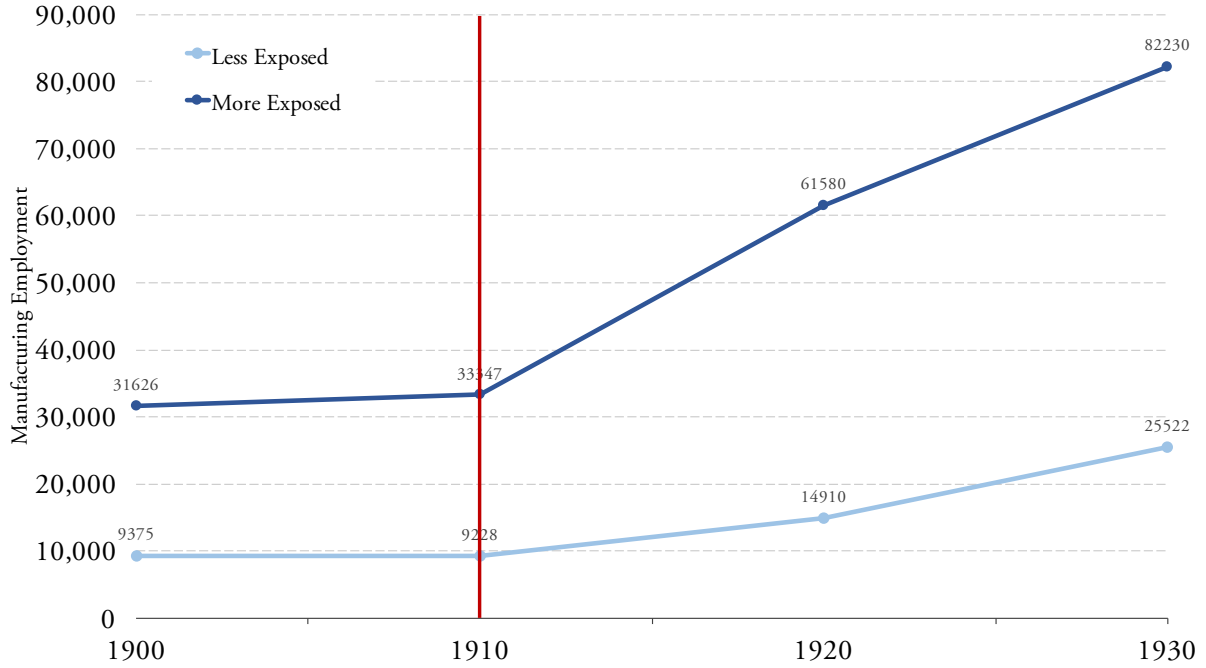
where  $Y_{i,s,1920}$  is the total labor income of sector  $s$  in province  $i$  in 1920, that is  $Y_{i,s,1920} = w_{i,s,1920}L_{i,s,1920}$  with  $w_{i,s,1920}$  referring to the wage in that province-sector and  $L_{i,s,1920}$  referring to the total number of employees, and finally  $distance_{i,Paris}$  refers to the shortest distance along the railroad network or maritime linkages between the capital of province  $i$  and Paris. The fitted line is depicted in figure 4. I find that each additional 100km distance to Paris translates into 4 percent lower income growth. This stylized fact is also robust at the sectoral level and controlling for labor market tightness - as proxied by the own sector size relative to the province size - as well as initial differences in comparative advantage - as proxied by the sectoral employment share in the national industry - as can be seen in regression table 7.

**Stylized Fact 4: Regional Dynamics & Industrial Capacity** To understand the differential impact that the shock had at the province level, I use the sectoral shocks to construct an exposure measure to the shock, i.e.

---

<sup>12</sup>As can be seen in the table Mining exports to non-belligerent countries all but disappeared in the period under considerations. According to the historical reports, this is not due to demand factors, but capacity constraints in Spain, a feature that is not inherent in the standard gravity approach.

Figure 3: Manufacturing Employment Growth and Shock Exposure



Notes: This figure shows the evolution of average manufacturing employment of most and least exposed provinces. Most and least exposed provinces are defined as above or below the median value for the exposure index defined below. The red line indicates the observation after which the WWI shock (1914-1918) is taking place. The data is taken from the population censi 1900-1930 and export statistics.

$$E_i \equiv \sum_s \frac{L_{i,s,1914}}{\sum_j L_{j,s,1914}} \times (g_{Spain,Bel,s} - g_{Spain,Non-Bel,s}) \times X_{Spain,France,s,1914}$$

where  $g_{Spain,Non-Bel,s}$  and  $g_{Spain,Bel,s}$  refers to the growth rate of exports to non-belligerent and belligerent countries respectively as calculated above. The difference is the excess growth in sector  $s$  associated with WWI. The exposure term summarizes therefore at the provincial level the expected incidence of the trade shock given the pre-existing industrial capacity within a sector proxied by the employment share in the *national* sectoral employment and given the estimated increase of French exports due to the WWI shock. In order to examine the impact of this exposure measure on regional dynamics and in order to illustrate pre trends I rely on additional data from the Spanish population censi on manufacturing employment. In order to analyze the responsiveness to the continuous exposure variable I examine a continuous treatment Diff-in-Diff specification in the spirit of [Acemoglu et al. \(2004\)](#), i.e.

$$\ln y_{it} = \delta_i + \delta_t + \gamma_1 \times d_{1920} + (\gamma_2 + \varphi \times d_{1920}) \times \ln dist_{i,Paris} + (\varphi_1 \times d_{1920} + \varphi_2 \times \delta_t) \times \ln E_i + \epsilon_{i,t}$$

The left hand side variable,  $y_{it}$  is manufacturing employment in province  $i$  at time  $t$ , where manufacturing employment is available in 1900, 1910, and 1920,  $\delta_i$  refers to the full set of province specific fixed effects,  $d_{1920}$  is a dummy for the 1920 which is the first observation after the WWI shock,

Figure 3 illustrates the results and the regression results are reported in table ???. The coefficient of interest is  $\varphi_1$  which is the responsiveness of manufacturing employment towards increases in the exposure measure - which measures a province's ability to exploit sector specific shocks given the scale of its prior industrial capacity in the affected industries. Comparing the 10th to the 90th percentile this gives an estimated effect of  $(8.65 - 5.89) * \varphi = 0.19$  log points. The regression as well as the figure above point towards parallel trends prior to the shock, as can be seen by the coefficient  $\varphi_2$  which is not significantly different from zero.

**Stylized Fact 5: Local Labor Supply can inhibit Regional Dynamics** In the presence of spatial labor market frictions, which would be consistent with the low level of decennial internal migration at the time in Spain (Silvestre; 2005), labor supply is partially localized and must be sourced from other sectors within the same province. This implies that the larger an industry's share in the local labor market the more limited the pool of workers it can source from. Regressing (nominal) income growth on the sectoral share of total provincial employment before the war which is defined as Employment Share of Sector in Province)  $\equiv \frac{L_{i,s,1914}}{\sum_r L_{i,r,1914}}$ , I find that an increase by 1 log point translates into .1 log points lower nominal growth rates. The linear fitted line can be seen in figure 8. This finding is robust to controlling for comparative advantage as proxied by the size of the province-sector in the national industry, and level size affect as proxied for by (log) employment in 1914 of that industry as can be seen in table 7.

## 5 Theoretical Framework

The theoretical framework is informed by the stylized facts shown above. As indicated by the spatial gradient, spatial frictions in the output and input market will play a

prominent role, thus shifting the attention towards economic geography models. Furthermore, the setting requires a multi-sectoral model to account for the sectoral heterogeneity of the shock. Finally, the last two stylized facts suggest that provinces compete for labor inputs and that labor supply can be - to some extent - localized. In order to accommodate that, I will extend the standard economic geography model to account for a fairly general set of labor market frictions, introducing switching costs that make labor sticky at the provincial and the sectoral level.

## 5.1 Setting

Consider an economy with a fixed number of  $I$  locations indexed by  $i, j, k \in \mathcal{N}$ . Locations are heterogeneous in their exogenously fixed housing supply,  $H_i$ , and their geographical location relative to one another. Each location produces goods in  $S$  sectors  $r, s \in \mathcal{S}$ . There are only two periods and the initial distribution at time 0 of the population across locations and sectors,  $[L_{i,s,0}]_{\forall i,s}$ , is given.

## 5.2 Labor Demand

Labor demand is being determined by a multi-sector Ricardian model with industry level economies of scale along the lines of [Kucheryavy et al. \(2016\)](#), that allows for intranational trade between provinces within a country and international trade with foreign countries. The only factor of production is labor. Each country has a representative consumer with upper tier Cobb Douglas preferences across housing - with an expenditure share  $\delta$  - and industry bundles, with industry specific expenditure shares given by  $\beta_r \in (0, 1)$ , such that  $\sum_r \beta_r = 1 - \delta$ . Trade costs are of the standard iceberg type implying that delivering a unit of any good in industry  $s$  from province  $i$  to province  $j$  requires shipping  $\tau_{ij,s} \geq 1$  units of the good. Trade shares take on the following functional form,

$$\lambda_{ij,s,t}(\mathbf{w}_{s,t}, \mathbb{L}_{s,t}) = \frac{S_{i,s,t} L_{i,s,t}^{\alpha_s} (w_{i,s,t} \tau_{ij,s})^{-\epsilon_s}}{\sum_k S_{k,s,t} L_{k,s,t}^{\alpha_s} (w_{k,s,t} \tau_{kj,s})^{-\epsilon_s}}$$

where  $\mathbf{w}_s$  and  $\mathbb{L}_s$  refers to the vector of sectoral wages and employment levels across provinces respectively,  $S_{i,s,t}$  is a province-sector specific productivity shifter,  $w_{i,s,t}$  are the province-sector specific wages, and  $L_{i,s,t}$  the quantity of labor employed, and  $\tau_{ij,s}$

refers to the iceberg trade cost as defined above. Higher labor densities increase productivity via the parameter  $\psi_s$ , which in turn increases trade shares mitigated via the *trade elasticity*,  $\epsilon_s$ , formally being defined as  $\epsilon_s \equiv -\frac{\partial \ln(\lambda_{ij,s}/\lambda_{ii,s})}{\partial \ln \tau_{ij,s}}$ . Together the effect can be summarised as  $\alpha_s \equiv \psi_s \times \epsilon_s$  which is the elasticity of changes in trade flows as a response to changes in employment size of a sector, which I refer to as *scale elasticity*. Finally, the trade elasticity  $\epsilon_s$  also governs the sensitivity of trade flows with regard to changes in the destination specific marginal cost pricing, in particular if they are driven by changes in the input cost, that is the local wage,  $w_{i,s,t}$ .

The current framework, which allows for industry level economies of scale, is consistent with a Ricardian model with external scale economies but is also sufficiently general to nest multiple other trade models including trade models that feature internal scale economies as pointed out by [Kucheryavyy et al. \(2016\)](#).<sup>13</sup> Within a given period the labor allocation is fixed. The static equilibrium can be defined as follows,

**Definition 1** (Static Equilibrium). *The static equilibrium within a period  $t$ , given the labor distribution, is given by goods market clearing, balanced trade and housing market clearing.*

$$w_{is,t}L_{is,t} = \sum_j \lambda_{ij,s}\beta_s Y_{j,t} \quad \forall (s,i) \in \mathcal{S} \times \mathcal{N} \quad (1)$$

$$E_{is,t} = \sum_j \lambda_{ji,s}\beta_s Y_{i,t} \quad \forall (s,i) \in \mathcal{S} \times \mathcal{N} \quad (2)$$

$$r_{i,t} = \frac{\delta Y_{i,t}}{H_{i,t}} \quad \forall (i) \in \mathcal{N} \quad (3)$$

### 5.3 Labor Supply

The initial allocation of households across locations and sectors,  $[L_{i,s,0}]_{\forall i,s}$ , is given. Between the first period -  $t = 0$  - and the second period -  $t = 1$  - Households can make a decision to move across provinces and sectors. The moving decision is based on a nested discrete choice, where workers first decide which province to move to - and implicitly to leave their own sector - and then upon arrival in the province decide which

---

<sup>13</sup>As [Kucheryavyy et al. \(2016\)](#) show, the framework can map into multi-sector variants of [Eaton and Kortum \(2002\)](#), [Krugman \(1979\)](#) and Melitz-Pareto type trade models ([Chaney; 2008](#))

sector to work in. Indirect utility is given by

$$V_{j,s,t} = \left( \frac{\rho_j w_{j,s,t}}{P_{j,t}^{1-\delta} r_{j,t}^\delta} \right) \times \kappa_{j,t} \times l_{s,t}$$

where  $\rho_j$  represent location specific amenities,  $r_{j,t}$  the market clearing rental rate for housing,  $P_{j,t}$  represents a local price index which aggregates sector level local price indices according to the Cobb Douglas preferences specified above, that is  $P_{i,t} = \tilde{\beta}_n \prod_s P_{i,s,t}^{\beta_s}$  and where the sectoral price index  $P_{i,s,t}$  is defined as follows,

$$P_{i,s,t} = \mu_{i,s} \left( \sum_{i \in \mathcal{N}} S_{i,s,t} L_{i,s,t}^{\alpha_s} (w_{i,s,t} \tau_{ij,s})^{-\epsilon_s} \right)^{-1/\epsilon_s}$$

where  $\mu_{i,s}$  and  $\tilde{\beta}_i$  are some constants, where  $S_{i,s,t}$  is a province-sector specific productivity shifter,  $w_{i,s,t}$  is the province-sector specific wage, and  $L_{i,s,t}$  the quantity of labor employed, and  $\tau_{ij,s}$  refers to the bilateral iceberg trade cost,  $\alpha_s$  is the scale elasticity and  $\epsilon_s$  the trade elasticity.<sup>14</sup> Finally,  $\kappa_{j,t}$  and  $l_{s,t}$  represent idiosyncratic preference shocks that capture preference heterogeneity at the micro-level. I adopt the assumption that they are Fréchet distributed.

**Assumption 1.** *The preference shocks are sequentially drawn and identically and independently distributed across provinces and sectors according to a Fréchet distribution with respective dispersion parameters  $\nu$  and  $\gamma$*

$$F(\kappa_{j,t}) = e^{-\kappa_{j,t}^{-\nu}}, \quad \nu > 1, \quad F(l_{s,t}) = e^{-l_{s,t}^{-\gamma}}, \quad \gamma > 1$$

Assumption 1 allows for convenient closed form solutions of the shares of workers across sectors and space.  $\nu$  and  $\gamma$  are the respective dispersion parameters and which will be shown to pin down the responsiveness of migration flows to changes in indirect utility. A household which in period  $t$  is residing in province  $i$  and working in sector  $s$  faces the following problem,

---

<sup>14</sup>The constant  $\mu_{n,k}$  depends on the specific model being adopted.  $\tilde{\beta}_n$  is the standard Cobb Douglas term

$$\max \left[ \underbrace{\mathbb{E}_t V_{i,s,t+1}}_{\text{Remain}}, \underbrace{\mathbb{E}_t \frac{V_{1,s,t+1}}{\mu_{i1}}, \dots, \mathbb{E}_t \frac{V_{I,s,t+1}}{\mu_{iI}}}_{\text{Remain Sec/Change Prov}}, \underbrace{\mathbb{E}_t \frac{V_{i,1,t+1}}{\mu_s}, \dots, \mathbb{E}_t \frac{V_{i,S,t+1}}{\mu_s}}_{\text{Change Sec/Remain Prov}}, \underbrace{\mathbb{E}_t \frac{V_{1,1,t+1}}{\mu_s \mu_{i1}}, \dots, \mathbb{E}_t \frac{V_{I,S,t+1}}{\mu_s \mu_{iI}}}_{\text{Change Prov \& Sec}} \right]$$

that is she can decide to remain in the current sector and in the current province, change just the sector, change just the province or change both. Effectively, the worker compares the indirect utility of remaining in the current province with the expected indirect utility of reallocating to any other province subject to incurring a switching cost, where  $\mu_{ij}$  and  $\mu_s$  are the geographical and sector specific switching costs that capture the difficulty of switching sectors and provinces.<sup>15</sup> The population that remains in a province is pinned down by the geographical mobility cost  $\mu_{ij}$  effectively discounting options that involve out migration. Similarly, the population that remains in a sector is pinned down by the sectoral mobility cost  $\mu_s$  effectively discounting options that involve sectoral switching.

I assume that this decision problem is being done sequentially with the worker first observing the location specific preference shocks,  $\kappa_t$ , but not yet knowing the vector of sector specific preference shocks,  $\iota_t$ . In the first stage the worker forms expectation over the maximized outcome in the second stage. Given the Fréchet distribution it can be shown that this implicit value has a closed form solution.

**Proposition 1.** *The expectation of the maximization problem over  $J$  alternatives, where the benefit accrued is  $\delta_i \times \epsilon_i$  and where  $\epsilon_i$  is Fréchet distributed with CDF  $F(x) = e^{-x^{-a}}$ , is given by the following expression,*

$$\sum_i E \left[ \max_i (\delta_i \times \epsilon_i) \right] = \left( \sum_j \delta_j^a \right)^{\frac{1}{a}} \Gamma \left( 1 - \frac{1}{a} \right)$$

where  $\Gamma(\cdot)$  is the Gamma function.

*Proof.* See appendix. □

This mirrors the implicit value commonly used in nested discrete choice estimations us-

---

<sup>15</sup>Different interpretations are possible: For agriculture the sectoral switching cost might absorb some of the cost of moving to a major urban center, for other sectors they might simply signify the loss of sector specific human capital. For the geographical part the reallocation cost might absorb the lost utility due to disrupted social connections, a psychic cost or the actual economic moving cost.



ing Gumbel distributed additive preference shocks instead of Fréchet distributed multiplicative preference shocks. Based on comparing these implicit values subject to geographical switching cost that are proportional to the expected utility in the destination and specific to bilateral pairs of provinces the Household chooses the optimal location to move towards. Therefore the upper level problem of a worker residing in sector  $s$  and in province  $i$  reduces to,

$$\max \left[ \mathbb{E}_t \frac{\tilde{V}_{1,t+1|s}}{\mu_{i1}}, \dots, \mathbb{E}_t \frac{\tilde{V}_{i,t+1|s}}{\mu_{ii}}, \dots, \mathbb{E}_t \frac{\tilde{V}_{L,t+1|s}}{\mu_{iL}} \right]$$

where  $\mu_{ij}$  is the bilateral spatial mobility cost, where  $\mu_{ii}$  is normalized to 1, and where the implicit value  $\tilde{V}_{i,t+1|s}$  indicates the expected utility obtained after observing the preference shocks in the second stage and making the utility optimizing decision. Due to sectoral switching costs the implicit value depends on the initial sector the worker is currently working in. The closed form is given by,<sup>16</sup>

$$\tilde{V}_{i,t+1|s} \propto \mathbb{E}_t \left( \frac{\rho_j}{P_{j,t+1}^{1-\delta} r_{j,t+1}^\delta} \right) \left( w_{i,s,t}^\gamma + \mu_s^{-\gamma} \sum_{k \neq s} w_{j,k,t+1}^\gamma \right)^{\frac{1}{\gamma}}$$

An attractive property of this formulation of the labor reallocation problem is that bilateral flows between provinces is primarily driven by a measure of aggregate attractiveness of the destination province rather than specifically tied to sector specific dynamics within that destination provinces. This is a more realistic choice in a setting where migrants in faraway provinces have little information about the specific conditions in specific sectors but might have some information about the general attractiveness of a destination. Crucially, the key determinant of the direction of migration flows is the relative size of spatial versus sectoral switching costs, pinning down to what extent labor adjusts between provinces rather than within provinces. I will return to this point during the quantitative analysis. Given the Fréchet distributed preference shocks, standard properties imply the following closed form for the shares of workers who move across provinces,

---

<sup>16</sup>In the current setting the Fréchet dispersion parameter  $\gamma$  is symmetric across locations, therefore we can abstract from additional multiplicative term that determines the scale of the expectation,  $\Gamma(1 - \frac{1}{\gamma})$ . However, one can easily extend the current setting to account for heterogeneity of local sectoral labor supply elasticities.

$$\sigma_{ij,s}(\tilde{\mathbf{V}}_{t+1|s}) = \frac{\left(\mathbb{E}_t \tilde{V}_{j,t+1|s} \times \frac{1}{\mu_{ij}}\right)^\nu}{\Omega_{i,s,t}}$$

where  $\Omega_{i,s,t} \equiv \sum_j \left(\mathbb{E}_t \tilde{V}_{j,t+1|s} \times \frac{1}{\mu_{ij}}\right)^\nu$  summarizes the option value of a person currently working in sector  $s$  and residing in province  $i$ , where  $\tilde{\mathbf{V}}_{t+1|s}$  is the vector of implicit values  $\tilde{V}_{j,t+1|s}$  as defined above, where  $\mu_{ij}$  refers to the geographical switching cost, and where implicit values depend on expected wages, rental rates, price indices and the sectoral switching cost  $\mu_s$ , finally  $\nu$  defines the elasticity with regard to changes in the implicit values or alternatively the switching cost.

Conditional on reallocating and upon arrival in the province the worker uncovers her vector of sector specific preference shocks,  $\iota_t$  and makes a choice selecting a sector. Again assuming Fréchet distributed preference shocks with dispersion parameter  $\gamma$ , one can obtain the following closed form for the share of workers that flow into industry  $r$  in province  $i$  and where prior to that in industry  $s$ ,

$$\sigma_{i,s,r}(\mathbf{w}_{i,t}) = \frac{\mu_s^{-\gamma} w_{i,r,t}^\gamma}{w_{i,s,t}^\gamma + \mu_s^{-\gamma} \sum_{k \neq s} w_{j,k,t+1}^\gamma} \quad \text{for } s \neq r$$

$$\sigma_{i,s,s}(\mathbf{w}_{i,t}) = \frac{w_{i,s,t}^\gamma}{w_{i,s,t}^\gamma + \mu_s^{-\gamma} \sum_{k \neq s} w_{j,k,t+1}^\gamma} \quad \text{for } s = r$$

where  $\mathbf{w}_{i,t}$  is the vector of wages in province  $i$  and  $w_{i,r,t}$  refers to the wage in sector  $r$  and in province  $i$ . Since the other determinants of indirect utility enter symmetrically across all options, they do not affect the sectoral shares. Finally, one can state the flows from province  $i$  and sector  $r$  to province  $j$  and sector  $s$ , as,

$$\sigma_{ij,sr}(\tilde{\mathbf{V}}, \mathbf{w}) = \sigma_{ij,s} \sigma_{j,s,r}$$

where  $\tilde{\mathbf{V}}$  represents the vector of expected indirect utilities across provinces, and where province-sector specific flows are separable between,  $\sigma_{ij,sr}$ , that is the bilateral flows between province  $i$  and province  $j$ , and the sorting into sector  $r$  within province  $i$ ,  $\sigma_{i,s,r}$ . Total labor supply is then given by a market clearing condition, that is,

$$L_{i,s,t+1} = \sum_{j,r} \sigma_{ji,rs}(\tilde{\mathbf{V}}, \mathbf{w}) L_{j,r,t}$$

## 6 Estimation

In order to use the model described in the previous section for a quantitative analysis of the World War I shock, one needs to obtain estimates of the key parameters. On the labor demand side, we need to obtain trade elasticities,  $\{\epsilon_s\}$ , scale elasticities,  $\{\alpha_s\}$ , and productivity shifters,  $\{A_{i,s}\}$ . On the labor supply side we need to estimate switching costs and geographical and sectoral supply elasticities. The estimation of the parameters determining labor demand can be done separately, since changes in the spatial equilibrium are sufficiently informative to estimate them. Given those estimates we can then estimate the parameters associated with the labor supply model.

### 6.1 Labor demand

The estimation of the key parameters that determine labor demand relies mainly on the labor market data - that is wages and employment size for each province-sector. I demonstrate how to use that data in conjunction with the model structure in order to estimate the key parameters that determine labor demand. In the first step I use a structural approach to separate out origin specific marginal cost prices and market access. In a second step, I then regress the obtained prices on wages and labor densities to obtain the structural parameters.

#### 6.1.1 Obtaining Origin-Prices

From the static (spatial) equilibrium, one can obtain the following two equations,

$$Y_{i,s} = \sum_j X_{ij,s} = \sum_j \tau_{ij}^{-\epsilon_s} p_{is}^{-\epsilon_s} P_{js}^{\epsilon_s} E_{js}$$

$$E_{i,s} = \sum_j X_{ji,s} = \sum_j \tau_{ji}^{-\epsilon_s} p_{js}^{-\epsilon_s} P_{is}^{\epsilon_s} E_{is}$$

where the first equation states that total income in province  $i$  and sector  $s$ ,  $Y_{i,s}$ , must equal the cumulative export sales for that sector, that is the sum of all export flows from the origin province  $i$  to any province  $j$ , i.e.  $\sum_j X_{ij,s}$ . Since export flows follow the gravity structure the second equality follows. The second equation states that total expenditure in province  $i$  on goods from sector  $s$ , must equal total incoming export flows from all origin provinces  $j$ , that is  $\sum_j X_{ji,s}$ . Combining and rearranging, one can obtain a system

of equations in terms of prices only,

$$p_{is}^{\epsilon_s} = \sum_j \tau_{ij}^{-\epsilon_s} \left( \sum_k \tau_{kj}^{-\epsilon_s} p_{ks}^{-\epsilon_s} \right)^{-1} \frac{E_{js}}{Y_{is}}$$

where  $p_{i,s}^{\epsilon_s}$  refers to the origin prices introduced above. Standard results in economic geography imply that this equation can be solved to find the unique vector of provincial origin prices (up to normalization) for each sector,  $\mathbf{p}_s^{\epsilon_s}$  (as shown and first introduced by [Allen and Donaldson \(2018\)](#)).

Using the labor market data before and after the war - that is for 1914 and 1920 - and using the housing market data to construct disposable income across provinces,  $E_{is} \equiv \beta_s Y_i$ , one can implement the inversion described in the previous paragraph. In the implementation, I first calculate the Cobb Douglas expenditure shares as the national income share of an industry out of aggregate labor income. This is theoretically consistent with one input economic geography model described above. The housing expenditure share  $\delta$  is obtained as described in section 3. I use the shortest distance along the railroad graph between Spanish provincial capitals and furthermore add France as an additional location, where the distance to France is the shortest distance to Paris across railroad and maritime linkages. The iceberg transport cost is calibrated to be,  $\tau_{ij} = \text{distance}_{ij}^{-1}$ , calibrating the distance elasticity to the canonical value of -1 ([Head and Mayer; 2013](#)). Since I do not have coherent labor market data for France, I only include the total value of sectoral exports<sup>17</sup> as additional demand into the economic geography system.

### 6.1.2 Price Regression

In the second step, I can use marginal cost pricing, which implies that  $p_{i,s} = \frac{w_{i,s}}{A_{i,s} L_{i,s}^{\alpha_s}}$ , to obtain a log-linear expression of prices as a function of sector-province employment levels and wages. Taking the first difference, I obtain the following equation,

$$\epsilon_s \log \frac{\tilde{p}_{is,t+1}}{\tilde{p}_{is,t}} = \delta_i + \epsilon_s \log \frac{\tilde{w}_{is,t+1}}{\tilde{w}_{is,t}} - \alpha_s \log \frac{\tilde{L}_{is,t+1}}{\tilde{L}_{is,t}} - \log \frac{\tilde{A}_{is,t+1}}{\tilde{A}_{is,t}} \quad (4)$$

where relative changes in origin-prices of sector  $s$  in province  $i$ ,  $\frac{\tilde{p}_{is,t+1}}{\tilde{p}_{is,t}}$ , are a function of

<sup>17</sup>French exports are at the yearly level while the labor market data is in terms of the hourly wage and only covers a subset of the overall economy of Spain. When introducing the exports into the model I divide the total value by  $54 \times 50$  to translate the value into hourly exports. Then I multiply it by the share of the industry that is represented in the sample, that is .44.

relative changes in wages and employment levels in that sector-province and where  $\tilde{x}$  indicates that the variable  $x$  has been normalized relative to a sector-specific baseline province. The responsiveness of origin prices with regard wages and employment levels is pinned down by the trade elasticity,  $\epsilon_s$ , and the scale elasticity  $\alpha_s$ , respectively. The scale elasticity itself is combination of productivity externalities and how these productivity externalities in turn translate into income gains, that is  $\alpha_s = \psi_s \times \epsilon_s$ . We can define the structural residual as  $\eta_{i,s,t} \equiv \log \frac{\tilde{A}_{i,s,t+1}}{\tilde{A}_{i,s,t}}$ , which is the unobserved productivity evolution at the sector-province level. Additionally, I include the full set of province specific fixed effects  $\delta_i$  to control for province specific confounding shocks.

**Endogeneity** A natural concern is the endogeneity of both wages,  $w_{i,s}$ , and employment,  $L_{i,s}$ . The model implies that as a result of increases in productivity,  $\frac{\tilde{A}_{i,s,t+1}}{\tilde{A}_{i,s,t}} > 0$ , labor demand will increase and move along the upward sloping labor supply curve, with increases in wages and employment levels as a result. This implies that the model structure indicates a positive correlation between the residual,  $\eta_{i,s,t}$ , and the wages and employment levels, which will in turn induce an upward bias for the estimation of  $\alpha_s$  and a downward bias for the estimation of  $\epsilon_s$ . The naive OLS results depicted in table 2 shows theoretically invalid negative trade elasticities and large estimates for the external scale parameter, consistent with the model implied bias. An instrument is therefore necessary to remedy the situation. The exclusion restriction for any instrument is that

$$E [\eta_{i,s,t} | \mathbf{z}_t] = E \left[ \log \frac{\tilde{A}_{i,s,t+1}}{\tilde{A}_{i,s,t}} | \mathbf{z}_t \right] = 0$$

where  $\mathbf{z}_t$  denotes the vector of instruments and  $\eta_{i,s,t} = \log \frac{\tilde{A}_{i,s,t+1}}{\tilde{A}_{i,s,t}}$  denotes the structural error as discussed above. The setting is more challenging than a standard endogeneity problem because of the presence of two - potentially correlated - endogenous variables. An appropriate instrument needs to induce sufficient independent and differential variation in the endogenous variables to separately identify their impact on the dependent variable. The model suggests that labor supply shifters interacted with the incidence of the shock can serve as a source of such variation. Intuitively, while the foreign demand shock translates into a labor demand shock that stems from the industries desire to expand their production, the curvature of local labor supply will determine whether the additional demand is being absorbed mostly into higher wages or larger sectoral size as measured by employment. As illustrated in the figures below.

Historical evidence as well as the stylized facts suggest that spatial frictions are high

Figure 4: Inelastic Labor Supply

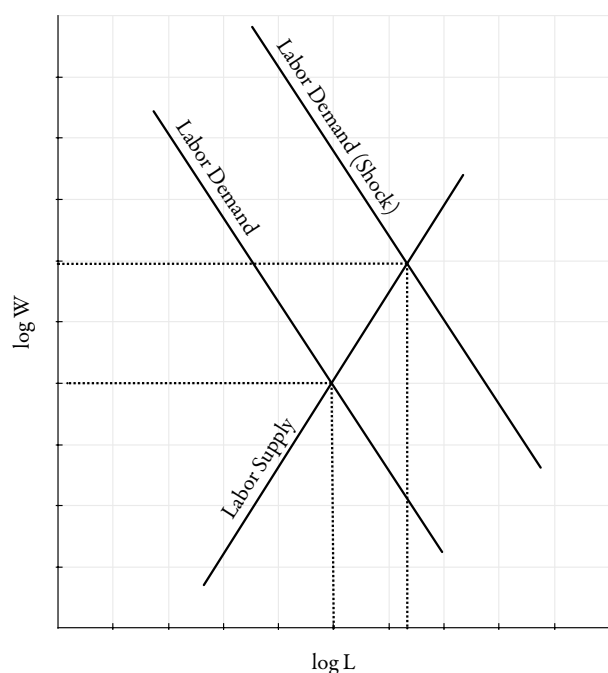
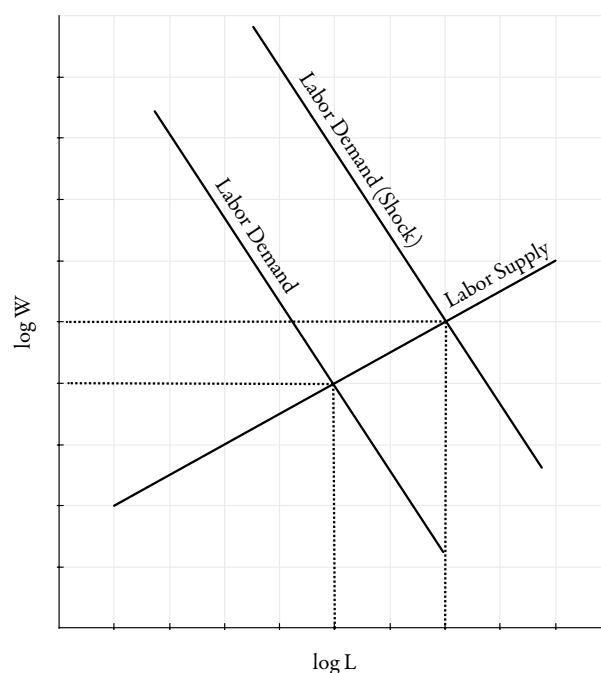


Figure 5: Elastic Labor Supply



Notes: The figure illustrates the underlying premise of the instrument. In the presence of two endogenous - potentially correlated - variables, independent variation is needed that differentially shifts the two variables. In the current setting a labor demand shock induces an outward shift of the labor demand curve from *Labor Demand* to *Labor Demand (Shock)*, inducing an increase in both wages and employment levels. The extent to which the shock is being absorbed by prices or quantities depends on the curvature of labor supply. If labor supply is tight - due to a small size of the local labor market - the curve will be upward sloping and wages will increase rather than employment levels. The opposite is true if labor supply is highly elastic.

and that labor supply is highly localized. I exploit this by using the (log) distance to Paris interacted with the (log) employment share of a sector within a province as a first instrument, where  $\log(\text{Employment Share of Sector in Province}) \equiv \log \frac{L_{i,s,1914}}{\sum_r L_{i,r,1914}}$  acts as a labor supply shifter and is interacted with distance to Paris as a reduced form proxy for differences in geographical advantages vis-a-vis the French destination market. A second instrument is given by a Harris Market Potential measure for the input market leaving own size out as a labor supply shifter, constructed as  $LMA_{i,s} = \sum_{j \neq i, r \neq s} \frac{1}{\text{distance}_{i,j}} L_{j,r}$ . The first stages are reported separately in table 6 and are sufficiently strong. Furthermore, there are no apparent pre trends neither along the distance margin nor along the sectoral shares as can be seen in the results for a regression that correlates wage growth prior to the war - that is between 1909 and 1914 - with distance to Paris and sectoral share, as can be seen in table ??.

**Results** The results can be seen in the table 1. The trade elasticities are theoretically consistent, positive and of comparable magnitude to sectoral trade elasticities currently found in the literature, though due to different aggregation and different time periods not directly comparable. The scale elasticity,  $\alpha_s$ , are mostly very imprecisely estimated with the result that for most sectors one cannot reject the presence of constant returns to scale, that is  $\alpha_s = 0$ . However, in some cases  $\alpha_s$  is significantly different from zero and negative, indicating decreasing returns to scale in chemicals, mining, metallurgy, metal works and textiles. These industries tend to require fixed installations, and thus decreasing returns in those sectors in the short and medium run seem plausible. The R squared is a natural measure of the fit of the model. To understand why, recall that prices solve the spatial equilibrium conditions, thus effectively functioning as residual income shifters, once one controls for market access differences. The R squared then measures how much of the variation can be explained by the log linear regression. The fit indicates that the model can explain half of the variation in the residual income shifter. Additionally, the model including labor densities performs much better than the model that only accounts for wage effects - as would be the case in the absence of any scale effects. The model without labor densities can only account for a quarter of the observed variation. Finally, the same estimation strategy can be used for the changes in labor market conditions from 1914 to 1925. The estimated scale elasticity,  $\alpha_s^{LR}$ , is reported in the table alongside the previous estimates. As can be seen, decreasing returns are no longer present in the industries in which they were present previously, suggesting decreasing returns to be a medium term phenomenon rather than a constant feature of these industries.

## 6.2 Labor supply

The estimation of the labor supply parameters proceeds in two steps and each step relies on different data sources.

### 6.2.1 Geographical Frictions

In the first step, I rely on data that shows the decennial change in the number of workers who live in a certain province but were born in another province, that is  $BAP_{i,j,t}$  for a worker who was born in province  $i$  but now lives in province  $j$ . The difference in this stock of foreign born workers,  $BAP_{i,j,t} - S \times BAP_{i,j,t-1}$  - adjusted for survivability rate  $S$  as explained in section 3 - is informative about the net inflow of foreign born workers,

either directly from the province under consideration or indirectly from other provinces. The data is adjusted so that the 1920s data shows the same number of total inhabitants born in a given province as the 1930s data, adding the additional population in their origin provinces. Using the closed forms from the previous section I can construct the model equivalent of this moment. The (estimated) stock of workers born in province  $i$  and currently residing in province  $k$  is given by,

$$\widehat{BAP}_{i,k,1930} = \sum_{j,s} \sigma_{jk,s}(\tilde{\mathbf{V}}_{1930|s}, \mathbf{w}) \times \pi_{i,s,1920} \times S \times BAP_{i,j,1920}$$

where  $\widehat{BAP}_{i,k,1930}$  refers to the simulated stock of workers born in province  $i$  and currently residing in province  $k$ ,  $\pi_{j,r,1920}$  refers to the industry share of industry  $r$  in province  $j$  in 1920 and where the closed form for the share of flows between province  $j$  and province  $k$  originating from sector  $s$  is given by  $\sigma_{ij,s}(\tilde{\mathbf{V}}_{t+1|s}) = \frac{\left(\mathbb{E}_t \tilde{V}_{j,t+1|s} \times \frac{1}{\mu_{ij}}\right)^v}{\Omega_{i,s,t}}$ . Implicitly, this is assuming that there is no sorting across industries of different groups of inhabitants, which in the absence on additional information is a necessary assumption. In the baseline estimation, I assume that wages and price indices follow a random walk. The geographical switching cost is calibrated as a function of distance that is

$$\mu_{ij} = \zeta_{cons} \times \zeta_i^1 \times distance_{ij}^{\zeta_{ij}^2}$$

where  $distance_{ij}$  is the shortest distance across railroad and maritime travelling routes from the province capital in  $i$  to the province capital in  $j$  in km. The structural estimation chooses the parameter vector  $\beta = (\zeta_1^1, \dots, \zeta_1^I, \rho_1, \dots, \rho^I, \zeta^2, \nu, \gamma, \mu_1, \dots, \mu_S)$  to match the observed moments, that is minimizing the error between imputed and observed quantities of workers born in another province,

$$\begin{aligned} \eta_{i,j}(BAP_{i,j,1930}, \beta) &= BAP_{i,j,1930} - \widehat{BAP}_{i,j,1930} \\ \hat{\beta} &= \arg \min_{\beta \in B} \boldsymbol{\eta}(BAP_{1930}, \beta)' \boldsymbol{\eta}(BAP_{1930}, \beta) \end{aligned}$$

where  $\boldsymbol{\eta}$  is the stacked vector of structural errors,  $\eta_{i,j}$ .

**Identification** The origin varying scalar,  $\zeta_i^1$ , determines the out-province migration share. Conditional on moving out of a province, the distance between the origin province and the destination province is informative about how geographical frictions



affect migration flows and thus determines the distance elasticity,  $\zeta^2$ . The incoming migration to specific provinces above and beyond what is justified by wage differences informs the province specific amenities,  $\rho_i$ . The responsiveness of in migration to dispersion in wages across sectors within a given province pins down the local supply elasticity,  $\gamma$ , while the response to dispersion of imputed indirect utilities across provinces informs the estimation of the spatial migration elasticity,  $\nu$ .

### 6.2.2 Sectoral Switching Costs

In order to estimate sectoral switching costs, I fit the model to changes in labor market conditions at the province-sector level from before to after the war. A key concern is that migration decisions were made during the war based on wage dynamics that are not part of the available data. In order to overcome this limitation I propose to use the estimated labor demand model together with sectoral trade data from 1915 to simulate the market clearing wages in the presence of the World War shock. I proceed by first using the 1914 data to impute the residual productivities,  $\{A_{i,s,1914}\}$ , and then feed in the trade shock to back out the simulated market clearing sectoral wage vectors,  $\hat{\mathbf{w}}_{s,1915}$ . Using these sectoral wage vectors as expected wages, and calibrating the spatial friction to the estimated values from the previous section, I use the closed forms to match the observed changes in employment size between 1914 and 1920,

$$\hat{L}_{i,s,1920} = \sum_{j,r} \sigma_{ji,rs}(\hat{\mathbf{w}}) L_{j,r,1914}$$

where  $\hat{L}_{i,s,1920}$  refers to the estimated stock of workers in province  $i$  and sector  $s$  in 1920, and  $L_{j,r,1914}$  refers to the observed size of industry  $r$  and province  $j$ , and  $\sigma_{ji,rs}(\hat{\mathbf{w}})$  is the closed form for migration flows between province  $j$  to province  $i$  and sector  $r$  to sector  $s$ . Recall that,

$$\sigma_{ij,rs}(\tilde{\mathbf{V}}, \hat{\mathbf{w}}) = \sigma_{ij,s} \sigma_{s,r,j}$$

that is the bilateral migration flows between sectors and provinces is a composite between outgoing migration between province  $i$  and province  $j$  in sector  $s$  and workers who upon arrival in province  $i$  sort into sector  $r$ . The structural error is given by,

$$\eta_{i,s}(\beta) = L_{i,s,1920} - \hat{L}_{i,s,1920}$$

$$\hat{\beta} = \arg \min_{\beta \in B} \eta(\beta)' \eta(\beta)$$

where  $\eta$  is the stacked vector of structural errors,  $\eta_{i,j}$  and where the structural procedure chooses  $\beta = (\mu_{agriculture,1}, \dots, \mu_{agriculture,I}, \mu_2, \dots, \mu_S, \gamma)$  to minimize the distance between the observed and the estimated employment size of each sector-province observation. Notice that the parameter vector includes province-specific switching costs for agriculture. The reason for this is twofold. Firstly, the switching costs associated with agriculture have a natural interpretation that is related to rural-urban migration, thus also involving some form of within province spatial friction. Since provinces differ in their size, this needs to be accounted for. Secondly, the changes of the agricultural sector are quantitatively important to match.

**Identification** With spatial frictions being calibrated, the size of the sectoral switching cost,  $\mu_s$ , is informed by the persistence of sectoral employment size in the presence of local wage disparities between sectors. An important caveat is that sectoral switching costs can only be identified in a scenario where workers do not reallocate despite a positive wage differential.

**Results** The results of the migration cost estimation are reported in table 8 in the appendix. Spatial frictions are prohibitively high implying low levels of internal migration with 2.7 percent of the population reallocating during the fitted period which is a gross measure. This is consistent with reported decennial *net* internal migration of 2.8 percent between 1911 and 1920 (Silvestre; 2005). Conditional on migrating distance is an important determinant with the composite distance elasticity,  $\zeta^2 \times \nu$ , giving a value of 2.38. Finally, labor is highly sticky, with a high degree of heterogeneity across sectors. Agriculture as a sector tends to be especially sticky across all provinces with a high degree of heterogeneity, nevertheless absolutely speaking agriculture releases most of the labor. This is to say that wage differentials are so large that high switching costs are necessary to justify the lack of mobility.

## 7 Quantitative Analysis: Spain without WWI

**Implementation** Having estimated the parameters that determine both labor supply and demand, one can now use the model to determine the counterfactual evolution of the Spanish economy in the absence of the WW1 shock. Since labor flows depend on

the expectations of utilities across province-sectors, and since those utilities themselves depend on the migration choices - via the scale economies - there is a potential for multiple equilibria in this class of model, and a necessity for equilibrium selection when conducting the counterfactual.<sup>18</sup> The baseline results presented here assume that wages and price indices follow a random walk and therefore are expected to remain at the level of the initial equilibrium observed in 1914. That is, workers coordinate using the current wages and price indices. Alternatives to that baseline can be explored. Conditional on implied reallocation patterns market clearing wages can be calculated and conclusions about impacts on income evolution can be drawn. In the following I compare the counterfactual 1920 wages and labor distribution with the observed state of the economy in 1920.

**Sectoral Employment Growth** One informative aspect of the counterfactual is to compare the aggregate industry sizes between the two scenarios. The results of such a comparison are presented in figure 13. The results indicate two important aspects: Firstly, there is high degree of reallocation from the agricultural sector towards the manufacturing sector, with the manufacturing sector as a whole growing by 1 percent as a result. A second important pattern is the heterogeneous response within the manufacturing sector with sectors that were particularly affected by the shock gaining substantially in size. Amongst those food, garments, textiles and metal works stand out, with the largest changes taking place in the textile sector.

**Regional Employment Growth** The same analysis can be conducted looking at province sizes rather than sectoral sizes. The results are presented in figure 16. There are very small difference in regional growth between the two scenarios, consistent with the finding that most of the adjustment is due to within provincial reallocation rather than between provincial allocation. Incidentally this is also consistent with a key characteristic of the migration choice framework highlighted above, that is that migration decisions do not respond effectively to individual industry dynamics but rather respond to the aggregate appeal of a destination, as captured in the estimation by the amenity values. Those amenities do not change in the counterfactual thus driving the patterns of the limited migration flows in either scenario.

---

<sup>18</sup>An alternative approach is to bound the possible outcomes by setting up the counterfactual problem as an MPEC Reguant (2016). This approach is currently being examined but the results are not yet available.

**Spatial Inequality** The model can also be used to calculate changes in nominal income aggregated at the sector-province level in the counterfactual. In the data the spatial gradient described in the reduced form section 4 led to a differential growth pattern between northern provinces (defined as above Madrid in terms of proximity to France), and southern provinces. Northern provinces experienced around 30 percent larger (nominal) income growth than southern provinces. The counterfactual without the War indicates only a minimal spatial gradient of 4 percent, residual productivity trends can explain a further 15 percent with the remaining 11 percent being attributed directly to the War. Since the model only allows a parametrically limited channel this can be understood as a lower bound for the effect of the War shock on spatial inequality. The exact patterns of incidence can be seen in map 16, indicating the differences in nominal income between the two scenarios which is indicative about the extent to which individual provinces managed to capture and monetize the demand shock effectively. The spatial gradient is visible, but is mitigate by provincial heterogeneity in sectoral specialization.

**Mobility costs and Market Integration** Finally, the model can be used to conduct counterfactuals on the real income levels if one allows for lower mobility costs. In order to simulate a reduction in the spatial mobility cost, I lower the bilateral travel distance between province capitals by 10%. Lowering mobility costs by 10% increases real income gains from the WWI-shock by an additional 3%, increasing the aggregate gains in real income from 20% to 23.59%. This is larger than the welfare gains from lower migration costs in the counterfactual non shock scenario where welfare would have only increased by 2.4%. This suggests that labor market integration and output market integration are complements. The reason why labor market integration and output market integration are complements is due to the fact that a more fluid labor market increases labor supply to the most productive industries and weakens localized competition for labor supply that in the presence of mobility cost can limit the extent to which the shock can be effectively exploited.

## 8 Conclusion

My primary interest was to examine to what extent labor market frictions can inhibit economic development of a country. I used a newly collected historical dataset that combines trade and labor market data, to examine a unique historical episode: A tem-

porary trade shock to a developing economy that prior to the shock only underwent slow structural transformation. I demonstrated the key features of the shock and its impact on regional development within Spain: The shock was temporary, sectorally heterogeneous, large and spatially biased. It induced spatially tilted regional development and affected provinces heterogeneously depending on their initial industrial specialization. I built a quantitative economic geography model that can account for the dynamic response to the temporary shock. A baseline economic geography model is extended to be better suited to match the regional dynamics of a temporary shock, by introducing and estimating labor market frictions that make employment sticky at the sectoral and provincial level as well as allowing for endogenous productivity feedbacks to determine the immediate productivity gains from reallocation.

An interesting aspect of the current work is that limited historical data can be complemented with structural models to improve both the estimation of objects of interest and in order to get further insights into phenomena that are not directly observed - as was done in this paper by obtaining unobserved sector-province labor reallocation patterns consistent with estimated migration costs and observed sectoral employment sizes.

The analysis suggests that high levels of labor market frictions and low immediate returns to reallocation due to the absence of scale economies and even the presence of decreasing returns in some industries prevented the Spanish economy from developing before the War. The shock induced reallocation across space and particularly between sectors within provinces, thus creating the fruitful preconditions for an economic take-off in the following decade. Finally, the analysis suggests that welfare gains from (output) market integration depend on the extent to which input markets are integrated.

This suggests four important conclusions: Firstly, labor market frictions are of primary importance for analyzing (spatial) development of a country or the lack thereof. Secondly, the relative size of different labor market frictions determines the pattern of development as well as the extent to which spatial arbitrage is possible between space and between sectors, making a quantitative understanding of these frictions important, in particular when analyzing patterns of spatial inequality. If spatial mobility costs are a reasonable concern in a developing country then policy makers need to take into account the distribution of labor as well as the spatial unevenness of the development process. Finally, labor market integration and output market integration ought to be considered in tandem to benefit from the complementary effects of both forms of mar-

ket integration.

## References

- Acemoglu, D., Autor, D. H. and Lyle, D. (2004). Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury, *Journal of Political Economy* **112**(3): 497–551.  
**URL:** <https://ideas.repec.org/a/ucp/jpolec/v112y2004i3p497-551.html>
- Adao, R., Costinot, A. and Donaldson, D. (2017). Nonparametric counterfactual predictions in neoclassical models of international trade, *American Economic Review* **107**(3): 633–89.  
**URL:** <http://EconPapers.repec.org/RePEc:aea:aecrev:v:107:y:2017:i:3:p:633-89>
- Allen, T. and Donaldson, D. (2018). Geography and path dependence, *Working Paper* .
- Blanchard, O. J. and Katz, L. F. (1992). Regional evolutions, *Brookings Papers on Economic Activity* **1**: 1–75.
- Chaney, T. (2008). Distorted gravity: The intensive and extensive margins of international trade, *American Economic Review* **98**(4): 1707–21.  
**URL:** <http://ideas.repec.org/a/aea/aecrev/v98y2008i4p1707-21.html>
- de Aduanas, D. G. (1910-1930). *Estadística general del comercio exterior de España con sus posesiones de Ultramar y potencias extranjeras*, Imprenta Nacional.  
**URL:** <https://books.google.fr/books?id=1aPIsshkkmYC>
- Dix-Carneiro, R. and Kovak, B. K. (2017). Trade liberalization and regional dynamics, *American Economic Review* **107**(10): 2908–46.  
**URL:** <http://www.aeaweb.org/articles?id=10.1257/aer.20161214>
- Eaton, J. and Kortum, S. (2002). Technology, geography, and trade, *Econometrica* **70**(5): 1741–1779.  
**URL:** <http://ideas.repec.org/a/ecm/emetrp/v70y2002i5p1741-1779.html>
- Ford, R. (1845). *A Hand-book for Travellers in Spain: And Readers at Home : Describing the Country and Cities, the Natives and Their Manners, the Antiquities, Religion, Legends, Fine Arts, Literature, Sports, and Gastronomy : with Notices on Spanish History*, number v. 1 in *A Hand-book for Travellers in Spain, and Readers at Home*, J. Murray.  
**URL:** <https://books.google.com.mx/books?id=PNExAQAAMAAJ>
- Hanlon, W. W. (2014). Temporary Shocks and Persistent Effects in the Urban System: Evidence from British Cities after the U.S. Civil War, *NBER Working Papers 20471*, National Bureau of Economic Research, Inc.  
**URL:** <https://ideas.repec.org/p/nbr/nberwo/20471.html>
- Harrison, J. (1978). *An Economic History of Modern Spain*, Manchester University Press.  
**URL:** <https://books.google.com/books?id=1hfpAAAIAAJ>
- Head, K. and Mayer, T. (2013). Gravity equations: Workhorse, toolkit, and cookbook, *CEPR Discussion Papers 9322*, C.E.P.R. Discussion Papers.  
**URL:** <http://ideas.repec.org/p/cpr/ceprdp/9322.html>
- Instituto de Reformas Sociales, L. (1916). *Informes de los Inspectores del Trabajo sobre la Influencia de la Guerra Europea en las Industrias Espanolas*.  
**URL:** <https://books.google.com/books?id=z1cPAQAIAAJ>

- Instituto Geográfico, C. y. d. E. S. (1912). *Censo de la población de España según el empadronamiento hecho en la península e islas adyacentes el 31 de diciembre de 1910...*, Censo de la población de España según el empadronamiento hecho en la península e islas adyacentes el 31 de diciembre de 1910, los hijos de M.G. Hernández.  
**URL:** <https://books.google.fr/books?id=7ikxAQAAMAAJ>
- Instituto Geográfico, C. y. d. E. S. (1922). *Censo de la población de España según el empadronamiento hecho en la península e islas adyacentes el 31 de diciembre de 1920...*, Censo de la población de España según el empadronamiento hecho en la península e islas adyacentes el 31 de diciembre de 1920, los hijos de M.G. Hernández.  
**URL:** <https://books.google.fr/books?id=7ikxAQAAMAAJ>
- Instituto Geográfico, C. y. d. E. S. (1932). *Censo de la población de España según el empadronamiento hecho en la península e islas adyacentes el 31 de diciembre de 1930...*, Censo de la población de España según el empadronamiento hecho en la península e islas adyacentes el 31 de diciembre de 1930, los hijos de M.G. Hernández.  
**URL:** <https://books.google.fr/books?id=7ikxAQAAMAAJ>
- Instituto Nacional de Prevision Social, L. (1930). *Clasificación general de industrias, oficios y comercios.*, 2 edn, Madrid. 006583.
- Juhasz, R. (2017). Temporary protection and technology adoption: Evidence from the napoleonic blockade, *CentrePiece - The Magazine for Economic Performance* 438, Centre for Economic Performance, LSE.  
**URL:** <https://ideas.repec.org/p/cep/cepcnp/438.html>
- Krugman, P. R. (1979). Increasing returns, monopolistic competition, and international trade, *Journal of International Economics* 9(4): 469–479.  
**URL:** <http://ideas.repec.org/a/eee/inecon/v9y1979i4p469-479.html>
- Kucheryavyi, K., Lyn, G. and Rodríguez-Clare, A. (2016). Grounded by Gravity: A Well-Behaved Trade Model with Industry-Level Economies of Scale, *NBER Working Papers* 22484, National Bureau of Economic Research, Inc.  
**URL:** <https://ideas.repec.org/p/nbr/nberwo/22484.html>
- Ministerio de Trabajo, C. e. I. (1927). *Estadística de los salarios y jornadas de trabajo referida al período 1914-1925*, number v. 1914-1930 in *Estadística de los salarios y jornadas de trabajo referida al período 1914-1925*, Sobrinos de sucesora de M. Minuesa de los Ríos, M. Servet.  
**URL:** <https://books.google.fr/books?id=TJ1CAAAAIAAJ>
- Nadal, J. (1975). *El fracaso de la revolución industrial en España, 1814-1913*, Ariel historia, Editorial Ariel.  
**URL:** <https://books.google.com/books?id=xqAYSxmeIJMC>
- Prados de la Escosura, L. (2016). Spain's Historical National Accounts: Expenditure and Output, 1850-2015, *Working Papers* 0103, European Historical Economics Society (EHES).  
**URL:** <https://ideas.repec.org/p/hes/wpaper/0103.html>
- Prados de la Escosura, L. (2017). *Spanish Economic Growth, 1850–2015*, Palgrave Studies in Economic History, Springer International Publishing.  
**URL:** <https://books.google.com/books?id=Jho0DwAAQBAJ>



- Redding, S. J. and Rossi-Hansberg, E. (2016). Quantitative Spatial Economics, *NBER Working Papers* 22655, National Bureau of Economic Research, Inc.  
**URL:** <https://ideas.repec.org/p/nbr/nberwo/22655.html>
- Reguant, M. (2016). Bounding Equilibria in Counterfactual Analysis, *Technical report*.
- Roldan, S. and Delgado, J. (1973). *La formación de la sociedad capitalista en España, 1914-1920*, number v. 2 in *Publicaciones del Fondo para la Investigación Económica y Social de la Confederación Española de Cajas de Ahorro: Fondo para la Investigación Económica y Social*, Confederacion Española de Cajas de Ahorro.  
**URL:** <https://books.google.com/books?id=SG7iXwAACAAJ>
- Silvestre, J. (2005). Internal migrations in Spain, 1877-1930, *European Review of Economic History* 9(2): 233–265.  
**URL:** <http://www.jstor.org/stable/41378422>
- Silvestre, J., Ayuda, M. I. and Pinilla, V. (2015). The occupational attainment of migrants and natives in Barcelona, 1930, *The Economic History Review* 68(3): 985–1015.  
**URL:** <http://dx.doi.org/10.1111/ehr.12090>

## A Tables

Table 1: Estimation Results - Labor demand parameters - Long Run

Industry	$\alpha_s$	Std Err	$\alpha_s^{LR}$	Std Err	$\epsilon_s$	Std Err
Agriculture	1.20	(2.60)	-0.93	(1.55)	4.71***	(1.74)
Books	-0.19	(1.08)	0.10	(1.54)	5.17***	(1.84)
Ceramics	-0.04	(1.32)	2.04**	(0.91)	5.34**	(2.12)
Chemicals	-0.76	(1.39)	-0.97	(1.09)	5.18***	(1.85)
Construction	0.92	(1.54)	0.11	(0.80)	4.22**	(1.99)
Decoration	0.84	(1.03)	-0.49	(1.51)	5.53***	(2.03)
Electricity	0.00	(1.24)	-0.17	(1.01)	5.47***	(1.89)
Food	0.05	(1.13)	0.72	(1.25)	4.58**	(1.80)
Forest	1.16	(4.71)	-6.10	(10.16)	4.84	(3.30)
Furniture	0.38	(0.87)	0.21	(1.19)	5.38**	(1.92)
Garments	0.28	(1.06)	0.41	(0.97)	4.44***	(1.79)
Glass	1.51	(2.92)	0.78	(1.42)	5.96***	(2.22)
Leather	1.79*	(1.07)	1.58	(1.52)	5.92***	(1.88)
Metal Works	-0.66	(0.88)	0.39	(1.43)	4.43***	(1.80)
Metallurgy	-0.98	(1.74)	-0.85	(1.75)	5.54***	(1.73)
Mines	-2.34*	(1.40)	2.80	(4.95)	5.99***	(1.83)
Paper	-3.06*	(1.89)	-6.86	(4.07)	3.21*	(1.93)
Public	1.34	(6.33)	1.61	(1.61)	5.37	(3.71)
Public Industry	22.21	(23.41)	1.18	(1.34)	11.99	(10.54)
Textiles	-0.88	(0.95)	0.48	(1.39)	4.04**	(1.75)
Tobacco	10.36*	(6.10)	2.78	(5.55)	1.43	(2.68)
Transport	1.51	(1.53)	0.71	(1.37)	4.78***	(1.85)
Varias	-0.95	(1.64)	-3.36	(3.77)	4.68**	(1.97)
Wood	0.53	(1.90)	1.00	(1.26)	4.84**	(1.99)
Observations			625			
R2			0.5892			
Province FE			✓			

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Notes: This table reports both the short run and long run results from the structural estimation of the labor demand parameters. The parameter  $\epsilon_s$  refers to the trade elasticity,  $\alpha_s$  for the composite external economies of scale parameter as discussed in the theory section. Additionally,  $\alpha_s^{LR}$  is reported, which is the corresponding scale elasticity if estimated for the 1914/1925 time frame instead of the 1914/1920 timeframe. The estimates are obtained via 2SLS instrumenting for employment size of sector  $s$  in province  $i$ ,  $L_{i,s}$  and wages  $w_{i,s}$ , using  $\log \text{distance}_{i,Paris} \times \log(\text{Employment Share of Sector in Province})$  as a first instrument, where  $\log(\text{Employment Share of Sector in Province}) \equiv \frac{L_{i,s,1914}}{\sum_r L_{i,r,1914}}$  works as a labor supply shifter and is interacted with distance to Paris as a reduced form proxy for differences in geographical advantages vis-a-vis the French destination market. A second instrument is given by a Harris Market Potential measure for the input market leaving own size out as a labor supply shifter, constructed as  $LMA_{i,s} = \sum_{j \neq i, r \neq s} \frac{1}{\text{distance}_{i,j}} L_{j,r}$ . The first stages for the 1914/1920 estimates are reported separately in table 6. The estimation is obtained on the sample that drops the 1% smallest industries, thus avoiding large leverage of outliers on estimates due to small measurement error in employment sizes and wages. Standard errors are obtained via bootstrap.

Table 2: Structural Estimation: Naive OLS

Industry	$\alpha_s$	T-Stats	$\epsilon_s$	T-Stats
Agriculture	0.91***	(6.97)	-1.20***	(35.61)
Books	1.00***	(8.39)	-0.99***	(13.16)
Ceramics	1.07***	(9.08)	-0.59***	(6.95)
Chemicals	0.85***	(9.79)	-0.37***	(5.35)
Construction	1.18***	(11.89)	-1.34***	(23.71)
Decoration	1.10***	(12.71)	-1.07***	(16.33)
Electricity	0.90***	(8.12)	-0.61***	(6.99)
Food	1.15***	(10.28)	-1.29***	(24.86)
Forest	0.87***	(9.83)	-3.30***	(22.60)
Furniture	1.04***	(9.86)	-0.98***	(16.02)
Garments	1.09***	(8.09)	-1.20***	(16.36)
Glass	0.99***	(5.80)	-1.05***	(12.06)
Leather	1.03***	(13.11)	-0.93***	(13.77)
Metal works	1.04***	(10.69)	-1.08***	(16.20)
Metallurgy	0.90***	(14.32)	-0.03	(0.46)
Mines	1.13***	(23.13)	0.77***	(13.56)
Paper	1.22***	(4.56)	-1.24***	(14.34)
Public	0.99***	(5.31)	-1.81***	(11.19)
Public Industry	0.61***	(6.26)	-2.80***	(25.83)
Textiles	1.06***	(9.65)	-1.24***	(17.07)
Tobacco	1.02***	(8.02)	-1.29***	(14.60)
Transport	1.06***	(10.10)	-1.06***	(15.39)
variants	0.98***	(3.61)	-0.89***	(6.91)
Wood	0.97***	(10.07)	-1.29***	(22.04)
Observations		625		
R2		0.8819		
Province FE		✓		

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Notes: This table reports the results of estimating the structural equation 4 without correcting for the endogeneity of wages and employment size of a sector. The estimation procedure is OLS. The dependent variable  $\ln p_{is}$  refers to  $\log \frac{p_{is,1920}}{p_{is,1914}}$  and the explanatory variables  $\ln w_{is}$  and  $\ln L_{is}$  refer to  $\frac{w_{is,1920}}{w_{is,1914}}$  and  $\frac{L_{is,1920}}{L_{is,1914}}$  respectively.

Table 3: Diff in Diff

	(1)	
	lworkers	
ldist	9.964***	(8.61)
treated	-0.0397	(-0.04)
treated=1 × ldist	0.00262	(0.02)
treated=1 × lexp	0.0686*	(2.41)
lexp × year_count	-0.00111	(-0.22)
Constant	-59.88***	(-7.45)
Observations	144	
Province FE	✓	

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Notes: This table reports the results of the diff-in-diff regression described in section

Table 4: Pre Trends

	(1)	
	lwage_growth	
ldist	-0.187	(-1.76)
lshare	-0.0140	(-0.83)
Constant	1.336	(1.77)
Observations	144	

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Notes: This table reports the results of a regression Source: Labor inspections (1909-1914)

Table 5: Spatial Gradient: Sectoral Estimates

	$\frac{Y_{i,s,1920}}{Y_{i,s,1910}}$	
Agriculture $\times \log(\text{DistanceParis})$	-0.619***	(-3.49)
Books $\times \log(\text{DistanceParis})$	-0.754***	(-4.01)
Ceramics $\times \log(\text{DistanceParis})$	-0.671***	(-3.55)
Chemicals $\times \log(\text{DistanceParis})$	-0.647***	(-3.44)
Construction $\times \log(\text{DistanceParis})$	-0.650***	(-3.59)
Decoration $\times \log(\text{DistanceParis})$	-0.717***	(-3.76)
Electricity $\times \log(\text{DistanceParis})$	-0.695***	(-3.66)
Food $\times \log(\text{DistanceParis})$	-0.671***	(-3.70)
Forest $\times \log(\text{DistanceParis})$	-0.793***	(-4.20)
Furniture $\times \log(\text{DistanceParis})$	-0.737***	(-3.90)
Garments $\times \log(\text{DistanceParis})$	-0.669***	(-3.70)
Glass $\times \log(\text{DistanceParis})$	-0.723***	(-3.70)
Leather $\times \log(\text{DistanceParis})$	-0.697***	(-3.70)
Metal Works $\times \log(\text{DistanceParis})$	-0.675***	(-3.71)
Metallurgy $\times \log(\text{DistanceParis})$	-0.668***	(-3.57)
Mines $\times \log(\text{DistanceParis})$	-0.653***	(-3.60)
Paper $\times \log(\text{DistanceParis})$	-0.700***	(-3.58)
Public $\times \log(\text{DistanceParis})$	-0.735***	(-3.66)
Public Industry $\times \log(\text{DistanceParis})$	-0.713***	(-3.64)
Textiles $\times \log(\text{DistanceParis})$	-0.678***	(-3.72)
Tobacco $\times \log(\text{DistanceParis})$	-0.802***	(-4.19)
Transport $\times \log(\text{DistanceParis})$	-0.664***	(-3.66)
Variats $\times \log(\text{DistanceParis})$	-0.737***	(-3.98)
Wood $\times \log(\text{DistanceParis})$	-0.684***	(-3.73)
$\log(\text{ShareInSector})$	0.0624	(1.02)
$\log(\text{ShareInProvince})$	-0.218**	(-2.79)
Constant	6.741***	(5.18)
Observations	685	

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

*Notes:* This table reports the results of a regression correlating nominal income growth between 1920 and 1910 at the sector province level with the (log) distance from the provincial capital to Paris. The distance measure is the shortest path along the railroad network and maritime linkages in kilometers. The regression allows for different intercepts for each sector. Additionally, the regression controls for the (log) employment share of sector  $s$  in province  $i$  in the national industry as a proxy for comparative advantage, as well as the (log) employment share of the sector within the province as a proxy for local labor market tightness.

Table 6: Labor Demand estimation: First stage

	(1)		(2)	
	$\log \frac{w_{is,1920}}{w_{is,1914}}$		$\log \frac{L_{is,1920}}{L_{is,1914}}$	
log(DistancetoParis) x log(ShareinProvince)	0.00569***	(4.81)	-0.0133***	(-8.86)
Log(LMA)	-0.0159	(-1.71)	0.0488***	(4.13)
Constant	0.827***	(11.55)	-0.386***	(-4.24)
Observations	657		657	
F Stat	15.53		46.08	

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Notes: This table reports the results of the first stage for estimating the structural equation 4. The first stage predicts the endogenous variables  $\log \frac{w_{is,1920}}{w_{is,1914}}$ , denoting (log) wage changes between 1920 and 1914 at the province-sector level, and  $\log \frac{L_{is,1920}}{L_{is,1914}}$ , denoting employment changes for the same time period at the province sector level. The first instrument is  $\log \text{distance}_{i,Paris} \times \log(\text{Employment Share of Sector in Province})$ , where  $\log(\text{Employment Share of Sector in Province}) \equiv \frac{L_{i,s,1914}}{\sum_r L_{i,r,1914}}$  works as a labor supply shifter and is interacted with distance to Paris as a reduced form proxy for differences in geographical advantages vis-a-vis the French destination market. A second instrument is given by a Harris Market Potential measure for the input market leaving own size out as a labor supply shifter, constructed as  $LMA_{i,s} = \sum_{j \neq i, r \neq s} \frac{1}{\text{distance}_{i,j}} L_{j,r}$ .

Table 7: Local Labor Supply and Income Dynamics

	$\frac{Y_{i,s,1920}}{Y_{i,s,1910}}$	
log(ShareInSector)	-0.0629	(-1.30)
log(ShareInProvince)	-0.173*	(-2.25)
log(EmploymentSize)	0.0933	(1.10)
Constant	0.860	(0.88)
Observations	637	

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Notes: This table reports the results from a regression of nominal income growth between 1920 and 1910 at the sector-province level on three different variables.  $\log(\text{Employment Share of Sector in Province}) \equiv \frac{L_{i,s,1914}}{\sum_r L_{i,r,1914}}$

Table 8: Results: Migration Cost Estimation

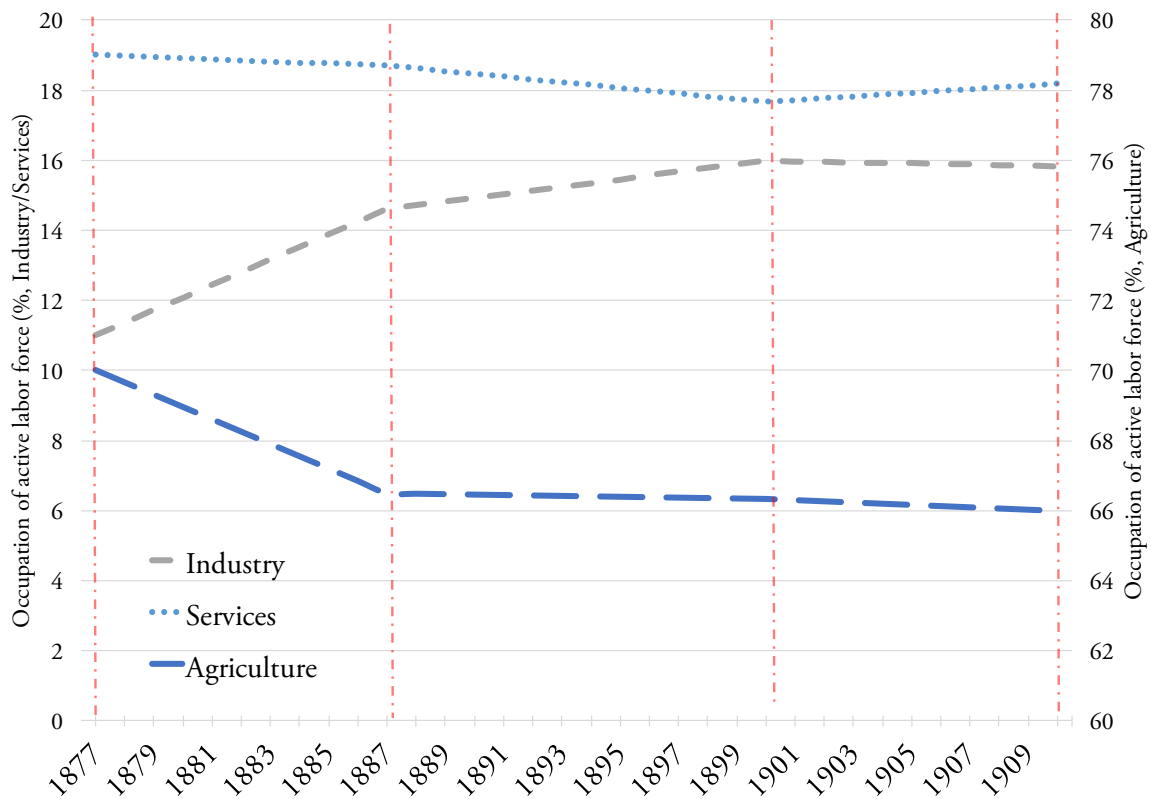
Province	$\rho_i$	$\log \zeta_i^1$	$\frac{1}{\mu_{agriculture,i}}$	Industries	$\frac{1}{\mu_s}$
Alava	1.09	0.92	0.03	Agriculture	0.2576
Albacete	0.20	3.30	0.00	Books	0.0452
Alicante	0.64	1.62	0.00	Ceramics	0.0601
Almeria	0.88	0.89	0.01	Chemicals	0.3394
Avila	0.65	1.95	0.01	Construction	0.0770
Badajoz	0.71	2.07	0.00	Decoration	0.3176
Balears	0.00	11.07	0.06	Electricity	0.0793
Barcelona	10.01	0.18	0.04	Food	0.2611
Burgos	1.25	1.36	0.01	Forrest	0.0852
Caceres	1.11	2.18	0.00	Furniture	0.3219
Cadiz	0.50	1.91	0.02	Garments	0.0642
Castellon	1.00	1.34	0.01	Glass	0.3366
Ciudad Real	1.18	2.31	0.00	Leather	0.6854
Cordoba	1.07	1.80	0.00	Metal Works	0.0001
Coruna	1.60	1.21	0.00	Metallurgy	0.1664
Cuenca	0.69	2.06	0.00	Mines	0.1801
Gerona	1.80	1.21	0.00	Paper	0.3458
Granada	0.33	2.88	0.00	Public	0.4204
Guadalajara	1.05	2.49	0.00	Public Industry	0.3710
Guipuzcoa	1.52	0.63	0.01	Textiles	0.0714
Huelva	0.88	1.51	0.00	Tobacco	0.0922
Huesca	1.56	1.01	0.00	Transport	0.1703
Jaen	1.09	1.41	0.00	Varias	0.5858
Leon	0.95	1.79	0.02	Wood	0.0000
Lerida	1.09	1.45	0.00		
Logrono	1.03	1.02	0.01		
Lugo	1.13	1.60	0.01	Elasticities and Constants	
Madrid	5.57	0.36	0.04	$\zeta^2$	1.49
Malaga	0.86	1.77	0.00	$\nu$	1.59
Murcia	0.93	0.95	0.00	$\gamma$	1.35
Navarra	1.22	1.45	0.02	$\log(\zeta_{cons})$	1.40
Orense	0.70	2.42	0.03		
Oviedo	0.79	2.06	0.02		
Palencia	0.59	1.79	0.03		
Pontevedra	1.63	1.21	0.02		
Salamanca	0.82	1.96	0.01		
Santander	0.83	0.75	0.02		
Segovia	0.90	2.02	0.00		
Sevilla	2.00	0.92	0.00		
Soria	1.07	1.31	0.01		
Tarragona	1.35	1.56	0.00		
Teruel	1.04	1.31	0.01		
Toledo	0.80	2.42	0.00		
Valencia	0.95	1.75	0.00		
Valladolid	0.98	1.31	0.03		
Vizcaya	1.09	0.67	0.04		
Zamora	0.64	2.18	0.01		
Zaragoza	0.39	2.93	0.01		

Notes: This table reports the results of the migration cost estimation. In the left column the amenity shifters associated with the different provinces are reported. Barcelona is normalized to 1, with the other provinces being expressed relatively to Barcelona. In the right column the sectoral switching cost parameter  $\mu_s$  is reported as well as the key elasticities pinning down spatial migration cost  $\mu_{ij} = \zeta_{cons} \times \zeta_i^1 \times distance_{ij}^{\zeta^2}$ . The parameters are obtained via minimum distance estimation and the procedure is described in detail in section ??.



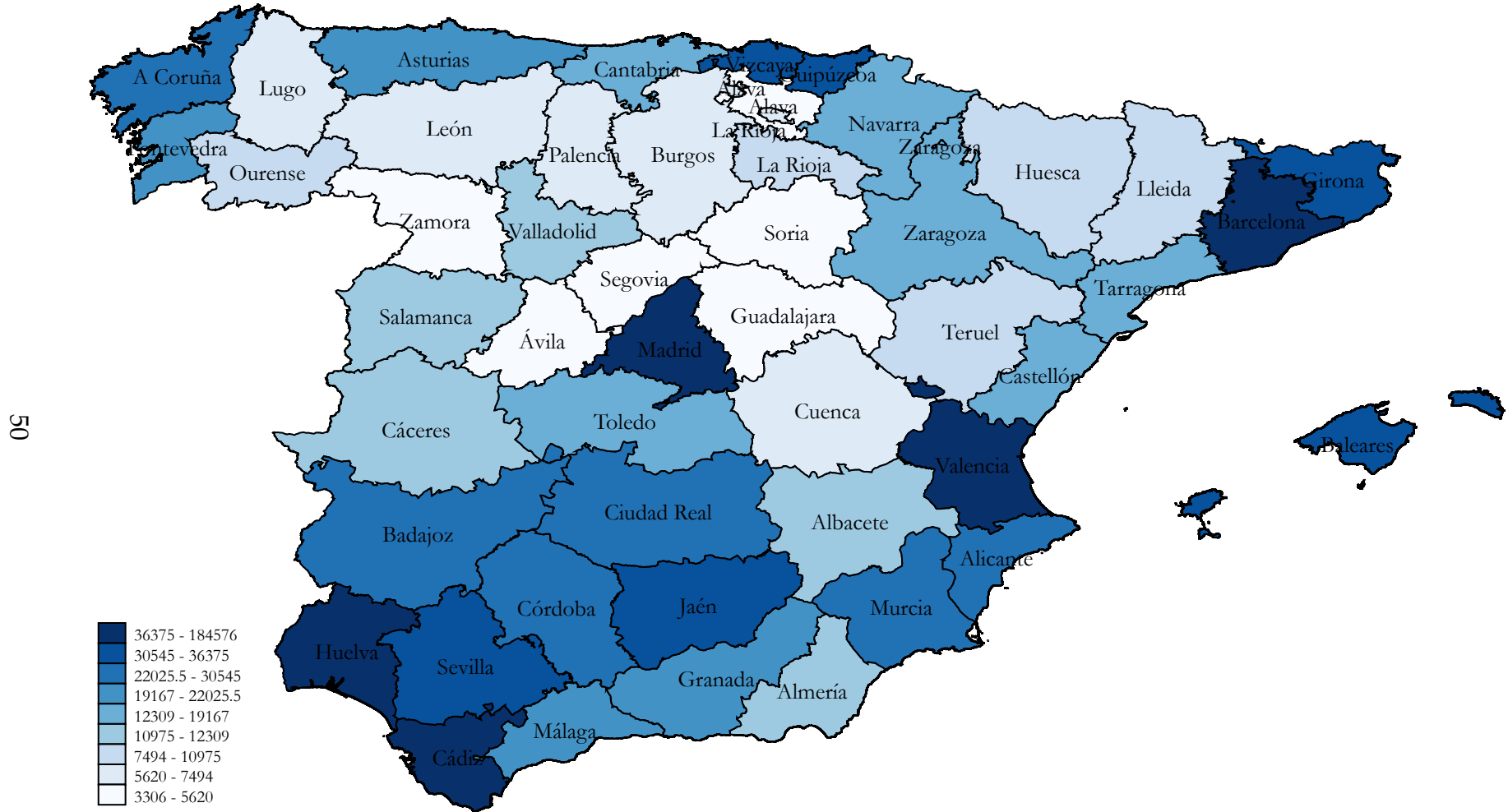
## B Figures

Figure 6: Structural Change in the 19th Century



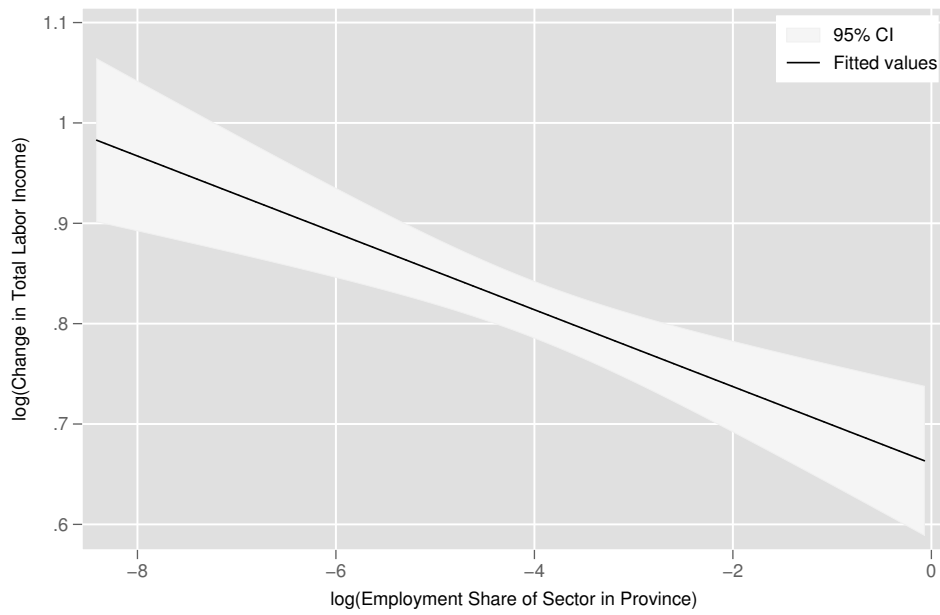
Notes: The figure depicts sectoral employment shares across the manufacturing sector/industry, agriculture and services. The shares are observed in the census data in 1877, 1887, 1900 and 1910 where census years are indicated by the red dotted line and the intervening years are imputed trend lines. Notice that while service and industry employment is plotted against the left y-axis, agricultural employment is plotted against right y-axis. The original computation of the aggregate employment share is due to [Harrison \(1978\)](#).

Figure 7: Spatial Distribution of Manufacturing Employment



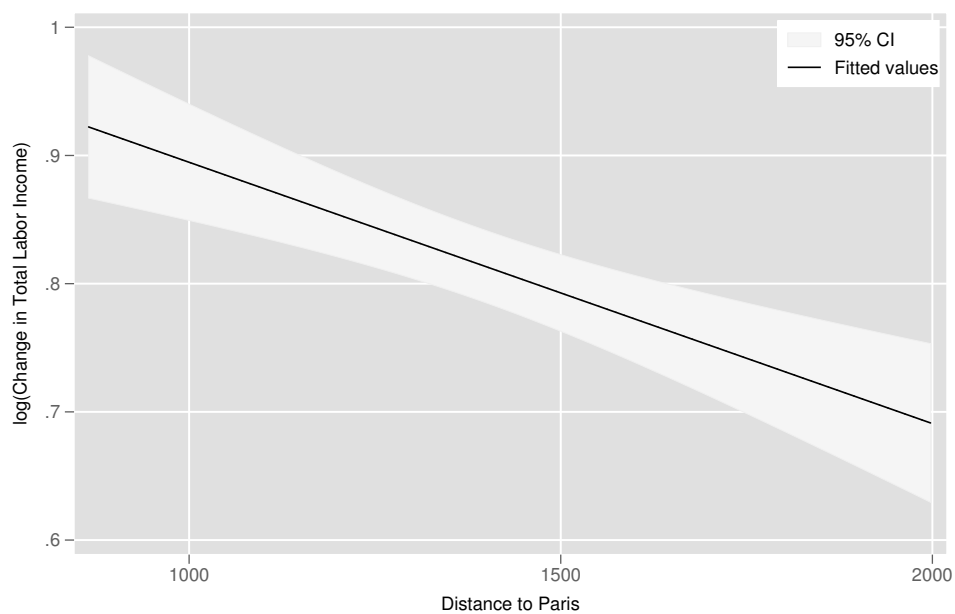
Notes: The map depicts total Manufacturing and Mining employment in the provinces in 1910 across Spain (without Canary Islands and North African possessions). The map is a choropleth with darker shaded colors depicting higher absolute numbers. The data is obtained from the Population census from 1910.

Figure 8: Local Labor Supply and Income Growth



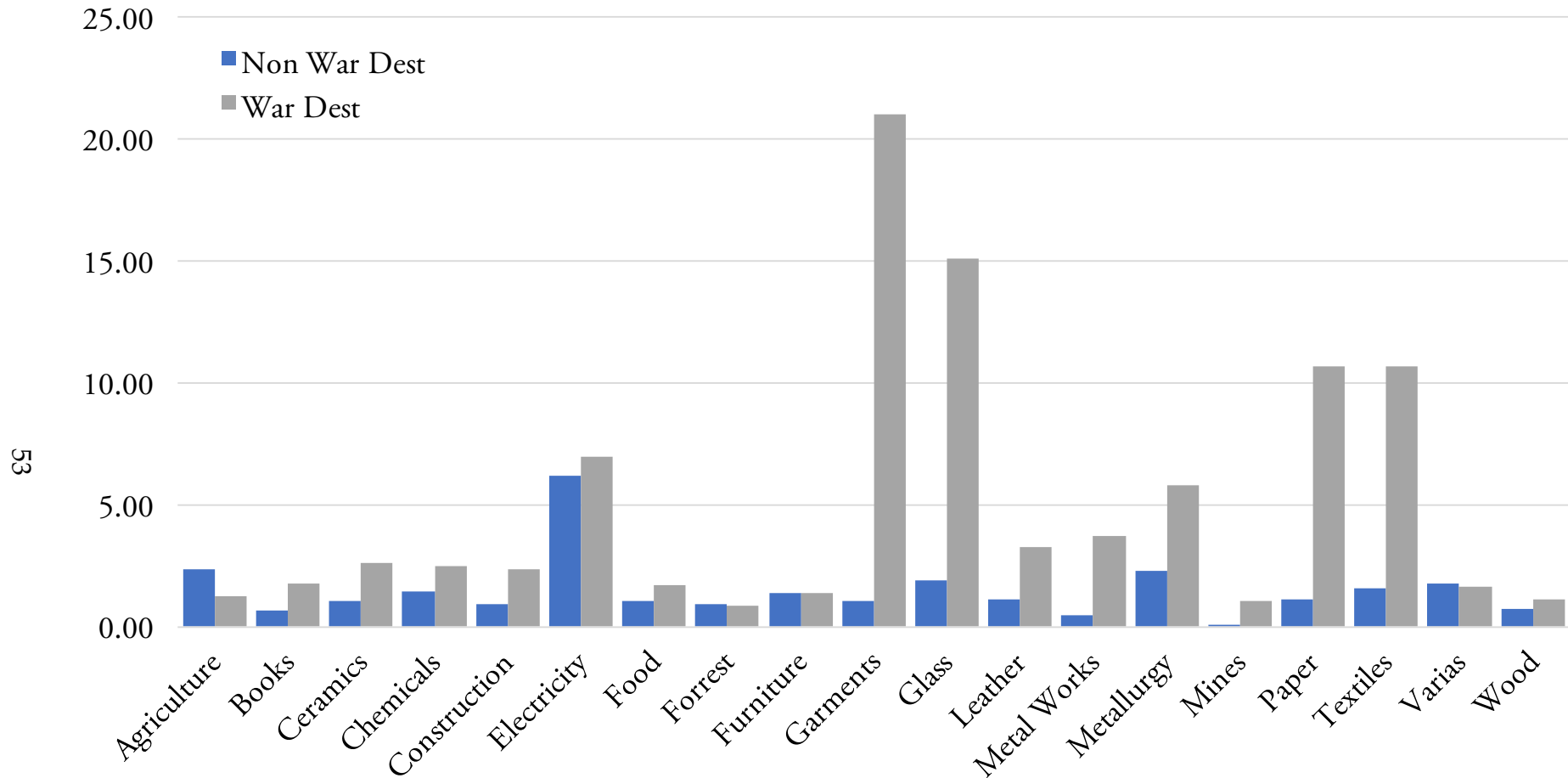
*Notes:* The graph shows the fitted line of a regression correlating (nominal) income growth at the sector province level between 1920 and 1914 with the log of the share of that sector in the total employed population in that province in 1914. Specifically, the variable on the x-axis is defined as  $\log(\text{Employment Share of Sector in Province}) \equiv \frac{L_{i,s,1914}}{\sum_r L_{i,r,1914}}$ . The data being used is the labor market panel introduced in the data section.

Figure 9: Spatial Gradient in Income Growth



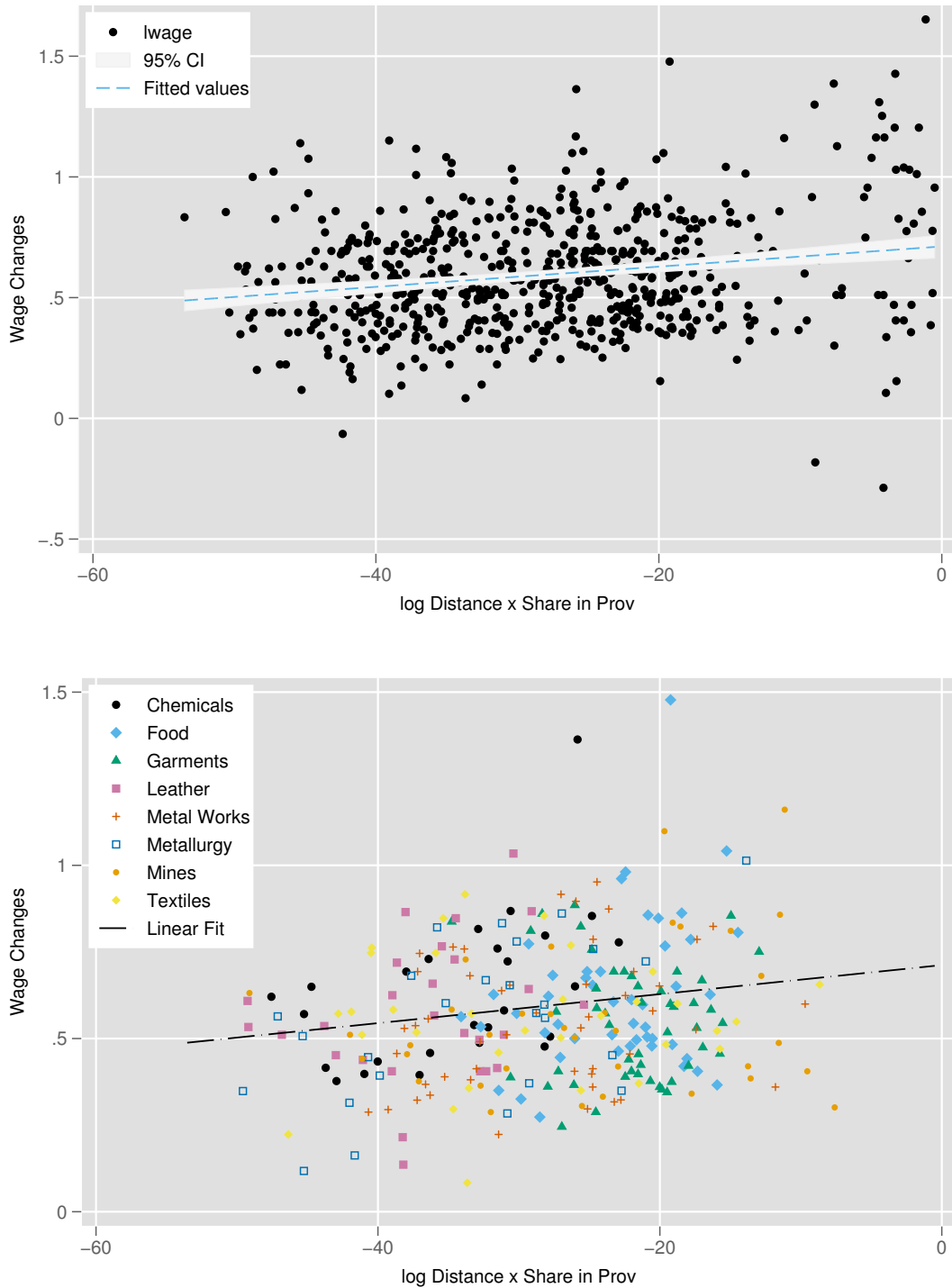
*Notes:* This figure reports the results of a regression correlating nominal income growth between 1920 and 1910 at the sector province level with the (log) distance from the provincial capital to Paris. The distance measure is the shortest path along the railroad network and maritime linkages in kilometers.

Figure 10: Sectoral Trade Growth: Belligerent vs Non Belligerent



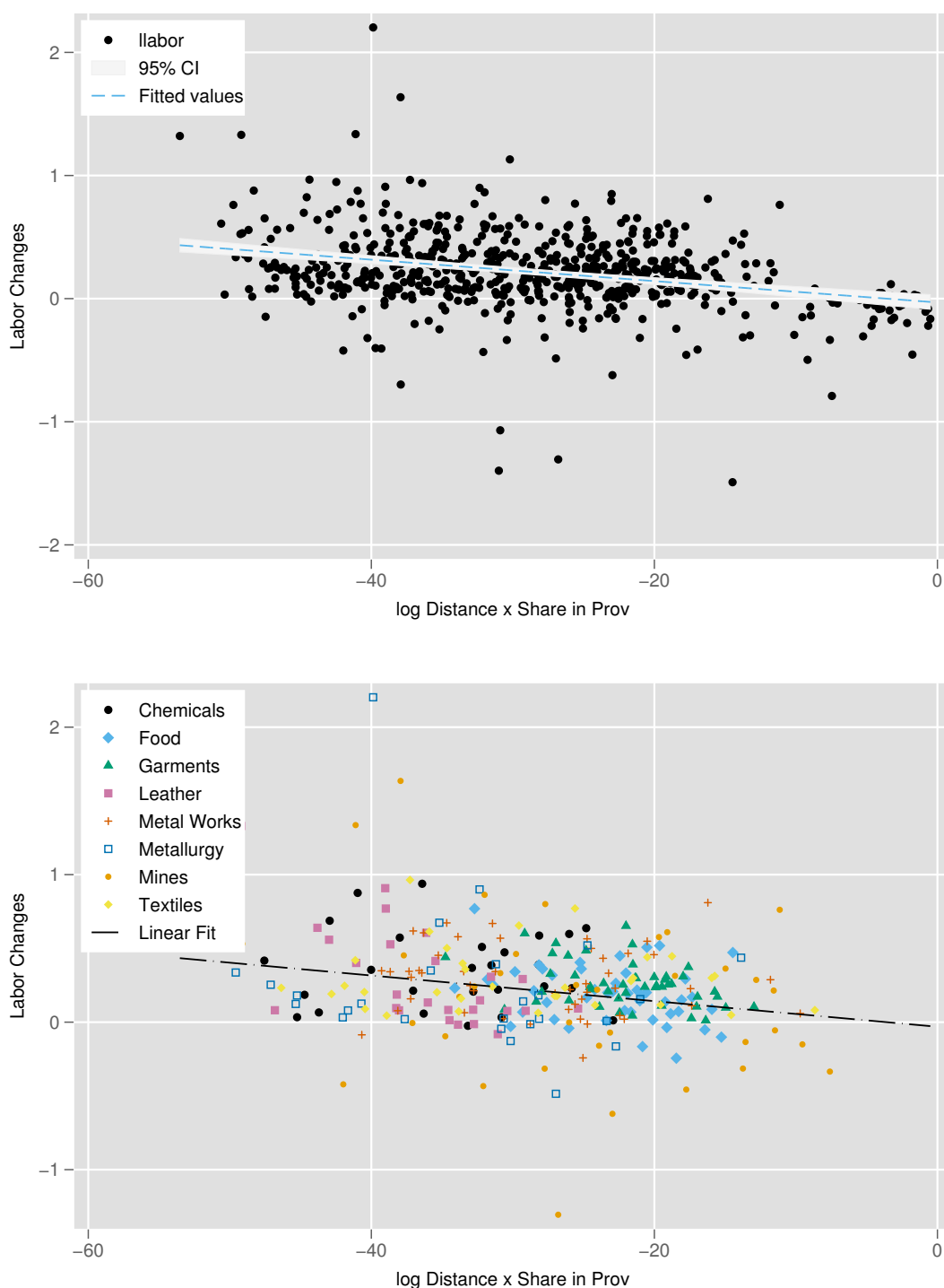
Notes: This figure reports the sectoral export growth for non belligerent destination countries - in blue - and belligerent destination countries in grey. The product level trade has been aggregated to sector level trade data to match the level of aggregation of the labor market panel. Growth rates are constructed by comparing the 1910 benchmark with average export values in 1915 and 1916, that is  $g_X^{War} \equiv \frac{1/2X_{Spain,War,1915} + X_{Spain,War,1916}}{X_{Spain,War,1910}}$  and correspondingly for non belligerent destinations. As discussed in section 4 I abstract from later years to avoid additional spatial frictions that perturbed international trade, in particular increased maritime warfare. To adjust for additional spatial disruptions of the frontline the belligerent countries are made up of France, Italy and the United Kingdom. The Non-belligerent countries exclude the United States and other later participants of WWI. The shock is being calculated using the official annual trade data in constant prices.

Figure 11: First stage for structural estimation (all industries)



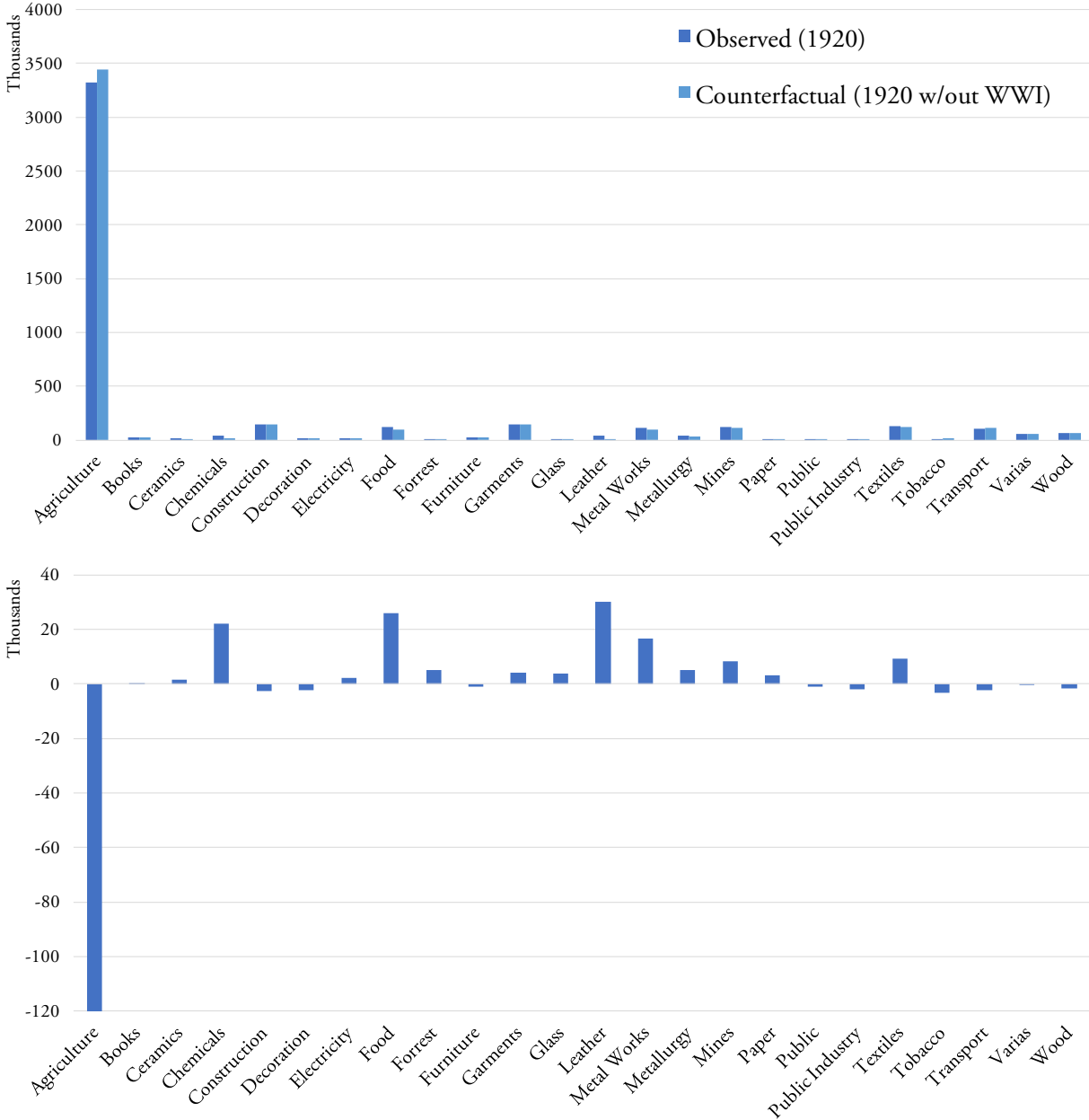
Notes: This figure shows the first stage regression predicting wage changes,  $\frac{w_{i,s,1920}}{w_{i,s,1914}}$  at the province-sector level, using  $\log distance_{i,Paris} \times \log(\text{Employment Share of Sector in Province})$  as an instrument, where  $\log(\text{Employment Share of Sector in Province}) \equiv \frac{L_{i,s,1914}}{\sum_r L_{i,r,1914}}$  works as a labor supply shifter and is interacted with distance to Paris as a reduced form proxy for differences in geographical advantages vis-a-vis the French destination market. Distance is calculated using the shortest path along a network of railroads and maritime linkages between province capitals in Spain and Paris in France. The figure depicts all industries. The data being used is the labor market panel introduced in section 3.

Figure 12: First stage for structural estimation (selected industries)



Notes: This figure shows the first stage regression predicting changes in employment size,  $\frac{L_{i,s,1920}}{L_{i,s,1914}}$  at the province-sector level, using  $\log distance_{i,Paris} \times \log(\text{Employment Share of Sector in Province})$  as an instrument, where  $\log(\text{Employment Share of Sector in Province}) \equiv \frac{L_{i,s,1914}}{\sum_r L_{i,r,1914}}$  works as a labor supply shifter and is interacted with distance to Paris as a reduced form proxy for differences in geographical advantages vis-a-vis the French destination market. Distance is calculated using the shortest path along a network of railroads and maritime linkages between province capitals in Spain and Paris in France. The figure depicts all industries. The data being used is the labor market panel introduced in section 3.

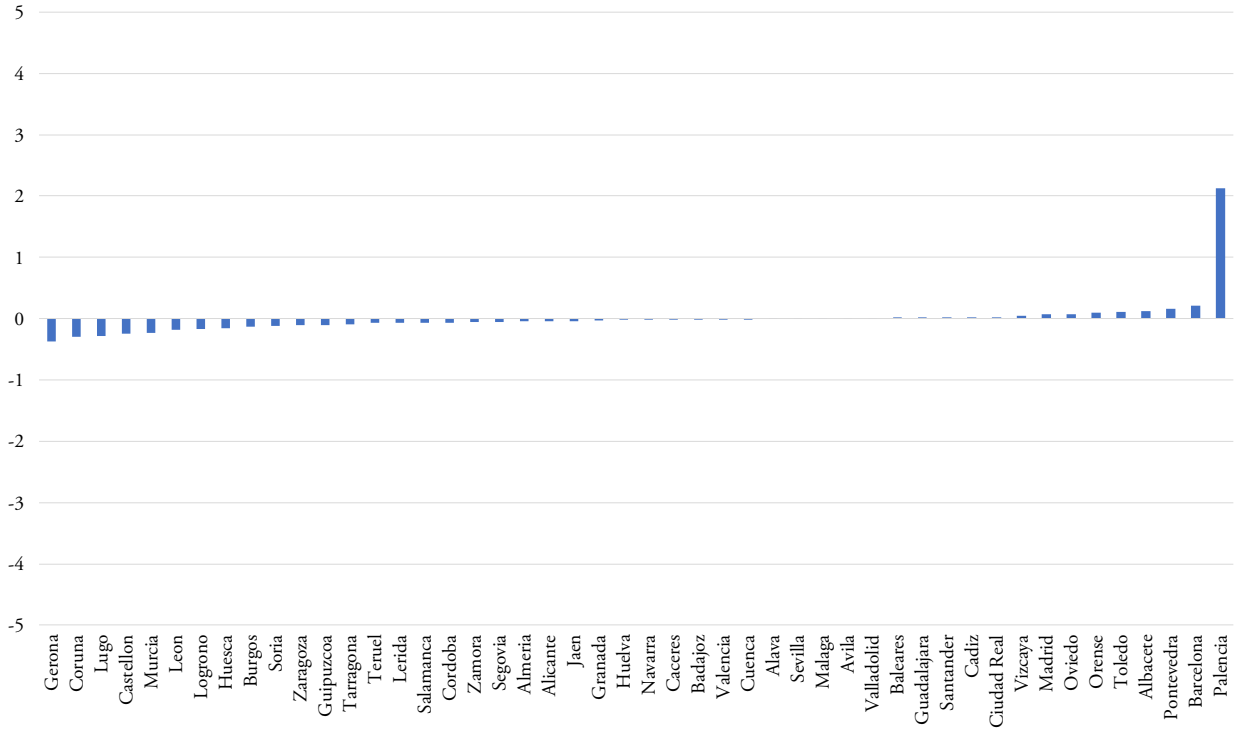
Figure 13: Counterfactual: National industry size (Employment levels and differences, 1920)



Notes: The upper graph depicts the aggregate sectoral employment for the observed data and the counterfactual simulation of Spain in the absence of WWI. The values are constructed in the following way,  $L_s \equiv \sum_i L_{i,s,1920}$  where  $L_{i,s,1920}$  refers to the observed employment size in province  $i$  and sector  $s$ . Similarly for the counterfactual,  $L_s^{CF} \equiv \sum_i L_{i,s,1920}^{CF}$  where  $L_{i,s,1920}^{CF}$  is the simulated counterfactual sectoral employment size using the estimated model as described in section 7. The lower graph shows the same figure in terms of difference between the counterfactual and observed data.

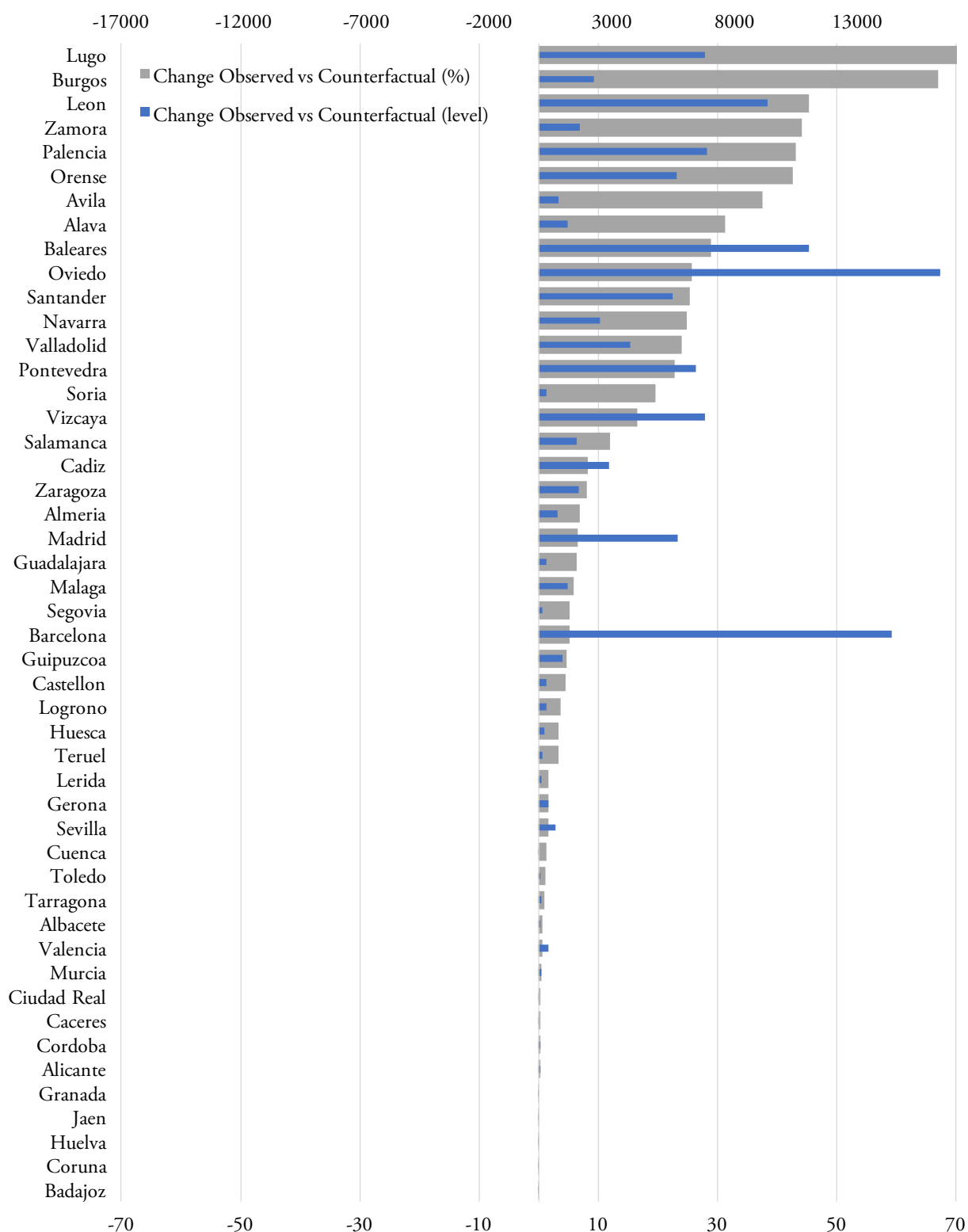


Figure 14: Counterfactual: Province size (Manufacturing and Agricultural Employment, 1920)



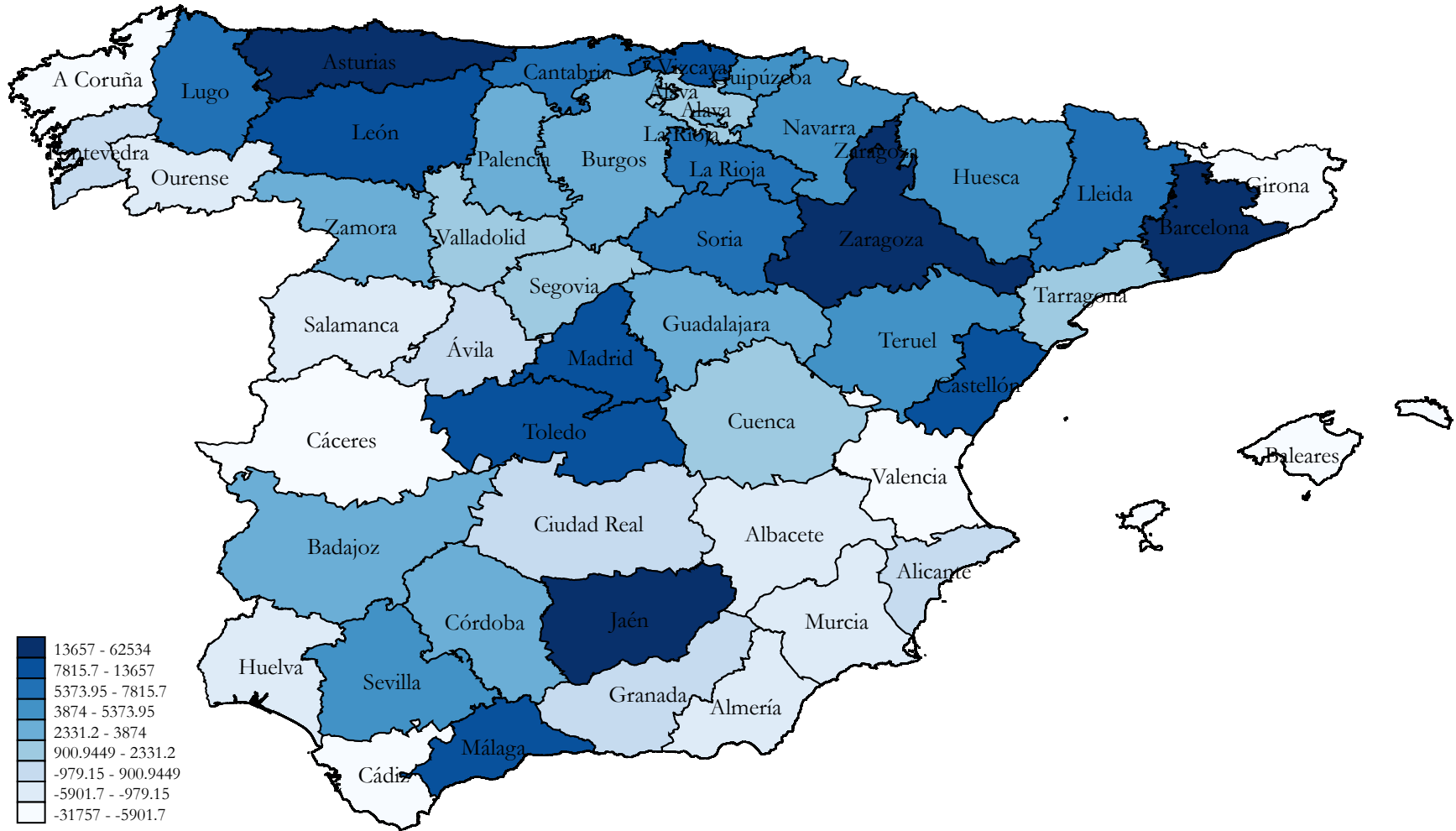
Notes: The figure depicts the change in aggregate provincial employment for the observed data and the counterfactual simulation of Spain in the absence of WWI. The values are constructed in the following way,  $L_s \equiv \sum_s L_{i,s,1920}$  where  $L_{i,s,1920}$  refers to the observed employment size in province  $i$  and sector  $s$ . Similarly for the counterfactual,  $L_s^{CF} \equiv \sum_s L_{i,s,1920}^{CF}$  where  $L_{i,s,1920}^{CF}$  is the simulated counterfactual sectoral employment size using the estimated model as described in section 7. Relative changes are indicated and calculated using those variables

Figure 15: Counterfactual: Manufacturing Employment (1920)



Notes: The figure depicts the change in manufacturing employment aggregated at the provincial level between the observed data and the counterfactual simulation of Spain in the absence of WWI. The values are constructed in the following way,  $L_s \equiv \sum_s L_{i,s,1920}$  where  $L_{i,s,1920}$  refers to the observed employment size in province  $i$  and sector  $s$ . Similarly for the counterfactual,  $L_s^{CF} \equiv \sum_s L_{i,s,1920}^{CF}$  where  $L_{i,s,1920}^{CF}$  is the simulated counterfactual sectoral employment size using the estimated model as described in section 7. Absolute and relative changes are indicated and calculated using those variables. The upper axis gives the relevant scale for absolute changes, while the lower axis gives the relevant scale for relative changes.

Figure 16: Counterfactual: Nominal Income Gains (1920)



59

Notes: The map depicts the provincial differences in nominal income in the observed data compared to the counterfactual without the war shock. The values for each province is calculated by first computing the nominal wages at the province level, that is  $Y_{i,1920} = \sum_r w_{i,r,1920} L_{i,r,1920}$  and subtracting the income in the counterfactual scenario,  $Y_{i,1920}^{CF} = \sum_r w_{i,r,1920}^{CF} L_{i,r,1920}^{CF}$ .

## C Data Sources

- *Censo de la población de España según el empadronamiento hecho en la península e islas adyacentes el 31 de diciembre de 1910* (**Instituto Geográfico; 1912**)
  - This publication contains population data disaggregated by profession for each province of Spain in 1910.
- *Censo de la población de España según el empadronamiento hecho en la península e islas adyacentes el 31 de diciembre de 1920* (**Instituto Geográfico; 1922**)
  - This publication contains population data disaggregated by profession for each province of Spain in 1920.
  - Furthermore, it also contains data on the origin of residents in each province that were born in another province.
- *Censo de la población de España según el empadronamiento hecho en la península e islas adyacentes el 31 de diciembre de 1930* (**Instituto Geográfico; 1932**)
  - This publication contains population data disaggregated by profession for each province of Spain in 1930.
  - Furthermore, it also contains data on the origin of residents in each province that were born in another province.
- *Estadística general del comercio exterior de España con sus posesiones de ultramar y potencias extranjeras (de Aduanas; 1910-1930)*
  - This publication contains trade records decomposed along destination countries and product type.
- *Estadística de salarios y jornadas de trabajo referida al periodo 1914-1925* (**Ministerio de Trabajo; 1927**)
  - This publication contains wage and quantity data by profession between for 1914, 1920 and 1925
- *Clasificación general de industrias, oficios y comercios 1931* (**Instituto Nacional de Prevision Social; 1930**)
  - This publication contains the official correspondence between industries and occupations.

## D Derivations

### D.1 Proof of Proposition 1

Let  $J$  be the number of alternatives. Depending on the values of the vector  $\epsilon = \{\epsilon_1, \dots, \epsilon_J\}$  the function  $\max_i(\delta_i \times \epsilon_i)$  takes on different values. First, examine the case where  $\max_i(\delta_i \times \epsilon_i) = \delta_1 \times \epsilon_1$ . That is, we

will integrate  $\delta_1 \times \epsilon_1$  over the set  $M_1 \equiv \{\epsilon : \delta_1 \times \epsilon_1 > \delta_j \times \epsilon_j, j \neq i\}$ :

$$\begin{aligned}
E_{\epsilon \in M_1} [\max_i (\delta_i \times \epsilon_i)] &= \\
&\int_{-\infty}^{\infty} (\delta_1 \times \epsilon_1) f(\epsilon_1) \left[ \int_{-\infty}^{\frac{\delta_1 \times \epsilon_1}{\delta_2}} \dots \int_{-\infty}^{\frac{\delta_1 \times \epsilon_1}{\delta_J}} f(\epsilon_2) \dots f(\epsilon_J) d\epsilon_2 \dots d\epsilon_J \right] d\epsilon_1 = \\
&\int_{-\infty}^{\infty} (\delta_1 \times \epsilon_1) f(\epsilon_1) \left( \int_{-\infty}^{\frac{\delta_1 \times \epsilon_1}{\delta_2}} f(\epsilon_2) d\epsilon_2 \right) \dots \left( \int_{-\infty}^{\frac{\delta_1 \times \epsilon_1}{\delta_J}} f(\epsilon_J) d\epsilon_J \right) d\epsilon_1 = \\
&\int_{-\infty}^{\infty} (\delta_1 \times \epsilon_1) f(\epsilon_1) F\left(\frac{\delta_1 \times \epsilon_1}{\delta_2}\right) \dots F\left(\frac{\delta_1 \times \epsilon_1}{\delta_J}\right) d\epsilon_1
\end{aligned} \tag{5}$$

The final term in the last equation is the first of J such terms in  $E[\max_i (\delta_i \times \epsilon_i)]$ . Specifically,

$$E \left[ \max_i (\delta_i \times \epsilon_i) \right] = \sum_i E_{\epsilon \in M_i} \left[ \max_i (\delta_i \times \epsilon_i) \right]. \tag{6}$$

Now we apply the functional form of the Fréchet distribution, where the CDF is given by  $F(x) = e^{-x^{-a}}$ , and the PDF is given by  $f(x) = ax^{-1-a}e^{-x^{-a}}$ , where  $a$  is the dispersion parameter. This gives,

$$\begin{aligned}
&E_{\epsilon \in M_i} \left[ \max_i (\delta_i \times \epsilon_i) \right] \\
&= \int_{-\infty}^{\infty} (\delta_i \times \epsilon_i) a \epsilon_i^{-a-1} e^{-\epsilon_i^{-a}} e^{-\left(\frac{\delta_i \epsilon_i}{\delta_2}\right)^{-a}} \dots e^{-\left(\frac{\delta_i \epsilon_i}{\delta_J}\right)^{-a}} d\epsilon_i \\
&= \int_{-\infty}^{\infty} (\delta_i \times \epsilon_i) a \epsilon_i^{-(a+1)} \prod_j e^{-\left(\frac{\delta_i \epsilon_i}{\delta_j}\right)^{-a}} d\epsilon_i \\
&= \int_{-\infty}^{\infty} (\delta_i \times \epsilon_i) a \epsilon_i^{-(a+1)} \exp\left(\sum_j -\left(\frac{\delta_i \epsilon_i}{\delta_j}\right)^{-a}\right) d\epsilon_i \\
&= \int_{-\infty}^{\infty} (\delta_i \times \epsilon_i) a \epsilon_i^{-(a+1)} \exp\left(\epsilon_i^{-a} \times \sum_j -\left(\frac{\delta_i}{\delta_j}\right)^{-a}\right) d\epsilon_i
\end{aligned} \tag{7}$$

where the second step comes from collecting one of the exponentiated terms into the product, along with the fact that  $\delta_j/\delta_i = 1$  if  $i = j$ . Now we define  $D_i \equiv \sum_j \left(\frac{\delta_i}{\delta_j}\right)^{-a}$  and make the substitution  $x = D_i \epsilon_i^{-a}$  so that  $dx = -a \epsilon_i^{-a-1} D_i d\epsilon_i \Rightarrow -\frac{dx}{D_i} = a \epsilon_i^{-(a+1)} d\epsilon_i$  and  $\epsilon_i = \left(\frac{x}{D_i}\right)^{-\frac{1}{a}}$ . Note that as  $\epsilon_i$  approaches infinite,  $x$  approaches 0, and as  $\epsilon_i$  approaches negative infinity,  $x$  approaches infinity.

$$\begin{aligned}
&E_{\epsilon \in M_i} \left[ \max_i (\delta_i \times \epsilon_i) \right] = \\
&\int_{\infty}^0 \left( \delta_i \left(\frac{x}{D_i}\right)^{-\frac{1}{a}} \right) \left(-\frac{1}{D_i}\right) \exp\{-x\} dx \\
&= \frac{1}{D_i} \int_0^{\infty} \left( \delta_i \left(\frac{x}{D_i}\right)^{-\frac{1}{a}} \right) e^{-x} dx
\end{aligned} \tag{8}$$

Recall that  $D_i \equiv \sum_j \left(\frac{\delta_i}{\delta_j}\right)^{-a} = \frac{\sum_j \delta_j^a}{\delta_i^a}$ . Notice that the familiar frechet choice probabilities  $P_i = \frac{\delta_i^a}{\sum_j \delta_j^a}$  are inverses of the  $D_i$ 's or in other words  $P_i = 1/D_i$ . Also note that  $\sum_i P_i = 1$ .

$$\begin{aligned} &= P_i \int_0^\infty \left( \delta_i \left( \frac{x}{D_i} \right)^{-\frac{1}{a}} \right) e^{-x} dx \\ &= P_i \delta_i D_i^{\frac{1}{a}} \int_0^\infty x^{\frac{1}{a}} e^{-x} dx \end{aligned}$$

The Gamma function is defined as  $\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx$ . This implies that the integral term is equal to  $\Gamma\left(1 - \frac{1}{a}\right)$ . Furthermore, since  $D_i^{1/a} = \frac{1}{\delta_i} \left(\sum_j \delta_j^a\right)^{\frac{1}{a}}$ , we obtain,

$$= \left( \sum_j \delta_j^a \right)^{\frac{1}{a}} \Gamma\left(1 - \frac{1}{a}\right) \times P_i$$

Finally summing over all alternatives,

$$\sum_i E_{\epsilon \in M_i} \left[ \max_i (\delta_i \times \epsilon_i) \right] = \left( \sum_j \delta_j^a \right)^{\frac{1}{a}} \Gamma\left(1 - \frac{1}{a}\right) \times \sum_i P_i = \left( \sum_j \delta_j^a \right)^{\frac{1}{a}} \Gamma\left(1 - \frac{1}{a}\right) \quad \square$$