# Macroeconomic Development, Rural Exodus, and Uneven Industrialization*

Tomás Budí-Ors
*EUI and Banco de España*

Josep Pijoan-Mas
*CEMFI and CEPR*

June 2025

**Abstract**

Economic development and industrialization are typically led by a few regions within a country. The initially laggard regions may catch up and industrialize —as in the U.S. 1880 to 1940— or they may fail to industrialize, experience a population exodus, and help industrialization elsewhere —as in Spain 1940 to 2000. To understand the relationship between internal migrations and uneven industrialization, we build a simple model of structural change with multiple locations and sectors where both internal migration and internal trade are costly. In the model, internal migrations change the relative labor demand across sectors at the local level and hence act as a force of within-region structural change and uneven paths to industrialization. We calibrate our economy to the development experience of Spain, and find that its large rural exodus and uneven regional industrialization were originated by a decline in migration costs towards the most industrial areas. More importantly, internal migration fully explains the lack of industrialization in laggard areas and accelerated growth at the aggregate level.

*JEL classification*: O41, O11, F16, N10, R11

*Keywords*: Growth, Structural Change, Internal Migration, Internal Trade

*The authors are grateful to comments by Paula Bustos, Nezih Guner, Berthold Herrendorf, Enrique Moral-Benito, Eduardo Morales, David Nagy, Fernando Parro, Michael Peters, Michael Waugh, Kei-Mu Yi, Miguel Zerecero, and Fabrizio Zilibotti, as well as by participants in the BSE Summer Forum, the ESSIM-CEPR conference, the Migration and the Macroeconomy conference at Bocconi, the STEG Theme Workshop, the UEA 12th European Meeting, the Winter Meetings of the Econometric Society, and seminars at Arizona State University, Bank of Spain, CEMFI, CREi, EUI, University of Connecticut, Universitat Jaume I, and University of Toronto. We also thank Marta Domínguez for superb research assistance. Berthold Herrendorf was an important inspiration for this paper, and he will be greatly missed. An earlier version of this paper has been presented under the title "Migration, Trade, and Structural Change". Emails: tomas.budiors@eui.eu, pijoan@cemfi.es

## 1 Introduction

The economic development of nations is characterized by a shift of employment away from agriculture and a slow cycle of industrialization and de-industrialization, see Herrendorf et al. (2014) for details. A salient feature of this process is that, at its early stages, agricultural and non-agricultural activities tend to happen in different locations within a country. For instance, industrialization was initially more intense in the North of France, the Northeast of the US, the Basque Country and Catalonia in Spain, or the Guangdong, Jiangsu, and Shanghai regions in China.

As a few regions in a country start to industrialize, different patterns of development may follow. On the one hand, the initially agrarian regions may experience large outmigrations flows towards the leading industrial hubs —a rural exodus— and fail to industrialize. We show this to be the case for Spain between 1940 and 2000, and we additionally document other development episodes with a large rural exodus like France (1872-1975) or China (2000-2015). On the other hand, the initially agrarian regions may not suffer large outmigration flows, may manage to catch up with the leaders, and also industrialize. Eckert and Peters (2022) show this to be the case for the US between 1880 and 1940, and we additionally document other development episodes without any rural exodus like Indonesia (1971-2010) or the Dominican Republic (1960-2010). In between these two polar patterns, one can also find a wealth of intermediate cases like India (1987-2011) or Brazil (1980-2010).

The goal of this paper is to uncover the economic forces shaping the emergence of the different patterns of development in terms of migrations and local industrialization. In particular, we show theoretically and quantitatively how internal migrations —arising from changes in spatial frictions or in region-specific sectoral productivities— are a key determinant of the sectoral structure of the local economy.

Our case study is the relatively recent development experience of Spain, which provides invaluable sector- and location-specific data over a complete process of economic development. We start by documenting that, from the 1950s, Spain experienced fast growth in income per capita, a strong reallocation of employment across sectors and space, a lack of industrialization in many regions, and an inverted-U shape evolution of spatial inequality. In particular, initial rural regions lost population in net terms (while the country almost doubled in size) and failed to create industrial jobs. At the same time, these initial rural regions created jobs outside agriculture in the (less tradable) services sector, and they eventually (partially) converged in income per capita with the rest.

Next, we build a model of structural change with multiple locations and sectors, where both the reallocation of workers across regions and the trade of goods and services are costly. In particular, we start from a simple model of structural change similar to Duarte and Restuccia (2010) where we add costly migration as in Artuç et al. (2010) and costly trade as in Eaton and Kortum (2002). This framework provides a high degree of tractability and is amenable to quantitative work. When spatial frictions tend to infinity, each region is like an independent closed economy. Sectoral reallocation at the regional level is driven by the standard income and price effects due to symmetric and asymmetric sectoral productivity growth within the region.[1] Sectoral reallocation at the country level is the sum of what happens in all its independent regions. Outside this limit case, different regions specialize in the production of different sectors following their comparative advantage and people move across regions following economic opportunities.

We first use the model to show theoretically how internal migration flows represent an engine of structural change at the local level by changing the sectoral composition of local labor demand. First, holding local wages fixed, we show that the elasticity of sectoral labor demand with respect to a population change is given by the share of local expenditure in the sector's local production. This implies that, in areas losing (gaining) population, labor is reallocated away from (toward) sectors that are more dependent on local consumers.[2] Second, a population outflow generates a local wage increase because, with trade, the decline in labor demand is lower than the decline in labor supply —since part of labor demand depends on consumers elsewhere. We demonstrate that the wage increase reduces labor demand more in sectors that are more exposed to trade, thereby inducing a reallocation of employment away from more tradable sectors. And third, the wage increase leads to sectoral reallocation due to the different income elasticity of sectoral demands. In particular, by increasing local wages compared to the case where the population does not move, outmigration accelerates the reallocation of employment away from agriculture and towards manufacturing and services. Overall, which force dominates is ultimately a quantitative question.

We then bring the model to the data through the development experience of Spain. We have time series data at the province and sector level for value added, prices, and employment, plus data on bilateral migration flows across provinces. These data allow to recover the time paths for sector-province productivity; sector-province-of-origin trade costs; and bilateral migration costs. Our strategy to achieve this identification represents a methodological contribution. The use of models with costly internal trade in the macro-development literature has been halted by the lack of time series data for internal bilateral trade flows, which is the information typically used to recover bilateral trade costs. We argue that in multi-sector models, the difference between sectoral employment shares and sectoral expenditure shares at the regional level reveals information, other things equal, about the tradability of each sector. Our strategy, similar to Eckert (2019), builds on Gervais and Jensen (2019), who show how to use region-industry trade surpluses to infer the trade costs for different goods and services in the US. We extend this logic to identify both sector-specific and region-specific trade costs that vary over time. Absent time series data for sectoral expenditure shares at the regional level, one can use the predictions from an estimated demand system, which is what we do.

[1] See Kongsamut et al. (2001) for the income effect and Baumol (1967), Ngai and Pissarides (2007) for the price effect. Duarte and Restuccia (2010), Herrendorf et al. (2013), Boppart (2014), Comin et al. (2021), and García-Santana et al. (2021) combine both mechanisms as we do here.

[2] This relates to Burstein et al. (2020), who show how heterogeneity in the tradability of an occupation's output determines its exposure to the effects of immigration shocks.

The calibrated model for Spain between 1940 and 2000 displays four main aggregate trends: an increase in sectoral productivities, a decline in trade costs, a non-monotonic evolution in the spatial dispersion of sectoral productivities, and a non-monotonic evolution of migration costs. The growth in productivity, mostly between 1950 and 1990, is largest in agriculture and smallest in services. The decline in trade costs is more apparent for agriculture than for manufacturing and services. These two patterns together generate an increase in aggregate productivity, a trend increase in the price of services relative to manufactures, and a trend decline in the price of agriculture relative to manufactures, which combined account for the bulk of structural change and aggregate growth. Sectoral productivity in manufacturing and services diverged across regions during the first decades and started to converge afterwards. This pattern led to the inverted-U shape evolution of income inequality across regions and contributed to the rural exodus by amplifying inequalities in the first half of the development process. Migration costs towards the regions experiencing the largest population gains fell during the 1950s, 1960s, and 1970s, coinciding with large government-funded cheap housing projects on the outskirts of big cities. This was a driver of the rural exodus, as differences in income per capita across regions were not enough to explain the large population movements. After that, migration costs increased everywhere, suggesting that there were still differences in real incomes across provinces which were not arbitraged away.

Next, we show that the large rural exodus of workers looking for better economic opportunities had a first order impact in the development process of each province while having moderate aggregate effects. To do so, we solve for two counterfactual economies without migration (infinite migration costs over all the period). In the first one, region-specific sectoral productivity paths remain as in the benchmark economy and do not react to population movements. In the second one, we allow for agglomeration economies in industry and decreasing returns to scale in agriculture such that region-specific sectoral productivity paths are affected by migrations.

Starting with the simpler case, we find that absent migrations, lagging regions do industrialize and leading regions specialize less in manufacturing. Indeed, there is structural change everywhere and the economic structure of regions looks more similar to each other. This result implies that the evolution of region-specific sectoral productivity and trade costs was conducive to industrialization in the initially rural areas, but the rural exodus impeded it. The reason is twofold. On the one hand, the relatively unproductive manufactures of laggard areas were very much dependent on local demand, which declined due to the outmigrations. On the other hand, migrations reduced the wage gap between leading and laggard regions, which allowed leading areas to remain competitive in the highly tradable manufacturing sector. This result is amplified when we allow for non-constant returns to scale in production because the rural exodus enhances industrial productivity in the leading regions at the expense of the laggard ones, an advantage that is inverted when there are no migrations. The aggregate effects of labor reallocation through the rural exodus are modest. In the counterfactual economy with no migration, GDP per capita would have grown 38 percentage points less (in a period in which output increased eightfold). When we allow productivity to react to migrations, the aggregate effects are similar. Therefore, we conclude that outmigrations were the key determinant of the lack of industrialization in laggard regions and moderately accelerated growth at the aggregate level.

Finally, we revisit the cross-country heterogeneity in development experiences through the lens of our analysis of Spain. We find that if migration costs had not changed over time (instead of falling), the pattern of migrations in Spain would have been similar to the US (1880-1940) —with initial agrarian regions not losing population in a systematic manner—, industrialization would have been more homogenous across regions, and GDP growth would have been lower. The reasons are as follows. Absent the fall in migration costs, more people remain in laggard areas. This favors their industrialization through the effects on local demand (which does not fall) and equilibrium prices (which prevent leading regions from exploiting their productivity advantage), but limits aggregate growth. Instead, if sectoral productivities had converged across regions since 1940 (rather than diverging) the rural exodus would have been similar to the one observed, but industrialization would have been much more even across regions.

### 1.1 Related literature

Our work relates to the literature on growth and structural change with spatial frictions. This literature has traditionally focused on frictions to either internal migration —see Garriga et al. (2017) or Ngai et al. (2019)— or internal trade —see Adamopoulos (2011), Herrendorf et al. (2012), or Gollin and Rogerson (2014)— by use of stylized models with two sectors (agriculture and non-agriculture) and two locations (rural and urban). Compared to this literature, we emphasize the interactions between migration and trade frictions and consider a quantitative exercise with multiple regions, three sectors, and lack of full sectoral specialization across regions. Caselli and Coleman (2001) analyzed the structural transformation of the US between 1880 and 1950, being the first paper in this literature to focus on its spatial dimension. In their model, agricultural workers need to acquire costly education to work in manufacturing. The decline in education costs acts as a decline in migration costs, promoting migration and structural change, a mechanism also highlighted by Porzio et al. (2022). We document that changes in education were not an important driver of rural migrations in Spain, see Appendix C.

Two recent papers are very much related to our research. First, Hao et al. (2020) study the growth episode in China between 2000 and 2015, which was characterized by large migrations from rural to urban locations and the reallocation of employment away from agriculture. Building on Tombe and Zhu (2019), they set up a model with both internal trade costs and migration costs as well as multiple locations. They show that the reduction in migration costs (the *hukou* system reform) and internal trade costs account for a substantial fraction of macroeconomic changes and spatial location of economic activity. Compared to this paper, we explore the impact of spatial frictions in the long run following a complete process of economic modernization (with a full cycle of industrialization and de-industrialization), and we emphasize the effects of migrations on the sectoral composition of local employment. The second one is Eckert and Peters (2022), who explore the role of migrations across counties in the US between 1880 and 1940. Different from Spain, they find that migration flows in the US do not follow any sectoral pattern. They write a model with constant spatial frictions and focus on the productivity catch-up between locations that makes rural areas industrialize instead of losing population. We allow both productivity and spatial frictions to change over time and discuss how they interact to generate the patterns of migrations and local industrialization observed in Spain.

Our paper highlights the effects of migrations on equilibrium regional wages and prices, and hence on the sectoral specialization of regions. In an extension to our main exercise, we also allow for regional migrations to affect sectoral productivity through decreasing returns to scale in agriculture and agglomeration economies in industry. The literature on growth and structural change has already studied how agglomeration economies in urban locations increase the productivity of certain sectors, see for instance Michaels et al. (2012), Desmet and Rossi-Hansberg (2014), Eckert (2019), or Nagy (2020). Different from ours, these papers put their focus on urbanization or the rise of the modern services economy.

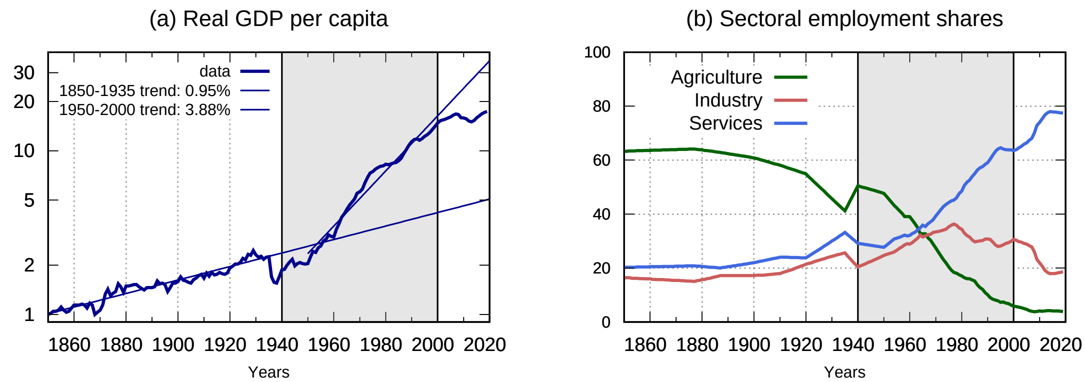
In terms of methodology, our work connects with two different literatures. First, Uy et al. (2014), Świecki (2017), Sposi (2019), or Lewis et al. (2021) among others use the Eaton and Kortum (2002) framework to study the effects of international trade on structural change. The effects of asymmetric trade cost reductions across sectors on relative prices and hence on structural change are already highlighted in these papers. We use this same framework to study internal trade and explore the interaction of trade and migration frictions, which these papers do not consider. And second, we relate to the literature that looks at the local labor market effects of aggregate shocks, say international trade shocks —see Artuç et al. (2010) or Caliendo et al. (2019)— or changes in migration costs —see Morten and Oliveira (2018). All these papers feature costly trade and/or costly migration across different locations within a country. Our work is different as it focuses on the long-run process of macroeconomic development.

Lastly, our paper also relates to a recent strand of work zooming into the development experience of selected countries. These country studies are useful for understanding specific factors affecting economic development and drawing lessons for current developing economies. For instance Song et al. (2011) study the case of China; Uy et al. (2014) the case of South Korea; Cheremukhin et al. (2017) the case of Russia; Fajgelbaum and Redding (2022) the case of Argentina; and Fan et al. (2021) the case of India. The development experience of Spain is one of the few success stories in the second half of the XXth Century outside Asia, and yet it has attracted little attention.

## 2 The Spanish development experience

Our period of study is 1940 to 2000. Spain started its industrialization process well before this, but it was only after 1950 that the process accelerated to become an episode of fast economic

Figure 1: The Spanish development experience




**Notes:** Panel (a) reports GDP per capita in real terms, normalized to 1 in 1850, and in log scale. Panel (b) reports the share of employment in each sector. Data from Prados de la Escosura (2017). The grey areas represent our period of study.

development leading to convergence with Europe.[3] Between 1850 and 1935, real GDP per capita increased steadily at an average rate of almost one percent per year, while the share of employment in agriculture fell from 63.3 to 41.2 percent, see Panels (a) and (b) in Figure 1. The Civil War between 1936 and 1939 interrupted this process. Economic development regained vigor in the 1950s: between 1950 and 2000 the average rate of growth increased to 3.88% per year, the share of employment in agriculture dropped from 50.5 to 5.9 percent, the share of employment in services increased from 29.1 to 63.5 percent, and the share of industrial employment experienced a classic hump over time, with an increase from 20.4 to 36.3 percent between 1940 and 1977. Despite the Spanish political idiosyncrasies of the period, this pattern of sectoral reallocation is standard among development episodes, see Herrendorf et al. (2014).[4,5]

Next, we zoom in on the spatial dimension of the development process. To do so we consider data at the province level, which is the smallest administrative unit for which we can gather information on employment, productivity, and prices at the sectoral level.[6] We identify three main facts. First, between 1940 and 2000 the country experienced massive migrations across provinces,

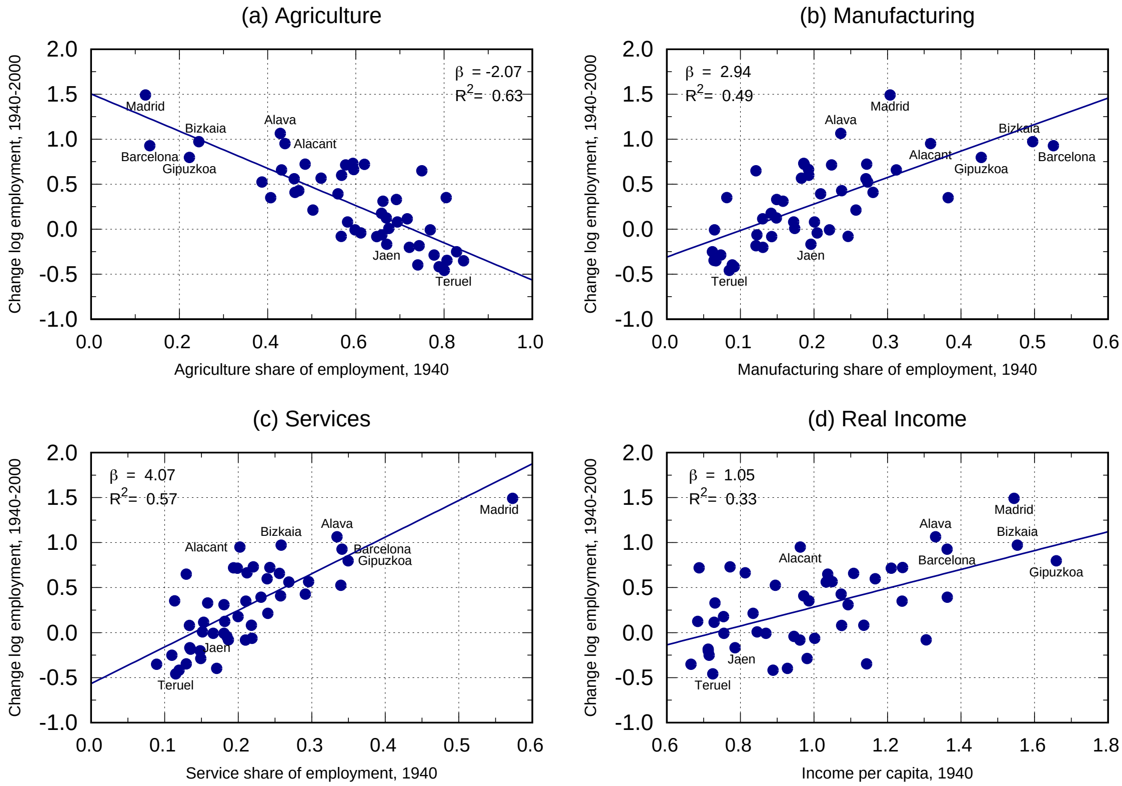
[3] We stop our analysis in 2000 because the process of economic modernization was already completed, and because of the peculiar characteristics in the years that followed, with the twin credit and housing booms and the ensuing severe recession. For the post-2000's growth process in Spain, see Díaz and Franjo (2016), Gopinath et al. (2017), García-Santana et al. (2020), Martín et al. (2021), and Almunia et al. (2021).

[4] See Prados de la Escosura and Rosés (2021) for a detailed account of the growth process in Spain since 1850.

[5] A salient feature of our period of study is a gradual opening to international trade. The process started in 1953 with the imports of machinery to modernize the industrial sector. In 1959, the "*Plan de Estabilización*" implemented a set of structural reforms, including the lifting of several restrictions to international trade. However, the recent work by Campos et al. (2022) estimates that the effective pace of trade liberalization in Spain between 1948 and 1985 was no different from the one experienced by the rest of European countries, and in terms of levels the Spanish border effect was comparable to the countries in the communist bloc. See Conesa et al. (2021) for a study of the role of international trade on the industrialization process of Spain since 1850.

[6] The territorial division of Spain into provinces dates back to 1833, with minimal changes since then. It is roughly equivalent to NUTS3 classification of Eurostat. The median employment of Spanish provinces was 135,000 workers in 1940 (9.2 million for the whole country) and 208,000 in 2000 (16.3 million for the whole country). To be consistent with the calibrated model later on, we use data for the 47 mainland provinces in the Iberian Peninsula (i.e. dropping data for the Canary and Balearic islands). Provinces are hence larger than local labor markets, and rural to urban migrations also happened within provinces. This is a dimension that we overlook in our study.

Figure 2: Employment growth and initial sectoral composition



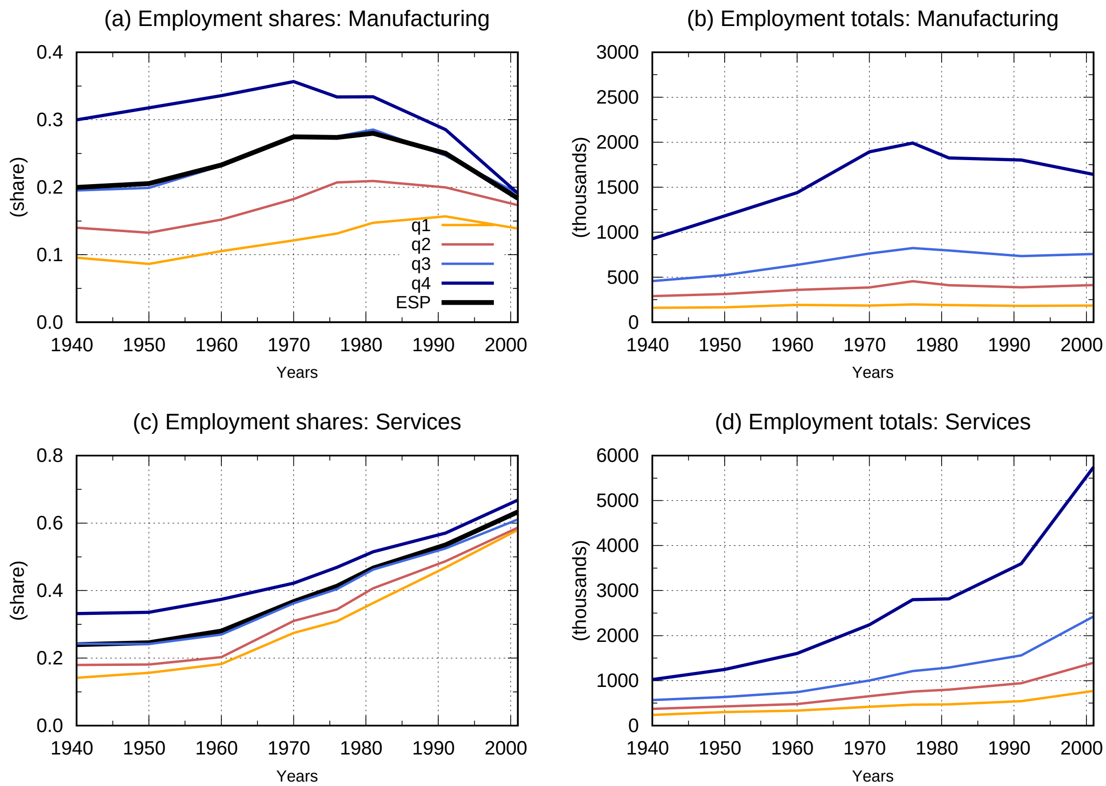

**Notes:** This figure plots the relative increase in employment between 1940 and 2000 (in logs) for all provinces, against the 1940 sectoral shares of employment, Panels (a) to (c), and against real income per capita (relative to the country average), Panel (d). Income per capita at province level deflated with the provincial price indices obtained with the model calibration of Section 4. Each panel also reports the slope of the relationship ($\beta$) and the share of variance in log employment growth explained by the corresponding x-axis variables ($R^2$).

mainly from poor agrarian provinces towards richer and more industrial or service-intensive ones.[7] In Figure 2, we plot the relative increase in employment between 1940 and 2000 against its sectoral composition and income per capita in 1940. On the vertical axis of any Panel we can see that there was a large heterogeneity in employment growth across provinces. Some provinces like Madrid, Alava, or Barcelona multiplied their total employment by a large factor, between two and four (employment grew between 100 and 150 log points), while others like Jaén or Teruel lost population over these 60 years.[8] The correlation between the initial sectoral employment composition and subsequent employment growth is strong: provinces with a larger share of agrarian employment lost population (Panel a) while provinces with a larger share of manufacturing or services grew in size (Panels b and c). Our estimated relationships imply that a province with 10 percentage points higher share of employment in agriculture in 1940 experienced a 20% smaller population growth between 1940 and 2000. Remarkably, 63% of the variance in employment growth across provinces is

[7]Richer provinces in 1940 had more labor in industry and services and less in agriculture, see Panels (a) to (c) in Figure H.1. We also note that in 1940 there was little relationship between income per capita and total employment across provinces, see Panel (d) in Figure H.1.

[8]The process of internal migration in Spain over this period is well documented in Bover and Velilla (2005).

Figure 3: Industrialization and lack thereof, 1940-2000



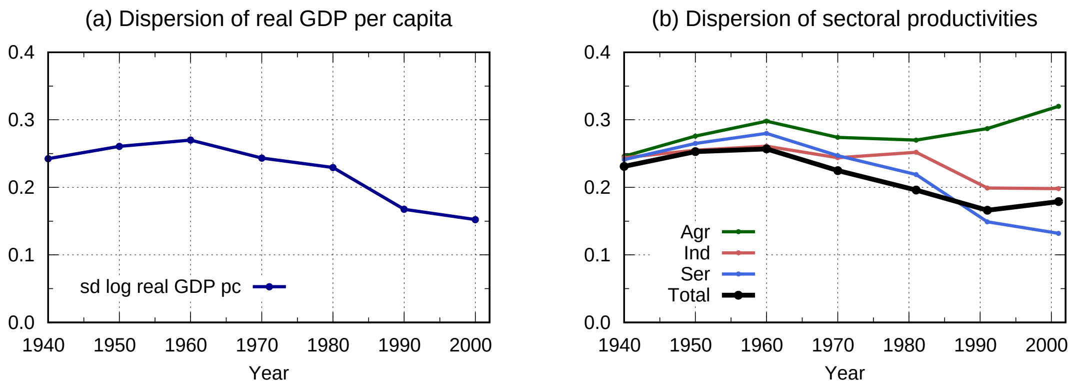

**Notes:** Panels (a) and (c) report the evolution of sectoral shares of employment in Manufacturing and Services while Panels (b) and (d) report the evolution of the number of workers in the same sectors. Provinces are grouped in four quartiles in terms of the change in their relative employment size within the country between 1940 and 2000. q1 corresponds to the bottom quartile (these provinces moved from 26.7% to 14.2% of total employment), q2 to the second (from 17.1% to 11.2%), q3 to the third (from 19.1% to 17.7%), and q4 is the top quartile (from 37.1% to 56.8%).

related to their initial agricultural share of employment. We also note that the correlation between initial income per capita and employment growth is strong (Panel d), which should not be surprising given the correlation between income per capita and sectoral specialization in 1940.

We can compare this pattern of internal migration to that of the development experience of other countries. To do so, we resort to the IPUMS International Census Database (and other sources) and gather data on employment at the region-sector level for 27 countries that experience a development episode. We find several development episodes with a rural exodus similar to Spain, as for instance France (1872 to 1975), Greece (1971 to 2011) and China (2000 to 2015). In contrast, in salient development episodes like the US (1880 to 1940), the Dominican Republic (1960 to 2010) or Indonesia (1971 to 2010), the initial sectoral composition of regions, while negatively related to regional employment growth, only explains a tiny fraction of its variation. In between, we have intermediate cases like India (1987 to 2015), Brazil (1980 to 2010), or Turkey (1985 to 2000). Further details can be found in Appendix D.

Second, the country experienced a process of industrialization that was largely uneven across regions. It is easy to overlook this fact because the evolution of relative industrial employment

Figure 4: Dispersion of income and productivity across provinces, 1940-2000




**Notes:** Panel (a) reports the standard deviation of the log of real income per capita. Panel (b) reports the standard deviation of the log of provincial productivities in each sector. Income per capita at province level is deflated with the provincial price indices obtained with the model calibration of Section 4.

within every province mimics the classic hump-shaped pattern. Yet, when one looks at total instead of relative employment by sector, it turns out that the initially agrarian regions did not industrialize: they just lost agrarian population who moved to the non-agrarian sectors in other provinces. This is illustrated in Figure 3. We classify all provinces into four groups according to the change in their relative size (employment) within the country between 1940 and 2000, see Appendix B for details. The provinces in the bottom quartile (q1) almost halved their size (shrinking from 26.7% to 14.2% of total employment in the country), while provinces in the top quartile (q4) experienced a large increase (from 37.1% to 56.7%). In Panel (a) we see how the four groups of provinces experienced a hump in the *share* of employment in manufacturing similar to the one for the overall country. Yet, in Panel (b) we see that provinces in the q1 and q2 groups hardly experienced any increase in the *levels* of manufacturing employment, so these provinces did not really industrialize. It is important to note that regions in the q1 and q2 groups did create non-agriculture jobs in net terms despite the outmigrations, though job creation was concentrated in the much less tradable construction and services sectors. For instance, Panels (c) and (d) in Figure 3 show the evolution of relative and total employment in services. This shows net employment growth in services, even in provinces in the q1 and q2 groups. Figure H.8 in Appendix H shows the same pattern for selected provinces within q1 and q2 groups.

Finally, spatial inequality in income per capita followed a classic inverted-U shape over time, increasing until 1960 and declining afterward, see Panel (a) in Figure 4. This path of inequality was first conjectured by Kuznets (1955) for individual income and later documented by Williamson (1965) or Lessmann (2014) among others for regional income. We also find that measured sectoral productivity diverged across provinces in the early decades, but it started to converge in 1970. This was mostly apparent in aggregate and services productivity, while agricultural productivity started to diverge again in the 1990s, see Panel (b) in Figure 4.

# 3 Model

We consider an economy with three sectors $j = a, m, s$ (agriculture, manufacturing, and services), and several regions $r = 1, 2, ..., R$. Within each sector there is a continuum of varieties indexed by $x \in [0, 1]$, which can be produced in any region by competitive producers that use labor $L$ as the only input of production. Productivity is region-, sector-, and variety-specific, and varieties are tradable between regions subject to trade costs. Labor can be costlessly reallocated across varieties and sectors within a region, but not across regions, as migration entails costs. This implies that there is a unique labor market within regions but segmented labor markets across regions. The model is static except for the law of motion of labor supply, which relates the distribution of population across regions at $t - 1$ and $t$ through the bilateral migration flows between the two periods. In addition, in the quantitative exercise of Section 4 we will allow several parameters to change over time, which will hence carry a $t$ subindex that we omit in this Section unless necessary.

## 3.1 Consumption and migration

At the beginning of every period, workers decide their region of work based on regional differences in real income, bilateral migration costs, and their idiosyncratic preference for each location. Once established in a region, they work and consume in that region. We can characterize their choices in a two-stage optimization problem, which we solve by backwards induction.

**Consumption decisions.** Workers in region $r$ have preferences for the consumption of goods in each sector $j$, denoted $c_{rj}$. In particular, preferences over $c_{rj}$ are characterized by a consumption basket $c_r$ that is implicitly defined by the following non-homothetic CES aggregator (Comin et al., 2021):

$$\sum_j \omega_j^{\frac{1}{\nu}} \left( \frac{c_{rj}}{c_r^{\epsilon_j}} \right)^{\frac{\nu-1}{\nu}} = 1 \tag{1}$$

where $\omega_j > 0$ are the CES taste parameters, $\nu \in (0, 1) \cup (1, \infty)$ is the constant elasticity of substitution between goods, controlling how changes in relative prices affect relative sectoral expenditures, and $\epsilon_j > 0 \ \forall j$ are parameters that introduce non-homotheticities in demand, controlling how changes real income affect relative sectoral expenditures. The budget constraint for a worker living in region $r$ is given by

$$P_{ra}c_{ra} + P_{rm}c_{rm} + P_{rs}c_{rs} \leq w_r \tag{2}$$

where $P_{rj}$ is the price index in region $r$ for sector $j$ final good $c_{rj}$ and $w_r$ is the nominal wage in region $r$. Workers maximize CRRA utility $U = \left(c_r^{1-\sigma} - 1\right) / (1 - \sigma)$, with $c_r$ implicitly defined by (1), subject to their budget constraint (2). This implies the following expenditure share on sector

$j$ goods by consumers in region $r$:

$$\frac{P_{rj}c_{rj}}{w_r} = \omega_j \left(\frac{w_r}{P_{rj}}\right)^{\nu-1} c_r^{(1-\nu)\epsilon_j} \tag{3}$$

where, given the budget constraint (2), $w_r$ is both the wage and the total expenditure per worker of region $r$. Given the regional wage $w_r$ and prices $P_{ra}$, $P_{rm}$, and $P_{rs}$, the consumption basket $c_r$ and the sectoral demands $c_{ra}$, $c_{rm}$, and $c_{rs}$ can be found combining equation (1) and equation (3) for all sectors.[9] Then, substituting the consumption basket into the utility function delivers the indirect utility that workers in region $r$ derive from consumption:

$$\mathcal{V}(w_r, P_{ra}, P_{rm}, P_{rs}) = \frac{\left[c_r(w_r, P_{ra}, P_{rm}, P_{rs})\right]^{1-\sigma} - 1}{1-\sigma}\,. \tag{4}$$

**Migration decisions.** At the beginning of each period, every worker in region $\ell$ decides to locate in the region $r$ that offers her the highest migration value $V^i_{\ell r} = \mathcal{V}(w_r, P_{ra}, P_{rm}, P_{rs}) - \mu_{\ell r} + \kappa\xi^i_r$, where $\mu_{\ell r}$ is a *route*-specific migration cost, capturing a notion of connectivity between regions $\ell$ and $r$, and $\xi^i_r$ is a taste shock that individual $i$ experiences for region $r$. This idiosyncratic shock is i.i.d. across regions and individuals, and captures the idea that workers may decide to live in a region for reasons unrelated to economic conditions. The extent to which idiosyncratic preferences determine migration patterns is therefore controlled by parameter $\kappa$. Assuming that $\xi^i_r$ is drawn from a Gumbel distribution, the share of workers of region $\ell$ that move to region $r$ between $t-1$ and $t$, $\rho_{\ell rt}$, is given by

$$\rho_{\ell rt} = \frac{\exp\left\{\frac{1}{\kappa}\big(\mathcal{V}(w_{rt}, P_{rat}, P_{rmt}, P_{rst}) - \mu_{\ell rt}\big)\right\}}{\sum_k^R \exp\left\{\frac{1}{\kappa}\big(\mathcal{V}(w_{kt}, P_{kat}, P_{kmt}, P_{kst}) - \mu_{\ell kt}\big)\right\}}\,. \tag{5}$$

This expression illustrates that workers move towards regions offering them higher real consumption (higher indirect utility) and to which it is easier to access (lower migration costs). It also shows that $\kappa$ limits the extent to which these economic forces determine the population of each region. The law of motion for the number of workers in each region is then given by

$$L_{rt} = (1+n_t)\sum_{\ell}^{R} \rho_{\ell rt} L_{\ell t-1}, \tag{6}$$

where $L_{rt}$ denotes region $r$ population at time $t$, $L_{\ell t-1}$ denotes region $\ell$ population at time $t-1$, and $n_t$ is population growth between $t-1$ and $t$. The labor supply in region $r$ is thus given by

[9] Note that substituting equation (3) for all sectors into equation (1) yields a single non-linear equation in $c_r$:

$$\left[\sum_i \omega_i \left(P_{ri} c_r^{\epsilon_i}\right)^{1-\nu}\right]^{\frac{1}{1-\nu}} = w_r$$

the inflows of workers from all regions (including itself) that optimally choose to work in $r$ plus population growth.

### 3.2 Production and trade

The (non-tradable) sector $j$ final good consumed by region $r$ workers comes from the aggregation of (tradable) intermediate varieties $q_{rj}(x)$ available in the region. We consider a standard CES aggregator with an elasticity of substitution parameter $\eta > 0$. One can micro-found this with a final good production sector, but we can as well think of it as a preference aggregator.

**Production of varieties.** Within each sector $j$ of region $r$, varieties are produced with a constant returns to scale technology that only uses labor, $y_{rj}(x) = A_{rj}(x)L_{rj}(x)$, where $A_{rj}(x)$ denotes the productivity of region $r$ at producing variety $x$ of sector $j$. Following Eaton and Kortum (2002), productivity $A_{rj}(x)$ is a random variable drawn from an independent region- and sector-specific Fréchet distribution with c.d.f. $F_{rj}(A) = \exp\{-T_{rj}A^{-\theta_j}\}$. The parameter $T_{rj}$ is region- and sector-specific and controls the average level of regional efficiency in the production of sector $j$ varieties. Due to specialization, region $r$ will tend to be a net exporter of sector $j$ goods if $T_{rj}$ is high relative to $T_{rk}$ for $k \neq j$. The shape parameter $\theta_j$ is sector-specific and common across regions, and governs the dispersion of productivity in the production of sector $j$ varieties. More productivity dispersion in a given sector (lower $\theta_j$) will provide room to further specialization, increasing intrasectoral trade across regions.

**Firm optimization and trade.** Regional trade is subject to iceberg trade costs. This means that $\tau_{r\ell j} \geq 1$ units of sector $j$ varieties must be shipped from region $r$ to region $\ell$ such that one unit arrives to $\ell$. Under perfect competition, cost minimization by firms implies that the price of a variety $x$ in sector $j$ that is offered by region $r$ producers to region $\ell$ consumers is $p_{r\ell j}(x) = \frac{w_r}{A_{rj}(x)}\tau_{r\ell j}$ —the marginal cost of production times the cost of shipping sector $j$ goods from $r$ to $\ell$. Consumers in region $\ell$ only purchase variety $x$ of sector $j$ from the region that can provide it at the lowest price, such that the price $p_{\ell j}(x)$ they pay is $p_{\ell j}(x) = \min_{r\in\{1,\ldots R\}} p_{r\ell j}(x)$. Hence, the distribution of sector $j$ variety prices paid by region $\ell$ consumers is a distribution over minimum prices, i.e. an extreme value distribution. Taking advantage of the properties of the Fréchet distribution, Eaton and Kortum (2002) show that the price $P_{rj}$ of the sector $j$ composite good (the price index of sector $j$ varieties) in region $r$ is given by

$$P_{rj} = \gamma_j \left[\sum_{\ell}^{R} \left(w_\ell \tau_{\ell r j}\right)^{-\theta_j} T_{\ell j}\right]^{-1/\theta_j}, \tag{7}$$

where $\gamma_j = \Gamma\left(\frac{\theta_j+1-\eta}{\theta_j}\right)^{1/(1-\eta)}$, and $\Gamma(.)$ is the gamma function. It can be shown that the share of region $\ell$'s expenditure in sector $j$ varieties produced by region $r$ is given by

$$\pi_{r\ell j} = \frac{\left(w_r \tau_{r\ell j}\right)^{-\theta_j} T_{rj}}{\sum_k^R \left(w_k \tau_{k\ell j}\right)^{-\theta_j} T_{kj}} , \tag{8}$$

which is larger if region $r$ has a low wage $w_r$, low trade costs $\tau_{r\ell j}$ or high productivity $T_{rj}$.

**Aggregation.** In equilibrium, the relevant objects are magnitudes at the region-sector level. Aggregating over all varieties $x$, we can define total output $Y_{rj}$ of sector $j$ in region $r$ (in units of the $rj$ consumption composite good) as $Y_{rj} \equiv \int y_{rj}(x) p_{rj}(x) dx / P_{rj}$.

### 3.3 Equilibrium

**Definition.** Given an initial distribution of workers across locations, $L_r^0$, a static equilibrium consists of region-specific wages $\{w_r\}$; consumption baskets $\{c_r\}$; bilateral migration flows $\{\rho_{r\ell}\}$; bilateral sector-specific trade flows $\{\pi_{r\ell j}\}$; and region- and sector-specific output $\{Y_{rj}\}$, prices $\{P_{rj}\}$, consumption $\{c_{rj}\}$ and employment allocations $\{L_{rj}\}$ such that: (a) workers and firms make optimal decisions (equations (3), (5), (7), (8) hold); (b) region-sector goods markets clear:

$$P_{rj} Y_{rj} = \sum_{\ell}^{R} \pi_{r\ell j} P_{\ell j} C_{\ell j} \qquad \forall r, j , \tag{9}$$

and (c) regional labor markets clear:

$$L_r = L_{ra} + L_{rm} + L_{rs} \qquad \forall r , \tag{10}$$

where labor supply $L_r$ is given by equation (6) and labor demands $L_{rj} \equiv \int L_{rj}(x) dx$ are implied by equation (9), see below.

**Discussion.** The goods market clearing in equation (9) requires that the value of sector-$j$ output produced in region $r$ equals the value that all regions, including itself, purchase from region $r$, where aggregate consumption of the sector $j$ good in region $r$ is defined as $C_{rj} \equiv c_{rj} L_r$. The labor market clearing in equation (10) requires that the labor supply in each region $r$, as given by equation (6), equals the labor demanded by region $r$ producers of all three sectors. The labor demand in each region $r$ and sector $j$ can be easily characterized as follows. First, note that constant returns to scale and perfect competition in the production of varieties imply that total revenues must equal total costs for all firms in each sector and region. As labor is the only input of production, total costs are simply labor costs and we can write $p_{rj}(x) y_{rj}(x) = w_r L_{rj}(x)$. Aggregating over producers we obtain,

$$P_{rj}Y_{rj} = w_r L_{rj}\,. \tag{11}$$

Next, one can use equation (11) to substitute total revenues by labor income in the goods market clearing condition (9) and obtain an expression for labor demand in sector $j$ and region $r$ as

$$L_{rj} = \frac{1}{w_r}\sum_{\ell}^{R} \pi_{r\ell j} P_{\ell j} C_{\ell j} \quad \forall r,\, \forall j. \tag{12}$$

Finally, note that the static nature of the consumer problem (workers in region $r$ can neither save nor borrow so $\sum_j P_{rj}Y_{rj} = \sum_j P_{rj}C_{rj}$) and the goods market clearing conditions imply that trade balances in each region:

$$\sum_{j\in\{a,m,s\}} \sum_{\ell\neq r}^{R} \pi_{r\ell j} P_{\ell j} C_{\ell j} = \sum_{j\in\{a,m,s\}} (1-\pi_{rrj}) P_{rj} C_{rj} \quad \forall r\,, \tag{13}$$

that is, the total value of exports must equal the total value of imports at the regional level. However, trade imbalances at the sectoral level, $Y_{rj} - C_{rj} \gtrless 0$, may exist in each region.

### 3.4 Structural change with spatial frictions

In order to study the process of spatial structural change, it will be useful to obtain production functions at the region-sector level. Given region-sector output $Y_{rj} \equiv \int y_{rj}(x) p_{rj}(x) dx / P_{rj}$, we can write the region-sector production function as $Y_{rj} = B_{rj} L_{rj}$ by defining $B_{rj} \equiv Y_{rj}/L_{rj} = \int \frac{L_{rj}(x)}{L_{rj}} \frac{p_{rj}}{P_{rj}} A_{rj}(x) dx$. Hence, in equilibrium, the aggregate sectoral productivity $B_{rj}$ in each region is obtained by averaging over the productivity of those varieties that survive country-wide competition (those with $L_{rj}(x) > 0$). Given this production function, equation (11) implies that in equilibrium

$$B_{rj} = \frac{w_r}{P_{rj}}, \tag{14}$$

which, given the expression for equilibrium prices $P_{rj}$ in equation (7), can be expressed as

$$B_{rj} = \gamma_j^{-1} \left(\frac{T_{rj}}{\pi_{rrj}}\right)^{1/\theta_j}. \tag{15}$$

This expression highlights two components of region-sector aggregate productivity. The first component is the exogenous $T_{rj}$, which determines the average productivity of region $r$ in producing goods in sector $j$. This would be the only relevant term if the region was closed to trade ($\pi_{rrj} = 1$) because region $r$ would need to produce all varieties on its own. As trade increases ($\pi_{rrj}$ declines) more goods are sourced from other regions, region $r$ can specialize in the subset of intermediate varieties for which it has a comparative advantage, and productivity increases due to selection.[10]

**Closed economy.** When our model economy is closed to internal trade and migration, each region is a closed economy independent from each other. Structural change within each region arises from the standard income and price effects due to symmetric and asymmetric sectoral productivity growth. Because $\pi_{rrj} = 1 \; \forall r, j$, this productivity growth just tracks the evolution of the exogenous productivity parameters $T_{rj}$, see equation (15). Structural change for the country is the sum of what happens in all its independent regions, while spatial income inequality is the result of asymmetric productivity growth across regions.

**Trade and structural change.** The effects of trade on structural change are well-studied in the literature on international trade and structural change cited in the Introduction. A symmetric decline in trade costs allows for regional specialization, which increases productivity due to selection ($\pi_{rrj}$ declines in equation (15)) and hence generates structural change through income effects (as long as income elasticities are different from one). When the decline in trade costs is asymmetric across sectors, relative prices change, and structural change follows as long as substitution elasticities are different from one.

**Migration and structural change.** The effects of migration on structural change are far less studied. In our model, migrations generate structural change both at the aggregate and at the local level. At the aggregate level, there is the mechanical composition effect of moving people from agrarian regions to industrial ones. This would be captured by the between-region component of a simple between-within decomposition of the evolution of country-wide sectoral shares, see Appendix G. At the local level, things are more interesting because population movements between locations change the sectoral composition of labor demand in each region, which induces local structural change. In general, this migration-led sectoral reallocation of employment does not cancel across regions and may have aggregate consequences. Hence, it appears in the within-region component of a mechanical between-within decomposition of the evolution of country-wide sectoral shares, incorrectly downplaying the role of migrations for aggregate structural change, see Appendix G.

To see how internal migrations generate local structural change, recall that the labor demand of sector $j$ in region $r$ is given by equation (12). A fraction of this demand is local (the term $\ell = r$ on the right-hand side) and a fraction is country-wide (the terms $\ell \neq r$). Suppose there is an outmigration from region $r$ (led by changes in either migration costs, trade costs, or productivity), and that $R$ is large enough such that we can abstract from the effect of outmigration on the country-wide fraction of labor demand. Then, we can state three theoretical results that we formally prove in Appendix E.1.

[10] See Finicelli et al. (2013), Sposi (2019) and our own derivations in Appendix E.1.

**Proposition 1.** *Holding prices fixed and assuming the number of regions $R$ is large enough, a decline in population $L_r$ in region $r$ leads to (a) a fall in labor demand $L_{rj}$ in all sectors $j$, and (b) a reallocation of labor away from sectors $j$ where local purchases represent a larger fraction of value added.*

Equation (12) shows that labor demands $L_{rj}$ in region $r$ decline mechanically in all sectors $j$ due to the fall in population $L_r$. The labor demand elasticity to this fall is given by

$$\frac{\partial L_{rj}}{\partial L_r}\frac{L_r}{L_{rj}} = \frac{\pi_{rrj}P_{rj}C_{rj}}{P_{rj}Y_{rj}} \leq 1 \tag{16}$$

and hence it is larger in sectors where local demand $\pi_{rrj}P_{rj}C_{rj}$ represents a larger fraction of total production $P_{rj}Y_{rj}$. In particular, local employment is reallocated away from sectors where this elasticity is larger. That is, local employment is reallocated away from sectors $j$ where region $r$ is less competitive in the sense that $\tau_{r\ell j}^{-\theta_j}T_{rj}$ is lower, which is the case for sectors $j$ in which region $r$ is less productive or less able to export.

**Proposition 2.** *Holding prices fixed and assuming the number of regions $R$ is large enough, a decline in population $L_r$ in region $r$ leads to (a) a fall in total labor demand $\sum_j L_{rj}$ that is lower than the fall in population $L_r$, and (b) an increase in the local wage $w_r$.*

The logic is clear. Without trade, all demand would be local. Then, if population leaves region $r$, the decline in labor demand is equal to the decline in labor supply and the regional wage does not change. With trade, at least a fraction of demand in each sector $j$ comes from outside region $r$, which means that outmigration reduces labor demand less than labor supply. Then, the wage has to increase to restore the equilibrium. Note the importance of the level of trade frictions for the quantitative effects of migration on the regional wage $w_r$.

In turn, the increase in the regional wage affects the sectoral employment composition in two different ways. First, it makes the local economy less competitive and lowers $\pi_{r\ell j}\ \forall \ell$ in equation (12). This effect is larger when sectoral trade costs are smaller, that is, the increase in the regional wage lowers demand relatively more in sectors that are more tradable. This is stated in the next result.

**Proposition 3.** *Holding expenditure shares constant, an increase in region $r$ wage $w_r$ reduces labor demand in all sectors $j$ of region $r$, more so in those sectors with lower trade costs, which reallocates labor towards less tradable sectors.*

Therefore, outmigration reallocates labor away from less tradable sectors due to the partial equilibrium effect but towards them due to the general equilibrium effect, with the net effect being ambiguous. However, outmigration unambiguously reallocates labor away from less productive sectors due to the partial equilibrium effect only. For instance, if manufacturing is relatively less productive in initially laggard areas, outmigration from these areas will reallocate employment away from manufacturing.

The second effect of an increase in the regional wage is that it makes region $r$ richer, which triggers changes in the composition of the local demand (the terms $P_{\ell j} C_{\ell j}$ in the right-hand side of equation (12) for $\ell = r$) due to the standard income effects.

## 4 Calibration

The calibration of the model requires choosing values for many parameters. For a given period, we have 8 preferences parameters ($\sigma$, $\nu$, $\epsilon_j$, and $\omega_j$); $R^2 + 1$ migration parameters ($\mu_{r\ell}$ and $\kappa$); $3R^2$ trade costs parameters ($\tau_{r\ell j}$); $3(R+1)$ productivity parameters ($\theta_j$ and $T_{rj}$); and the rate of population growth $n$.

We choose $R = 47$ (the 47 contiguous provinces in mainland Spain) and we want to match data at 7 different points in time (from 1940 to 2000 every ten years). The model is static conditional on the initial geographic distribution of labor, but the endogenous evolution of population plus time-changing parameters can generate rich dynamics. Hereafter, we will explicitly add a time subscript to the relevant parameters and variables of interest. We hold constant all parameters related to the elasticities of preferences, migration, and trade (that is $\sigma$, $\nu$, $\epsilon_j$, $\kappa$, and $\theta_j$). Instead, we allow the sectoral preference shifters $\omega_{jt}$, the bilateral migration costs $\mu_{r\ell t}$, the bilateral sector-specific iceberg trade costs $\tau_{r\ell jt}$, and the region- and sector-specific productivity index $T_{rjt}$ to vary over time. We also allow the rate of population growth $n_t$ to vary over time.

We start by setting $\sigma = 1$ and $\eta = 4$. The curvature of the utility function $\sigma$ does not play much of a role in a static model without uncertainty, so we use log for simplicity. The elasticity of substitution across varieties $\eta$ has no quantitative relevance, so we choose it such that it satisfies the technical condition $1 + (1-\eta)/\theta_j > 0$. We set $\theta_a = 8.8$, $\theta_m = 5.6$, and $\theta_s = 6.2$ based on the estimates of Sposi (2019). This implies that productivity dispersion, and hence the potential gains from trade, are highest in manufacturing and lowest in agriculture, with services in between.[11] Next, we normalize $\mu_{rrt} = 0$ and $\tau_{rrt} = 1$ $\forall r, t$. Finally, we choose $n_t$ to match aggregate employment growth at the country level between $t-1$ and $t$. Then, the calibration proceeds in three steps. In the first step, we estimate the preference parameters to match the aggregate evolution of sectoral employment over time. In the second step, trade costs and productivity parameters are calibrated jointly to match, every period, data on employment and productivity at the region-sector level. Finally, in the third step, we recover the migration costs parameters that perfectly rationalize the observed migration flows. In principle, one would need to know the migration elasticity $\kappa$ and costs $\mu_{r\ell t}$ to compute the labor supply for given wages in Step 2. However, observing the migration flows $\rho_{r\ell t}$ allows us to invert the labor supply equation (6), which means that we can take it as given when estimating sectoral productivity and trade costs in Step 2.[12]

[11] Alternatively, we have used $\theta_a = \theta_m = \theta_s = 4$ based on the estimate of Simonovska and Waugh (2014), and the results —available upon request— look very similar.

[12] The separation between labor supply and labor demand in the calibration stage also allows to retrieve the productivity and trade cost parameters that are consistent with labor allocations in 1940, despite not having 1930-1940 bilateral migration flows.

### 4.1 Data

We have data on (a) employment, nominal value added, and prices at region-sector level from 1940 to 2000 in 10-year periods; and (b) bilateral migration flows between all provinces from 1960 to 2000, also in 10-year periods (which we extend back to 1940, see Appendix A.2). All the data sources are detailed in Appendix A.1. We note that sectoral employment and value added shares at the province level are different from each other in the data, while they are restricted to be the same in the model. For consistency between the model and the data, we keep the sectoral split of employment and ignore the one of value added. Hence, we redefine the sectoral value added in each province as $P_{rjt}Y_{rjt} = \frac{L_{rjt}}{L_{rt}} P_{rt}Y_{rt}$, where $P_{rt}Y_{rt}$ is the provincial value added in our data. Then, using equation (11) we can infer data on regional wages as $w_{rt} = P_{rjt}Y_{rjt}/L_{rjt} = P_{rt}Y_{rt}/L_{rt}$. Additionally, using equation (14) we can infer data on region-sector productivity as $B_{rjt} = w_{rt}/P_{rjt}$.

### 4.2 First step: preferences

The preference parameters $\nu$, $\epsilon_j$, and $\omega_{jt}$ determine the sectoral composition of expenditure in each region, see equation (3). We do not have data on the sectoral composition of expenditure by province and, due to internal trade, the sectoral composition of employment or value added in each region will be different from the sectoral composition of expenditure. Hence, we cannot directly use equation (3) with province-level data to estimate the parameters governing the demand system. Instead, we can aggregate equation (3) across regions to obtain an expression relating the country-level expenditure on sector $j$ goods to regional-level wages and prices, which we do observe. Then, treating Spain as a closed economy, such that the total expenditure on sector $j$ goods equals the total value of production of sector $j$ goods, we can use data on sectoral value added at the country level to estimate the parameters of the demand system.[13] In particular, multiplying both sides of (3) by $w_r L_r$, aggregating over provinces, and applying the equilibrium condition (9) we obtain the following expression for the aggregate value added share of sector $j$ in year $t$

$$\frac{\mathrm{VA}_{jt}}{\mathrm{VA}_t} = \omega_{jt} \sum_r \left(\frac{w_{rt}}{P_{rjt}}\right)^{\nu-1} c_{rt}^{(1-\nu)\epsilon_j} \frac{\mathrm{VA}_{rt}}{\mathrm{VA}_t}\,, \tag{17}$$

where $c_{rt}$ is implicitly defined by the aggregator (1) and $\mathrm{VA}_{jt} \equiv \sum_r \left(\frac{L_{rjt}}{L_{rt}} P_{rt}Y_{rt}\right)$, $\mathrm{VA}_{rt} \equiv P_{rt}Y_{rt}$, and $\mathrm{VA}_t \equiv \sum_r P_{rt}Y_{rt}$. This expression gives us 2 independent equations per time period (so, 14 equations). As it is common in the literature, to estimate the income and price elasticities from time-series variation, we need to set constant the preference shifters $\omega_{jt}$. Hence, we have 5 parameters to estimate: $\omega_a$, $\omega_s$, $\epsilon_a$, $\epsilon_s$, and $\nu$ (without loss of generality we normalize $\omega_m = 1$ and $\epsilon_m = 1$). We do so with data on $w_{rt}$, $P_{rjt}$, $L_{rjt}/L_{rt}$, and $P_{rt}Y_{rt}$ and by non-linear least squares,

[13] Exploiting the closed-economy assumption for estimation of preferences is standard in this liteature, e.g. Herrendorf et al. (2013). The equality of sectoral value added and expenditure shares further requires that value added shares in investment are similar to those in consumption. This is also a standard assumption in the literature, but see García-Santana et al. (2021) and Herrendorf et al. (2021) for recent examples where they differ.

solving for $c_{rt}$ in each iteration combining equations (1) and (3). Next, because we need the predicted aggregate sectoral expenditures to be consistent with aggregate sectoral labor incomes (such that we can estimate regional productivity and trade costs in Step 2 below), we recover the time-varying $\omega_{jt}$ that perfectly match the aggregate sectoral value added shares at every decade given the estimated elasticities $\nu$ and $\epsilon_j$.[14]

**Estimated parameters and model fit.** The estimated demand system with constant $\omega_j$ reproduces well the aggregate sectoral shifts in the value added of the Spanish economy between 1940 and 2000, see Panel (a) in Figure H.2. The parameter estimates are reported in Table 1. We estimate an elasticity of substitution $\nu$ close to zero, which implies that goods from different sectors are poor substitutes. Hence, changes in relative prices have very small effects on relative quantities and translate almost one-to-one into changes in relative sectoral expenditure. The estimated value of the term $\epsilon_j$, which controls the income elasticity of sectoral demand, is higher in services than in agriculture, and in both cases lower than in manufacturing. For the region with the average real income, these estimates generate income elasticities of 0.56, 1.47, and 1.29 for agriculture, manufacturing, and services, respectively, in 1940.[15] In the year 2000, these elasticities are 0.42, 1.10, and 0.97. This implies that income growth reallocates expenditure away from agriculture and towards manufacturing and, to a lesser extent, services. This force is more than offset by the decline in the price of manufactures relative to services, which shifts non-agriculture expenditure from manufacturing towards services.[16] While the estimated demand system provides a good fit to the evolution of value added shares over time, a perfect match requires time-varying $\omega_{jt}$. The time-varying $\omega_{jt}$ that we recover from this correction procedure can be found in Table H.1. To grasp the importance of changes in $\omega_{jt}$ over time, panel (b) in Figure H.2 plots the evolution of value added shares when only $\omega_{jt}$ changes, keeping wages and prices constant at their 1940 values. We can see that $\omega_{jt}$ slightly shifts expenditure away from agriculture and into services, particularly for the last decade of our data.[17]

[14]Our calibration strategy in Step 2 takes the predicted local sectoral expenditure as given and looks for the productivity and trade costs that match the sectoral labor income observed in the data. To be able to match the sectoral labor income, the predicted aggregate sectoral expenditure has to perfectly match the observed aggregate sectoral labor income. Otherwise, we cannot recover productivity and trade costs that generate local sectoral labor incomes that, once aggregated, are consistent with the data.

[15]The income elasticity of sectoral demand is given by $\partial \log c_{rjt} / \partial \log w_{rt} = \nu + (1-\nu)(\epsilon_j/\bar{\epsilon})$ where $\bar{\epsilon} = \sum_j \frac{P_{rjt} c_{rjt}}{w_{rt}}$ is the expenditure-weighted average of the non-homotheticity parameters.

[16]Preference estimates with sectoral value added data tend to find $\nu \simeq 0$ as we do here, see Herrendorf et al. (2013). Different from standard results, we find $\epsilon_m > \epsilon_s$ whereby income growth reallocates employment from manufacturing to services. We find similar results with Stone-Geary preferences.

[17]The increase in $\omega_{st}/\omega_{at}$ in the 1990s may reflect an estimated income effect that is weaker in this decade than the observed reallocation from agriculture to services. One potential explanation is that the current account deficit increased sharply in the 1990's, which means that consumption expenditure increased more than GDP. Our closed economy model is fed with the increase in GDP, not consumption expenditure, so it misses part of the true income effect.

Table 1: Demand system parameters

| $\omega_a$ | $\omega_m$ | $\omega_s$ | $\nu$ | $\epsilon_a$ | $\epsilon_m$ | $\epsilon_s$ | RMSE a | RMSE m | RMSE s |
|---|---|---|---|---|---|---|---|---|---|
| 0.024 | 1.0 | 0.382 | 0.005 | 0.379 | 1.0 | 0.88 | 0.053 | 0.081 | 0.107 |

**Notes:** the table reports the estimated parameters of the demand system restricted to have constant $\omega_j$. The parameters are estimated by non-linear least squares using equation (17) and data on sectoral value added. The table also reports the root mean squared error (RMSE) for each sector.

### 4.3 Second step: trade costs and productivity

**Trade costs.** It is standard in models of international trade to calibrate the iceberg trade costs $\tau_{r\ell jt}$ using equation (8) and data on bilateral trade flows across countries. Data on trade flows within countries, however, are not typically available for development episodes. Hence, the use of models with costly internal trade for macro-development questions requires a different empirical strategy. We develop a novel strategy to calibrate $\tau_{r\ell jt}$ based on the mismatch between sectoral expenditure and production shares in each region. In particular, the difference between sectoral employment and sectoral expenditure shares at the local level, other things equal, provides information on the magnitude of trade costs, see Gervais and Jensen (2019). Intuitively, if the economy is closed to trade, all consumption is local, and the sectoral shares in expenditure and in production are equal to each other in all provinces. With low trade costs, instead, provinces consume according to their local prices and wages and produce according to their comparative advantage.

Using equation (8) to substitute away trade flows, the sectoral labor demands (equation 12) can be written as

$$L_{rjt} = \frac{1}{w_{rt}} \sum_{\ell}^{R} \left( \frac{(w_{rt}\tau_{r\ell jt})^{-\theta_j} T_{rjt}}{\sum_k^R (w_{kt}\tau_{k\ell jt})^{-\theta_j} T_{kjt}} \right) P_{\ell jt} C_{\ell jt} \quad \forall r, \forall j. \tag{18}$$

We have data on sectoral employment $L_{rjt}$ and on wages $w_{rt}$ for each region. Sectoral expenditure in each region $P_{\ell jt}C_{\ell jt}$ is given by equation (3) using the preference parameters estimated in the first step together with price and wage data. Hence, equation (18) gives us $3R$ moment conditions, while we have $3R(R-1)$ bilateral trade costs $\tau_{r\ell jt}$ and $3R$ productivity parameters $T_{rjt}$ to recover. We will deal with the $T_{rjt}$ using $3R$ extra moments in the next paragraph, but we still have more trade costs parameters than equations. This happens because the full matrix of bilateral trade costs is not necessary to determine labor demands at the region-sector level. Therefore, to obtain identification, we need to reduce the dimensionality of the matrix of trade costs. First, we parameterize the iceberg costs as $\log \tau_{r\ell jt} = (\hat{\tau}_{rjt} d_{r\ell})$, where $\hat{\tau}_{rjt}$ captures the ability of region $r$ at exporting sector $j$ goods at time $t$, and $d_{r\ell}$ is a time-invariant term capturing origin-destination fixed effects. This means that we allow the trade costs (a) to be asymmetric between origin-destination routes, $\tau_{r\ell jt} \neq \tau_{\ell rjt}$, and (b) to vary over time by sector-origin. Second, we use data on road distance (normalized between 0 and 1) between $r$ and $\ell$ capital cities in the year 2000 to calibrate the $d_{r\ell}$ terms outside the model. All in all, for every period $t$, equation (18) gives us $3R$ conditions and we have $3R$ trade cost parameters $\hat{\tau}_{rjt}$ to pin down. In terms of identification, on the one hand, the correlation of sectoral expenditure and employment shares across provinces for each sector $j$ reveals information about the average values of $\hat{\tau}_{rjt}$ in sector $j$. In the data, there is a steady decline in this correlation for all three sectors after 1950, which will help recover declining sectoral trade costs (see Table H.2 in Appendix H). On the other hand, the correlation of sectoral expenditure and employment shares across sectors for each region $r$ reveals information on the average values of $\hat{\tau}_{rjt}$ in region $r$. That is, other things equal, provinces with similar expenditure and production shares in all sectors are inferred to face higher trade costs.

**Productivity.** Absent data on internal trade flows, our data on average sectoral productivity by region $B_{rjt}$ is not enough to directly recover $T_{rjt}$ from equation (15). Hence, we need to plug equation (8) into (15) to obtain

$$B_{rjt} = \gamma_j^{-1} \left( \sum_{k}^{R} \left( \frac{w_{rt}}{w_{kt} \tau_{krjt}} \right)^{\theta_j} T_{kjt} \right)^{1/\theta_j} \qquad \forall r, \forall j \tag{19}$$

which gives us $3R$ non-linear equations in as many unknowns $T_{rjt}$ for every time period $t$ given the trade costs $\tau_{r\ell jt}$ and the wage and average productivity data $w_{rt}$ and $B_{rjt}$.

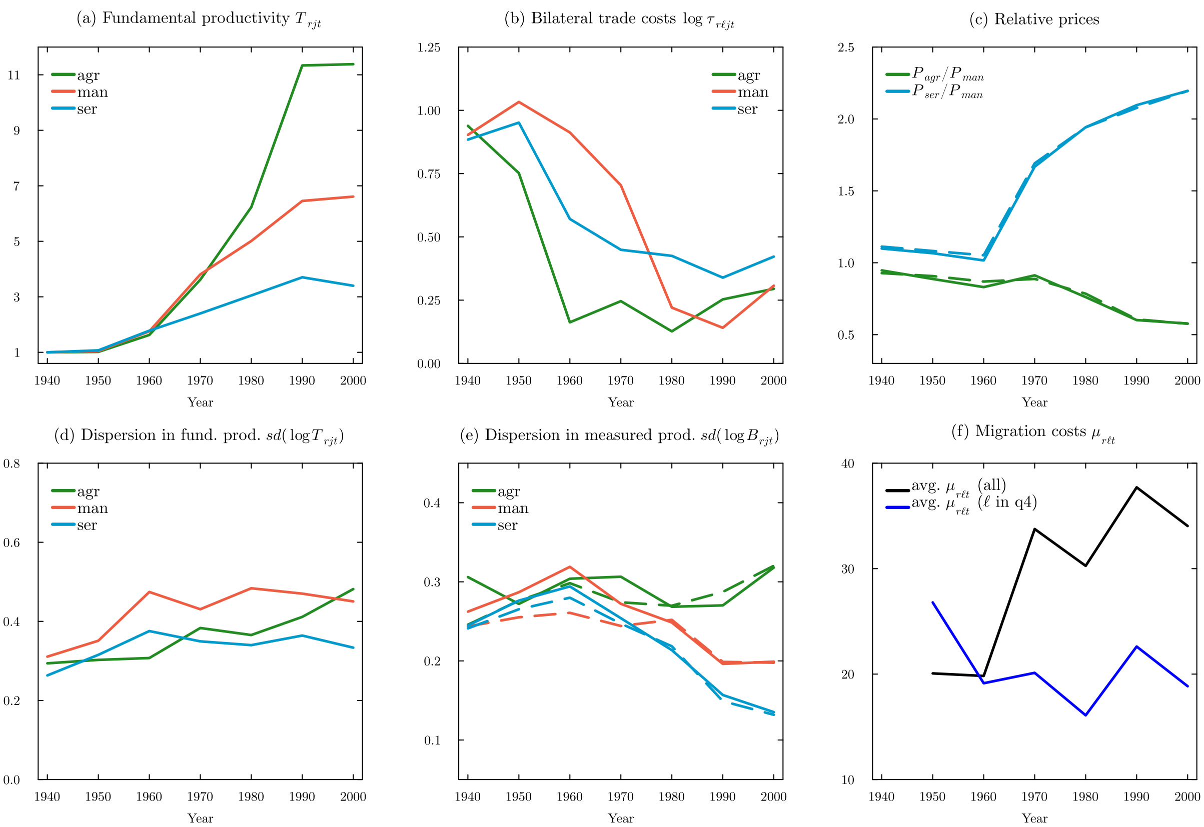
**Estimation and model fit.** We aim to find the values of $\hat{\tau}_{rjt}$ and $T_{rjt}$ that solve the non-linear system of equations given by (18) and (19) for each region in each sector $j$, *separately for every period*. In practice, we minimize the sum of square residuals of both equations, rather than trying to find the zero of the large non-linear system, as we find this strategy to work better. The model fit to the targeted data moments on employment and average productivity in each region and sector is presented in Figures H.3 to H.6. Overall, the model fits very well the targeted moments.

**Estimated parameters.** In order to summarize all the information of the estimated parameters, we report the average (across all provinces $r$) of $T_{rjt}$ for every sector $j$ and year $t$ in Panel (a) of Figure 5, and the average (across all origin-destination pairs $r\ell$) of all the values of $\tau_{r\ell jt}$ below the 75th percentile of their distribution for every sector $j$ and year $t$ in Panel (b) of Figure 5.[18] The estimated parameters display two important trends: an increase in sectoral productivity (mostly between 1950 and 1990) and a decline in trade costs (mostly between 1950 and 1980).[19] The productivity growth is largest in agriculture and smallest in services, while the decline in trade costs

[18]We report the average of a left truncation of the distribution of trade costs because these are the values that are more relevant for trade, as very high trade costs are typically associated with routes with no trade.

[19]The productivity stagnation in the 1940s is consistent with the lack of economic development in that decade, see Figure 1 in Section 2. The productivity slowdown in the 1990s (and beyond) is already documented by García-Santana et al. (2020) among others. The fall in trade costs can be understood as a result of large investments in transport equipment and infrastructure. For instance, a publicly-funded program to improve the surface and increase the width of the most-used 5,000 km of the road network was in place between 1967 and 1974, while the construction of the first highways (6,000 km) connecting the main cities of the country kick started in 1968, see Ventosa (2017).

Figure 5: Estimated trends

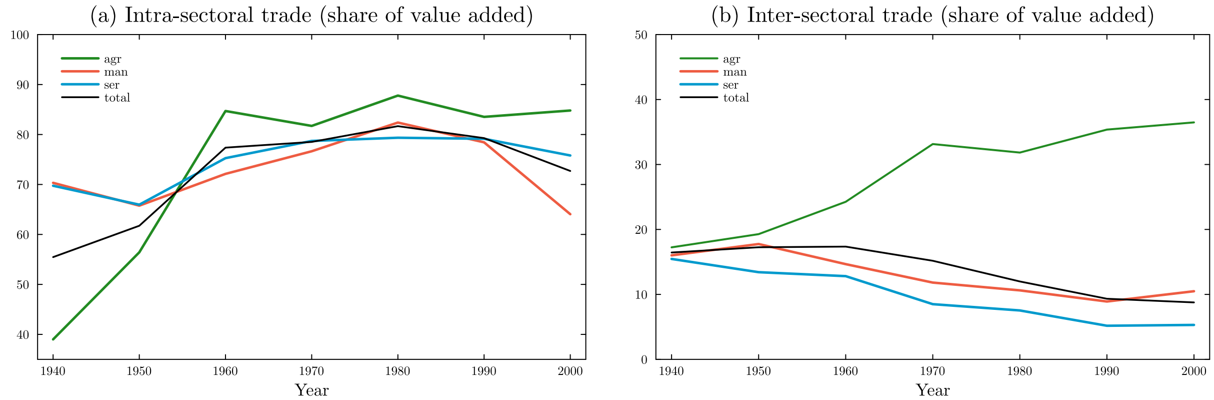



**Notes:** Panel (a) displays the estimated sectoral productivity parameters averaged across all provinces, normalizing 1940 to 1 in each sector. Panel (b) reports the average (across all origin-destination pairs) of the estimated trade costs parameters below the 75th percentile of their distribution. Panel (c) shows the relative price of agriculture and services to manufacturing averaged across provinces in the data (solid line) and in the model (dashed line). Panels (d) and (e) report the standard deviation across regions of the log of fundamental productivity $T_{rjt}$ and the log of labor productivity $B_{rjt}$, with dashed lines representing the data in the latter. Panel (f) reports the estimated migration cost parameters averaged across all origin-destinations—black line—and averaged across destinations in the top quartile of relative employment growth between 1940 and 2000.

is larger and earlier in time for agriculture than for manufacturing and services. Both changes in productivity and trade costs affect relative prices across sectors, as reported in Panel (c) of Figure 5. The asymmetric productivity growth is the main driver of the increase in the price of services relative to manufactures, while the decline in the price of agriculture relative to manufactures is driven first by the asymmetric fall in trade costs and then by the asymmetric growth in productivity, see Figure F.1. These patterns summarized by averages are common to most provinces, see Figure H.7 in Appendix H.

**Implied trade volumes.** The fall in trade costs and changes in the dispersion of productivity generate a rise in trade volumes. We define intrasectoral trade as the fraction of sectoral expenditure not produced locally, $\sum_r (1 - \pi_{rrjt}) P_{rjt} C_{rjt}$, see Appendix E.1. This is the notion of trade in Eaton and Kortum (2002) whereby intrasectoral heterogeneity in productivity across provinces generates Ricardian trade. As we can see in panel (a) of Figure 6, intrasectoral trade as a share of sectoral

Figure 6: Trade across provinces




**Notes:** Panel (a) reports predicted intrasectoral trade as a share of sectoral value added. For the year 2000, the dots represent intrasectoral trade in manufacturing (red) and agriculture (green) in the data. Panel (b) shows predicted intersectoral trade as a share of sectoral value added.

value added increases in all three sectors. This increase is much more apparent for agriculture, while for manufacturing and services it follows a hump-shape, given that aggregate value added reallocates over time toward these sectors. We can additionally define intersectoral trade for sector $j$ as the difference between sectoral expenditure and sectoral value added, $\sum_r \frac{1}{2} |P_{rjt}C_{rjt} - P_{rjt}Y_{rjt}|$, see Appendix E.1. In Panel (b) of Figure 6 we see how intersectoral trade declines relative to sectoral value added in manufacturing and services, which reflects growing convergence between sectoral expenditure and value added shares across provinces in these sectors. Instead, intersectoral trade increases in agriculture due to the fact that the declining production of agriculture gets concentrated over time in fewer, specialized regions.

### 4.4 Third step: migration costs

In order to recover the migration elasticity $\kappa$ and the bilateral migration costs $\mu_{r\ell t}$, we use our data on the bilateral migration flows $\rho_{r\ell t}$ for each decade. Many papers in the migration literature parameterize $\mu_{r\ell t}$ as a function of the distance between regions. We prefer to keep these costs non-parametric for two reasons. First, while trade costs matter, there may be other factors influencing connectivity between regions. For instance, existing networks of previous migrants from the same hometown or the availability of cheap housing at destination are typically important. Likewise, the availability of public subsidies in poorer regions or the strength of family networks at origin may also be relevant.[20] And of course, all these aspects may vary over time. Second, by keeping migration costs non-parametric, we can perfectly match the data on migration flows, which allows us to separate this part of the calibration from the algorithm in Step 2.

As it is common in this literature —see for instance Artuç et al. (2010)— we start by estimating

[20]This argument is reinforced by the fact that the correlation between bilateral migration flows and distance is quite weak in the Spanish data, between -0.28 and -0.30 depending on the year.

$\kappa$ from the observed correlation between migration flows and differences in regional value functions. In particular, using equation (5) we can write

$$\log \rho_{r\ell t} - \log \rho_{rrt} = \frac{1}{\kappa}\big(\mathcal{V}(w_{\ell t}, P_{\ell at}, P_{\ell mt}, P_{\ell st}) - \mathcal{V}(w_{rt}, P_{rat}, P_{rmt}, P_{rst})\big) - \frac{\mu_{r\ell t}}{\kappa} \tag{20}$$

where $\mu_{rrt}$ is normalized to 0 and the non-linear part of equation (5) (the denominator) is differenced away. This expression shows that $\kappa$ regulates how many people move from region $r$ to region $\ell$ (as compared to those that stay in $r$) given the difference in value between them. With data on wages and prices, and with the parameters of the utility function estimated before, we can construct the value of living in each location. Then, we estimate $\frac{1}{\kappa}$ using a Poisson Pseudo Maximum Likelihood (PPML) estimator, instrumenting $\mathcal{V}(w_{rt}, P_{rat}, P_{rmt}, P_{rst})$ with the fundamental productivity parameters recovered in Step 2.[21]

Our estimation delivers $1/\kappa = 0.434$ (s.e. 0.070). As we are using log utility, this number corresponds to the elasticity of migration flows to real income for the homothetic limit where $\epsilon_j$ are all equal to one.[22] The time evolution of the bilateral migration costs that we recover is reported in Panel (f) of Figure 5. We report the average for every period $t$ of the $\mu_{r\ell t}$ over all routes $r\ell$ (black line). The migration costs increase all over the period. However, many of these terms are economically irrelevant because the associated bilateral routes are insignificant in terms of population movements. For this reason, we also report the average of the $\mu_{r\ell t}$ for all the routes $r\ell$ whose destination $\ell$ is one of the provinces in the top quartile of the distribution of relative employment growth over the 1940-2000 period (blue line). We can see that the migration costs towards these high-growth provinces fell sharply in 1960 (corresponding to the migration flows between 1950 and 1960), remain stable in 1970 and fall again in 1980. Then, they increase back between 1980 and 1990. This means that internal migrations towards the most dynamic areas between 1950 and 1980 —the time of the rural exodus— were partly fuelled by a decline in migration costs. One possible explanation for this decline is the government-led construction of cheap housing for migrant workers in the (then) outskirts of cities like Barcelona, Bilbao, or Madrid.[23] Another one is the accumulation of migrant networks from the same location. Later on, the increase in migration costs from the 1980s reveals that, despite differences in real wages

[21] Estimating $\kappa$ by OLS is problematic. First, estimates of $1/\kappa$ may be biased because differences in $\mathcal{V}(w_{rt}, P_{rat}, P_{rmt}, P_{rst})$ across regions are an equilibrium outcome that depends on the value of unobserved migration costs $\mu_{r\ell t}$. Second, the outcome variable $\rho_{r\ell t}/\rho_{rrt}$ may take values very close to zero, as the flow between some region pairs is negligible, leading to heteroskedastic error terms and biased estimated of $1/\kappa$, see Silva and Tenreyro (2006).

[22] This elasticity is a bit lower than elasticities typically found in the literature. For instance, using a similar model to ours, Tombe and Zhu (2019) estimate an elasticity of 1.5 for China by use of 5-year migration flows, while Morten and Oliveira (2018) estimate an elasticity of 4.5 for Brazil by use of 10-year migration flows.

[23] In the early 1950s, migrants to the big cities settled in self-built shanty towns. In 1957, the newly created *Ministerio de la Vivienda* approved the *Plan de Urgencia Social*, with the explicit objective of building cheap (legal) housing for the migrants arriving in big cities. Construction was done by the private sector, fuelled by public subsidies and cheap land provided by the government through selected rezoning, see López Simón (2022) for details. Between 1950 and 1980, the number of residential dwellings in the provinces of Barcelona, Madrid, and Vizcaya multiplied by a factor of 3.5 or more, as compared to a factor of 2 for the rest of the country. See Table 6.7 in Carreras and Tafunell (2005).

across provinces persisted, workers were not moving. We interpret this post-1980 increase in the migration costs towards the most dynamic areas as the result of the development of the welfare state in Spain, which equalized after tax-transfer incomes.[24] In our model, this would show up as in increase in the migration costs from poor areas to richer ones.

### 4.5 The calibrated economy *vs* the data

We report a selection of development statistics for the benchmark economy in Column (2) of Table 2. By construction of our estimation strategy, they are very close to the actual data, see Column (1). Real income grows by a factor of 8.31. The employment shares of agriculture and services decline and increase in 46 and 39 percentage points respectively, while the employment share of manufacturing follows a hump shape, increasing 14 percentage points until 1970 and declining by 7 percentage points afterwards. The uneven industrialization across provinces can be seen by comparing the increase in manufacturing employment between 1940 and 1970 for q1 provinces (those that lose the most population) and q4 provinces (those that gain the most population). While for the q1 group this increase is only 8% of its total employment in 1940, it amounts to 43% for the q4 group. The rural exodus is summarized by three statistics: the increase in the coefficient of variation employment across provinces of 63 percentage points, the regression coefficient of log regional employment growth between 1940 and 2000 on the initial agricultural share of -2.05 (which corresponds to the slope of Panel (a) in Figure 2), and the $R^2$ of this regression of 0.60. The evolution of spatial inequality is described by the changes in the standard deviation of regional income, which falls by 24 p.p. after 1960.

### 4.6 Engines of development

Given the calibrated model, one can analyze the sources of growth in Spain by decomposing how changes in productivity $T_{rjt}$, migration costs $\mu_{r\ell t}$, and trade costs $\tau_{jr\ell t}$ affect different long run outcomes. This analysis is not the main goal of our paper, so we relegate it to Appendix F.1. As a summary, we mention that the growth in productivity generates most of the changes in relative prices and aggregate productivity that drive structural change and growth. Changes in trade costs complement productivity growth to complete the reallocation of employment out of agriculture. The effects of changes in migration costs are analyzed in the next Section.

[24] Bentolila (1997) documents that social protection expenditures rose from 18% in 1980 to 24% of GDP in 1993. In addition, a special transfer system was implemented to protect unemployed workers in agriculture in the provinces of Andalucía and Extremadura paying 75% of the minimum wage for up to 300 days to individuals having worked for at least 40 days within the year.

# 5 Quantitative results

We start this Section by examining how the rural exodus shaped the development process of Spain (Section 5.1). To do so, we consider counterfactual economies where no migrations are possible. At the regional level, we find that internal migrations —and not the evolution of comparative advantage— is the main reason behind the asymmetric industrialization of Spain. At the aggregate level, we find that internal migrations accelerated growth by reallocating labor toward more efficient uses, yet the overall effect was moderate.

Next, we study the driving forces of the rural exodus in Spain, and relate it to the cross-country evidence presented in Figure D.1 (Section 5.2). We find that the decline of migration costs towards the leading regions (Section 5.2.1) and the regional divergence in productivity (in manufacturing and services) during the first half of the development process (Section 5.2.2) were on their own powerful sources of migrations and of uneven paths to regional industrialization, respectively.

## 5.1 The role of the rural exodus

To explore the role of the rural exodus in the development experience of Spain, we solve for a counterfactual economy in which workers cannot migrate (migration costs tend to infinity) and hence the relative size of provinces remains as in 1940. The main challenge in this exercise is what to assume about the evolution of labor productivity when the economy faces substantially different migration flows. In our first exercise, we assume that the region-specific sectoral productivity paths $T_{rjt}$ evolve as in the calibrated economy and hence do not react to population movements. This is a strong assumption, but one that allows us to focus on the partial effect of migration flows on development. Second, we allow for agglomeration economies in industry and decreasing returns in agriculture, which results in industrial productivity growing more in areas receiving more population. Our results below show that the main insights for the effects of migration on local development are already apparent in the first case, although there are some interesting quantitative differences in the second one.[25]

### 5.1.1 Exogenous productivity paths

When productivity paths $T_{rjt}$ evolve as in the calibrated economy (they do not depend on local employment), we find that without population movements the initially lagging regions do industrialize,

[25] Migrations may affect labor productivity for other reasons too. First, the allocation of capital across regions and time may depend on the allocation of labor, see for instance Kleinman et al. (2021). In this situation, capital stocks may follow population changes, increasing the labor productivity in regions attracting more migrants. Second, there may be selection of individuals of better skills into migration as in Lagakos and Waugh (2013) such that regions attracting migrants receive more skilled workers (although for the case of Spain we show very little selection on education, see Appendix C). Considering any of these extensions comes at the cost of a more complicated model. The results in this Section show that most of the interesting action already happens with productivity paths that are invariant to population movements.

Table 2: Counterfactual exercises (changes between 1940 and 2000)

| | | | No rural exodus | | Early convergence | | |
|---|---|---|---|---|---|---|---|
| | Data (1) | Bench (2) | CRS (3) | Non-CRS (4) | Man (5) | Man, Ser (6) | Ini $\mu_{r\ell t}$ (7) |
| Real income pc | 8.16 | 8.31 | 7.95 | 7.92 | 8.31 | 8.24 | 7.92 |
| Agr share | −0.45 | −0.46 | −0.46 | −0.47 | −0.46 | −0.46 | −0.46 |
| Man share: 40-70 | 0.14 | 0.14 | 0.14 | 0.15 | 0.18 | 0.16 | 0.14 |
| Man share: 70-00 | −0.08 | −0.07 | −0.07 | −0.07 | −0.11 | −0.08 | −0.07 |
| Ser share | 0.39 | 0.39 | 0.39 | 0.39 | 0.38 | 0.39 | 0.39 |
| Man emp q1: 40-70 | 0.07 | 0.08 | 0.19 | 0.26 | 0.38 | 0.28 | 0.16 |
| Man emp q4: 40-70 | 0.44 | 0.43 | 0.27 | 0.25 | 0.26 | 0.33 | 0.31 |
| $CV(L_r)$ | 0.65 | 0.63 | 0.0 | 0.0 | 0.59 | 0.56 | 0.2 |
| $\hat{\beta}$ | −2.07 | −2.05 | 0.0 | 0.0 | −1.92 | −1.88 | −0.08 |
| $R^2$ | 0.63 | 0.6 | 0.0 | 0.0 | 0.58 | 0.57 | 0.0 |
| Sd($\log C_r$): 40-60 | 0.02 | −0.02 | −0.01 | −0.01 | −0.1 | −0.13 | −0.01 |
| Sd($\log C_r$): 60-00 | −0.21 | −0.24 | −0.23 | −0.23 | −0.19 | −0.16 | −0.22 |

**Notes:** Each column reports level changes in the corresponding variable between 1940 and 2000, except other time frame indicated in the corresponding row. For real income pc we report the ratio. "$\hat{\beta}$" and "$R^2$" are the coefficient and the share of explained variance of a regression of employment growth on initial agriculture share. "Man emp q1: 40-70" is employment change in manufacturing in q1 provinces relative to total employment of those provinces in 1940. Columns (1) and (2) correspond to the data and the calibrated economy respectively. Columns (3) and (4) correspond to economies with no migration, with constant and non-constant returns to scale production functions respectively. Columns (5) and (6) show the outcomes of economies in which productivity converges across regions since 1940 in manufacturing and in manufacturing and services respectively. Column (7) shows the outcomes of an economy with migration costs fixed at their 1940 values.

while the initially leading regions specialize much less in manufacturing. In particular, absent migrations, provinces in the q1 and q4 groups would have increased their manufacturing employment between 1940 and 1970 in 19% and 27% of their initial population size, while in the benchmark economy this increase is 8% and 43% respectively, see Column (3) in Table 2 (or Figures H.8 and H.9 in Appendix H for a few province by province examples). Therefore, we can conclude that the evolution of productivities and trade costs in the Spanish development episode was conducive to industrialization in the initially rural areas, but migrations prevented this from happening. The rural exodus is behind the uneven industrialization of Spanish regions. We discussed the economic mechanisms in Section 3.4. First, manufactures in initially laggard areas were not very competitive and hence depended very much on local demand, which the rural exodus depressed. Second, migration flows eroded the wage gap between leading and laggard regions, which made the manufactures of leading regions much more competitive country-wide, eventually taking over most of the market.

In the aggregate, the rural exodus contributed to real income growth, while having very limited effects on structural change. In particular, real income increases 36 percentage points more in the benchmark economy than in the no migration counterfactual, as workers are forced to remain in regions that are less productive in manufacturing and services. Yet, the aggregate reallocation of employment across sectors remains unchanged. On the one hand, keeping workers in laggard regions makes the country poorer, which limits the reallocation of employment into manufacturing due to income effects. On the other hand, it hinders aggregate productivity growth in manufacturing, which limits the reallocation of employment out of manufacturing due to price effects. In the aggregate, these two opposing forces tend to compensate each other, and the overall effect of the rural exodus on structural change is limited.

### 5.1.2 Non-constant returns to scale

Next, we allow for non-constant returns to scale in sectoral production functions. Our quantitative model can easily accommodate this exercise. In particular, we can redefine the scale parameter $T_{rjt}$ as $T_{rjt} \equiv \widetilde{T}_{rjt} L_{rj}^{\theta_j \alpha_j}$, where $\widetilde{T}_{rjt}$ reflects the exogenous component of productivity and $L_{rjt}^{\theta_j \alpha_j}$ reflects departures from constant returns to scale whenever $\alpha_j \neq 0$. Then, the average sectoral output $Y_{rjt}$ can be written as $Y_{rjt} = B_{rjt} L_{rjt}^{1+\alpha_j}$. With an estimate of $\alpha_j$ for each sector, the estimated productivity paths $T_{rjt}$ from Section 4.3, and the observed allocation of workers $L_{rjt}$, we can recover the time paths $\widetilde{T}_{rjt}$ that are invariant to population changes. This allows to separate productivity $T_{rjt}$ between a "fundamental" exogenous component $\widetilde{T}_{rjt}$ and a component that varies with $L_{rjt}$ in counterfactual exercises.

Estimating values for $\alpha_j$ is notoriously difficult as good instruments are needed to isolate exogenous variation in population. Lacking such instruments, we take some values from outside sources. In the macro-development literature, it is quite standard to set $\alpha_a < 0$, reflecting decreasing returns to scale in agriculture due to the fixed land factor. Typical values for $\alpha_a$ range from 0.30 (see Gollin et al. (2007) or Restuccia et al. (2008), who use estimates from Hayami and Ruttan (1985)) to 0.09 (Lagakos et al., 2023), so we decide to set an intermediate value of $\alpha_a = -0.15$. For manufacturing, we resort to the estimate of Desmet et al. (2018), a classic reference in the literature, and set $\alpha_m = 0.06$. Finally, for lack of a better alternative, we leave $\alpha_s = 0$ in services. In Appendix F.2 we discuss the difference between the calibrated productivity paths $T_{rjt}$ and the ones inferred with non-constant returns to scale production functions, $\widetilde{T}_{rjt}$.

Then, we revisit our counterfactual exercise of no migration with non-constant returns to scale production functions, see Column (4) in Table 2. Overall, we find that the qualitative effects are the same and that the quantitative results get reinforced. The initially lagging (leading) regions experience a stronger (weaker) industrialization when there is no migration compared to the CRS case because the increase (decrease) in manufacturing employment in these provinces improves (worsens) their manufacturing comparative advantage. Indeed, with agglomeration economies and no migration, q1 provinces experience a slightly stronger industrialization process than q4 provinces, as the manufacturing employment increase between 1940 and 1970 equals 26% and 25% of initial employment respectively. This can be seen by comparing Figure H.10 (H.11) in Appendix H for the case with agglomeration with Figure H.8 (H.9) for the case with CRS. In terms of growth, the rural exodus is a slightly larger contributor to the increase in aggregate income when we allow for non-constant returns to scale, as migrations toward leading regions raise aggregate productivity beyond the mechanical increase due to shifting population from less to more productive regions.

### 5.2 The origins of the rural exodus

In order to uncover the economic forces leading to the rural exodus in Spain, we draw on the experiments in Appendix F.1 and focus on the heterogeneous evolution of migration costs and sectoral productivity across regions.

#### 5.2.1 Changes in migration costs

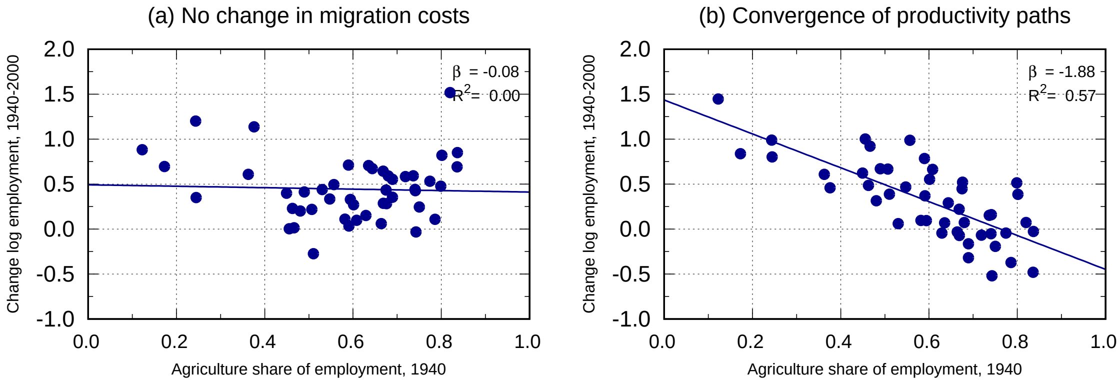
One important feature of our calibrated model is the decline of migration costs $\mu_{r\ell t}$ towards the most dynamic regions in 1960, 1970, and 1980 (see Section 4.4). To explore the importance of this decline, in this Section we solve for a counterfactual economy in which migrations costs $\mu_{r\ell t}$ remain constant at their values in 1940, while the other time-changing parameters (productivity $T_{rjt}$ and trade costs $\tau_{jr\ell t}$) evolve as in the benchmark economy. We report the main results in Column (7) of Table 2.

First, we find that keeping migration costs constant has important effects on the pattern of migrations, see Column (7) in Table 2. In particular, there is a much smaller increase in the dispersion of population across regions, with the coefficient of variation of population across provinces increasing by only 20 percentage points (compared to 63 in the benchmark). More importantly, the relationship between the initial agricultural employment share and employment growth vanishes, as the regression coefficient $\hat{\beta}$ becomes -0.08 (-2.26 in the benchmark), with an $R^2$ of 0.00 (0.60 in the benchmark). We illustrate this in Panel (a) of Figure 7. The results of this counterfactual exercise for Spain are consistent with the findings of Hao et al. (2020) showing that the reform of the *hukou* system in China —an effective decline of migration costs— was an important driver of the Chinese rural exodus between 2000 and 2015. As discussed above (see footnote 23), the decline in migration costs towards the most prosperous regions can be related to the cheap housing policy developed by the Spanish government in the outskirts of big cities.

Second, with constant migration costs the industrialization paths are less uneven across regions. In particular, the increase in industrial employment represents 16 and 31 percent of initial employment in the q1 and q4 provinces respectively, compared to 8 and 43 percent in the calibrated economy. That is, the economy with constant migration costs finds it hard to allocate labor in the industrial sectors of the provinces leading the development process.

Finally, we find that the decrease in migration costs has limited aggregate effects, consistent with the results of the no migration counterfactual in the previous Section. In particular, the economy with constant migration costs grows 39 percentage points less than the benchmark economy, as workers do not move to the most productive regions.

Figure 7: Employment growth and initial sectoral composition: counterfactual economies




**Notes:** This figure plots the relative increase in employment between 1940 and 2000 (in logs) for all provinces, against the 1940 agriculture share of employment for different model economies. Panel (a): an economy with migration costs constant to their 1940 values; Panel (b): an economy where sectoral productivities start converging across regions since 1940. Each panel also reports the slope of the relationship ($\beta$) and the share of variance in log employment growth explained by the corresponding x-axis variables ($R^2$).

### 5.2.2 Early convergence in productivity

Labor productivity in manufacturing and services diverged across provinces during the first two decades of the Spanish development episode and only started to converge after 1960 (see Panel (b) in Figure 4). This pattern increased the gap between leading and lagging regions at the start of the developing process, generating the inverted-U-shaped profile of inequality across provinces documented in Panel (a) in Figure 4. Our calibrated economy recovers productivity paths $T_{rjt}$ generating a similar pattern of regional inequality, see Panel (d) in Figure 5. In this Section, we want to quantify the role played by the initial divergence of sectoral productivity across provinces in the development process of Spain, in particular on its rural exodus and uneven regional industrialization. To do so, we generate counterfactual sector-specific productivity paths $T'_{rjt}$ for manufacturing and services in each Spanish province such that sectoral productivity starts converging in 1940 —at the same rate as services productivity converged between 1960 and 2000— while aggregate sectoral productivity still grows as in the data, see Appendix F.3 for details. We leave the evolution of productivity in the agriculture sector as in the calibrated economy.

We start by considering convergence only in manufacturing productivity —while letting services and agriculture productivity evolve as in the calibrated economy— and report the main results in Column (5) of Table 2. The convergence in manufacturing productivity is able to generate convergence in real income per capita across provinces before 1960 (the standard deviation of income per capita decreases 10 log points, compared to 2 log points in the benchmark). However, it barely reduces the rural exodus (the slope of the relationship between initial agriculture share and employment growth is -1.92 and the $R^2$ 0.58, compared to -2.07 and 0.60 in the benchmark). Convergence in manufacturing productivity reverses the uneven pattern of industrialization across regions, as the relative increase in manufacturing employment in q1 provinces is higher than in q4 provinces. At the aggregate level, manufacturing experiences a bigger hump over the course of development than in the benchmark, as both the increase pre-1970 and the decline post-1970 of the manufacturing share are higher than in the benchmark. Before 1970, higher manufacturing productivity in q1 regions pushes employment towards manufacturing due to income effects; after 1970, it reallocates employment away from manufacturing due to price effects.

When we extend the convergence in productivity to both manufacturing and services, these results are amplified. First, convergence in real income per capita is stronger before 1960 (its standard deviation declines by 13 log points). This further diminishes the incentives to outmigrate from the initially rural areas, producing a smaller increase in the dispersion of population across provinces (a 56 percentage points increase in the coefficient of variation) and a smaller correlation between initial agriculture shares and employment growth (the slope of the relationship is -1.88 and the $R^2$ 0.57). Yet, the resulting migration pattern as a function of the initial agicultural share is quite similar to the benchmark economy, see Panel (b) of Figure 7. Eckert and Peters (2022) argue that early convergence of productivity across counties explains the lack of rural exodus and homogenous local industrialization in the US. Our results show that early convergence of sectoral productivity would have generated a more even pattern of industrialization in Spain, but still unable to prevent the rural exodus.

## 6 Conclusions

As countries develop, their regions industrialize at an uneven pace. We have shown how migrations from laggard to leading regions help explain the lack of industrialization of the former and the fast path of industrialization of the latter. More generally, we have shown how these migrations contribute to the overall process of growth and macroeconomic development of a country.

We have started by looking at the development experience of Spain between 1940 and 2000, characterized by fast growth in income per capita, a large structural transformation, and mass migration from rural areas, which failed to industrialize, to early industrial hubs. Our simple model of structural change with multiple locations and sectors shows that the large rural exodus in Spain originated mainly from a decline in the migration costs towards the most prosperous regions during the 1950s, 1960s, and 1970s, complemented by a divergence of sectoral productivities across regions during these same decades. The rural exodus completely explains the lack of industrialization in laggard areas, which shows how population movements can be a relevant force of local structural change.

We think of our paper as a first step towards understanding the role of the heterogeneous incidence of spatial frictions on economic development and structural change across countries. In this sense, this research agenda can help shed light on the heterogeneous paths of development documented by Rodrik (2016) and Huneeus and Rogerson (2020), among others. Our results focus on how population movements change local wages and prices, which in turn change the patterns of comparative advantage across regions. A limitation is that we take the estimated time paths of productivity, migration costs, and trade costs as exogenous. Further research can explore how to endogenize them, which would affect the counterfactual analyses. We took a first step in this direction by exploring how the time paths of region-specific sectoral productivity vary with employment due to agglomeration economies in manufacturing or decreasing returns to scale in agriculture. While qualitative results do not change, the quantitative effects of a rural exodus on the local development experience of regions are amplified. Future research could endogenize the joint evolution of migration costs and regional productivities, which is a promising explanation for the cross-country differences in local development patterns.

## References


ADAMOPOULOS, T. (2011): "Transportation Costs, Agricultural Productivity, and Cross-Country Income Differences," *International Economic Review*, 52, 489–521. (Cited on page 4.)

ALMUNIA, M., P. ANTRÀS, D. LÓPEZ-RODRÍGUEZ, AND E. MORALES (2021): "Venting Out: Exports During a Domestic Slump," *American Economic Review*, 111, 3611–3662. (Cited on page 6.)

ARTUÇ, E., S. CHAUDHURI, AND J. MCLAREN (2010): "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 100, 1008–1045. (Cited on pages 1, 5, and 23.)

BAUMOL, W. J. (1967): "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis," *American Economic Review*, 57, 415–426. (Cited on page 2.)

BENTOLILA, S. (1997): "Sticky Labor in Spanish Regions," *European Economic Review*, 41, 591–598. (Cited on page 25.)

BOPPART, T. (2014): "Structural Change and the Kaldor Facts in a Growth Model with Relative Price Effects and Non-Gorman Preferences," *Econometrica*, 82, 2167–2196. (Cited on page 2.)

BOVER, O. AND P. VELILLA (2005): "Migrations in Spain: Historical Background and Current Trends," in *European Migration: What Do We Know?*, ed. by K. Zimmermann, Oxford University Press, chap. 11. (Cited on pages 7 and A.2.)

BURSTEIN, A., G. HANSON, L. TIAN, AND J. VOGEL (2020): "Tradability and the Labor Market Impact of Immigration: Theory and Evidence for the U.S." *Econometrica*, 88, 1071–1112. (Cited on page 2.)

CALIENDO, L., M. DVORKIN, AND F. PARRO (2019): "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, 87, 741–835. (Cited on page 5.)

CAMPOS, R., I. REGGIO, AND J. TIMINI (2022): "Thick Borders in Franco's Spain: the Costs of a Closed Economy," Banco de España Documento de Trabajo 2209. (Cited on page 6.)

CARRERAS, A. AND X. TAFUNELL, eds. (2005): *Estadísticas Históricas de España: Siglos XIX-XX*, Fundación BBVA, Bilbao, second ed. (Cited on page 24.)

CASELLI, F. AND W. J. COLEMAN (2001): "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation," *Journal of Political Economy*, 109, 584–616. (Cited on page 4.)

CHEREMUKHIN, A., M. GOLOSOV, S. GURIEV, AND A. TSYVINSKI (2017): "The Industrialization and Economic Development of Russia through the Lens of a Neoclassical Growth Model," *Review of Economic Studies*, 84, 613–649. (Cited on page 5.)

Comin, D., D. Lashkari, and M. Mestieri (2021): "Structural Change with Long-run Income and Price Effects," *Econometrica*, 89, 311–374. (Cited on pages 2 and 10.)

Conesa, J. C., M. J. Delventhal, P. S. Pujolas, and G. Raveendranathan (2021): "The Cost of Trade Disruptions at Different Stages of Development," Mimeo. (Cited on page 6.)

Desmet, K., D. K. Nagy, and E. Rossi-Hansberg (2018): "The Geography of Development," *Journal of Political Economy*, 126, 903–983. (Cited on page 28.)

Desmet, K. and E. Rossi-Hansberg (2014): "Spatial Development," *American Economic Review*, 104, 1211–1243. (Cited on page 5.)

Díaz, A. and L. Franjo (2016): "Capital goods, measured TFP and Growth: The case of Spain," *European Economic Review*, 83, 19–39. (Cited on page 6.)

Duarte, M. and D. Restuccia (2010): "The Role of the Structural Transformation in Aggregate Productivity," *Quarterly Journal of Economics*, 125, 129–173. (Cited on pages 1 and 2.)

Eaton, J. and S. Kortum (2002): "Technology, Geography and Trade," *Econometrica*, 70, 1741–1779. (Cited on pages 1, 5, 12, and 22.)

Eckert, F. (2019): "Growing Apart: Tradable Services and the Fragmentation of the U.S. Economy," Mimeo. (Cited on pages 2 and 5.)

Eckert, F. and M. Peters (2022): "Spatial Structural Change," Mimeo. (Cited on pages 1, 4, 31, A.7, and A.20.)

Fajgelbaum, P. and S. Redding (2022): "Trade, Structural Transformation and Development: Evidence from Argentina 1869-1914," *Journal of Political Economy*, 130, 1249–1318. (Cited on page 5.)

Fan, T., M. Peters, and F. Zilibotti (2021): "Service-Led or Service-Biased Growth? Equilibrium Development Accounting across Indian Districts," NBER Working Paper 28551. (Cited on pages 5 and A.7.)

Finicelli, A., P. Pagano, and M. Sbracia (2013): "Ricardian Selection," *Journal of International Economics*, 89, 96–109. (Cited on page 15.)

Franck, R. and O. Galor (2021): "Flowers of Evil? Industrialization and Long Run Development," *Journal of Monetary Economics*, 117, 108–128. (Cited on page A.7.)

García-Peñalosa, C. and V. Bignon (2022): "The toll of tariffs: Protectionism, Education and Fertility in Late 19th Century France," Unpublished Manuscript. (Cited on page A.7.)

García-Santana, M., E. Moral-Benito, J. Pijoan-Mas, and R. Ramos (2020): "Growing like Spain: 1995-2007," *International Economic Review*, 61, 383–416. (Cited on pages 6 and 21.)

García-Santana, M., J. Pijoan-Mas, and L. Villacorta (2021): "Investment Demand and Structural Change," *Econometrica*, 89, 2751–2785. (Cited on pages 2 and 18.)

Garriga, C., A. Hedlund, Y. Tang, and P. Wang (2017): "Rural-Urban Migration, Structural Transformation, and Housing Markets in China," Mimeo. (Cited on page 4.)

Gervais, A. and J. B. Jensen (2019): "The Tradability of Services: Geographic Concentration and Trade Costs," *Journal of International Economics*, 118, 331–350. (Cited on pages 2 and 20.)

Gollin, D., S. Parente, and R. Rogerson (2007): "The Food Problem and the Evolution of International Income Levels," *Journal of Monetary Economics*, 54, 1230–1255. (Cited on page 28.)

Gollin, D. and R. Rogerson (2014): "Productivity, Transport Costs and Subsistence Agriculture," *Journal of Development Economics*, 104, 38–48. (Cited on page 4.)

Gómez-Tello, A., A. Díez-Minguela, J. Martínez-Galarraga, and D. Tirado (2019): "Regional Prices in Early Twentieth-Century Spain: a Country-Dummy Approach," *Cliometrica*, 13, 245–276. (Cited on page A.1.)

Gopinath, G., S. Kalemli-Ozcan, L. Karabarbounis, and C. Villegas-Sanchez (2017): "Capital Allocation and Productivity in South Europe," *Quarterly Journal of Economics*, 132, 1915–1967. (Cited on page 6.)

Hao, T., R. Sun, T. Tombe, and X. Zhu (2020): "The Effect of Migration Policy on Growth, Structural Change, and Regional Inequality in China," *Journal of Monetary Economics*, 113, 112–134. (Cited on pages 4, 29, and A.7.)

Hayami, Y. and V. W. Ruttan (1985): *Agricultural Development: An International Perspective*, Baltimore: Johns Hopkins University Press. (Cited on page 28.)

Herrendorf, B., R. Rogerson, and A. Valentinyi (2013): "Two Perspectives on Preferences and Structural Transformation," *American Economic Review*, 103, 2752–2789. (Cited on pages 2, 18, and 19.)

——— (2014): "Growth and Structural Transformation," in *Handbook of Economic Growth*, ed. by P. Aghion and S. Durlauf, Elsevier Science Publishers, vol. 2, chap. 6, 855–941. (Cited on pages 1 and 6.)

——— (2021): "Structural Change in Investment and Consumption: a Unified Approach," *Review of Economic Studies*, 88, 1311–1346. (Cited on page 18.)

Herrendorf, B., J. A. Schmitz, and A. Teixeira (2012): "The Role of Transportation in U.S. Economic Development: 1840-1860," *International Economic Review*, 53, 693–716. (Cited on page 4.)

HUNEEUS, F. AND R. ROGERSON (2020): "Heterogeneous Paths of Industrialization," NBER Working Paper 27580. (Cited on page 31.)

KLEINMAN, B., E. LIU, AND S. J. REDDING (2021): "Dynamic Spatial General Equilibrium," NBER Working Paper 29101. (Cited on page 26.)

KONGSAMUT, P., S. REBELO, AND D. XIE (2001): "Beyond Balanced Growth," *Review of Economic Studies*, 68, 869–882. (Cited on page 2.)

KUZNETS, S. (1955): "Economic Growth and Income Inequality," *American Economic Review*, 1–28. (Cited on page 9.)

LAGAKOS, D., A. M. MOBARAK, AND M. E. WAUGH (2023): "The Welfare Effects of Encouraging Rural–Urban Migration," *Econometrica*, 91, 803–837. (Cited on page 28.)

LAGAKOS, D. AND M. WAUGH (2013): "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economic Review*, 103, 948–980. (Cited on page 26.)

LESSMANN, C. (2014): "Spatial Inequality and Development - Is There an Inverted-U Relationship?" *Journal of Development Economics*, 106, 35–51. (Cited on page 9.)

LEWIS, L. T., R. MONARCH, M. SPOSI, AND J. ZHANG (2021): "Structural Change and Global Trade," *Journal of the European Economic Association*, 20, 476–512. (Cited on page 5.)

LÓPEZ SIMÓN, I. (2022): *Los Olvidados: Marginalidad Urbana y Fenómeno Quinqui en España (1959-1982)*, Madrid: Marcial Pons Historia. (Cited on page 24.)

MARTÍN, A., E. MORAL-BENITO, AND T. SCHMITZ (2021): "The Financial Transmission of Housing Booms: Evidence from Spain," *American Economic Review*, 111, 1013–1053. (Cited on page 6.)

MICHAELS, G., F. RAUCH, AND S. REDDING (2012): "Urbanization and Structural Transformation," *Quarterly Journal of Economics*, 127, 535–586. (Cited on page 5.)

MORTEN, M. AND J. OLIVEIRA (2018): "The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City," NBER Working Paper 22158. (Cited on pages 5 and 24.)

NAGY, D. K. (2020): "Hinterlands, City Formation and Growth: Evidence from the U.S. Westward Expansion," Mimeo. (Cited on page 5.)

NGAI, R. AND C. PISSARIDES (2007): "Structural Change in a Multisector Model of Growth," *American Economic Review*, 97, 429–443. (Cited on page 2.)

NGAI, R., C. PISSARIDES, AND J. WANG (2019): "China's Mobility Barriers and Employment Allocations," *Journal of the European Economic Association*, 17, 1617–1653. (Cited on page 4.)

Porzio, T., F. Rossi, and G. Santangelo (2022): "The Human Side of Structural Transformation," *American Economic Review*, 112, 2774–2814. (Cited on page 4.)

Prados de la Escosura, L. (2017): *Spanish Economic Growth, 1850-2015*, London: Palgrave MacMillan. (Cited on pages 6 and A.1.)

Prados de la Escosura, L. and J. R. Rosés (2021): "Accounting for Growth: Spain, 1850-2019," *Journal of Economic Surveys*, 35, 804–832. (Cited on page 6.)

Restuccia, D., D. T. Yang, and X. Zhu (2008): "Agriculture and Aggregate Productivity: a Quantitative Cross-country Analysis," *Journal of Monetary Economics*, 55, 234–250. (Cited on page 28.)

Rodrik, D. (2016): "Premature Deindustrialization," *Journal of Economic Growth*, 21, 1–33. (Cited on page 31.)

Silva, J. M. C. S. and S. Tenreyro (2006): "The Log of Gravity," *The Review of Economics and Statistics*, 88, 641–658. (Cited on page 24.)

Simonovska, I. and M. Waugh (2014): "The Elasticity of Trade: Estimates and Evidence," *Journal of International Economics*, 92, 34–50. (Cited on page 17.)

Song, Z., K. Storesletten, and F. Zilibotti (2011): "Growing Like China," *American Economic Review*, 101, 202–241. (Cited on page 5.)

Sposi, M. (2019): "Evolving Comparative Advantage, Sectoral Linkages, and Structural Change," *Journal of Monetary Economics*, 103, 75–87. (Cited on pages 5, 15, and 17.)

Świecki, T. (2017): "Determinants of Structural Change," *Review of Economic Dynamics*, 24, 95–13. (Cited on page 5.)

Tombe, T. and X. Zhu (2019): "Trade, Migration, and Productivity: A Quantitative Analysis of China," *American Economic Review*, 109, 1843–1872. (Cited on pages 4 and 24.)

Uy, T., K.-M. Yi, and J. Zhang (2014): "Structural Change in an Open Economy," *Journal of Monetary Economics*, 60, 667–682. (Cited on page 5.)

Ventosa, J. R. (2017): "La configuración de la red: 1960-2017. Crecimiento y modernización de las carreteras españolas," *Revista del Ministerio de Fomento*, 674, 72–79. (Cited on page 21.)

Williamson, J. G. (1965): "Regional Inequality and the Process of National Development: a Description of Patterns," *Economic Development and Cultural Change*, 13, 3–45. (Cited on page 9.)

Online Appendix

# Macroeconomic Development, Rural Exodus, and Uneven Industrialization

Tomás Budí-Ors    Josep Pijoan-Mas

June 2025

## Table of Contents

## Appendix A: Data

### A.1 Data sources

Our analysis combines data from different sources, which we detail in the following paragraphs.

**Employment.** Data on regional employment for each sector between 1940 and 2000 comes from the Spanish Population Census, conducted every ten years by the National Statistics Office (*Instituto Nacional de Estadística*). We aggregate employment in agriculture, hunting, forestry and fishing and classify it as agriculture; employment in manufacturing, mining, construction and utilities and classify it as manufacturing; and employment in trade, transport, business, government and personal services and classify it as services.[26] In Figure 1 Panels (a) and (b) we use historical data from Prados de la Escosura (2017), which can be accessed here.

**Migration flows.** Information on bilateral migration flows is retrieved as well from the Census, which reports, for each province, the number of people that lived in a different province in the previous census wave, separating this number by migrants' region of origin since 1960 (i.e. we know, for instance, the number of people living in Barcelona in 1970 who were living in València in 1960, and the same for every pair of provinces). The procedure we follow to go from observed bilateral migration flows to our model-consistent bilateral migration flows is discussed in detail in Section A.2 of the Appendix.

**Value added and prices.** Data on regional value added and price indices by sector are obtained from the regional accounts prepared by the research department of BBVA.[27] To get a time series of regional price levels, we combine the regional price indices, which tell us how prices changed over time for each province and sector, with a cross-section of regional price levels for 1930. The latter is obtained from micro data on the prices of a common basket of goods across Spanish provinces, and were gathered by the *Instituto de Reformas Sociales*, a government institution in charge of assessing the material living conditions of the labor force at the time. We direct the interested reader to Gómez-Tello et al. (2019), to whom we are thankful for kindly sharing the data with us.

[26]For most years, we obtain the data at four-sector level (agriculture, manufacturing, construction, and services) directly from INE, which follows the aggregation criteria outlined above.

[27]Starting in 1957, the research department of BBVA published, every two years, a volume with the main economic aggregates of Spanish regions. In 2003, they released a volume with revised information and longer time coverage, with information dating back to 1930, which can be downloaded here.

## A.2 From data on migration flows to model counterparts

Our data on inter-regional migration flows $\hat{\rho}_{\ell rt}$ spans 1960 to 2000. We make two corrections to these data.

The first one comes from the fact that the observed migration flows $\hat{\rho}_{\ell rt}$ do not perfectly square with the data on employment, that is, $L_{rt}$ is not exactly equal to $(1+n_t)\sum_{\ell}^{R}\hat{\rho}_{\ell r}L_{\ell t-1}$.[28] To find the migration flows $\rho_{\ell rt}$ that are consistent with the data on employment, we simply minimize the Euclidean distance with respect to the observed $\hat{\rho}_{\ell rt}$ subject to the law of motion of employment being satisfied. In other words, we search for the most similar matrix to the observed matrix of bilateral migration flows that replicates the observed changes in the distribution of labor across regions. We further impose the constraints that the entries in $\rho_{\ell rt}$ are non-negative and that the elements in each row add up to 1. Then, the problem we solve is:

$$
\begin{aligned}
\min_{\rho_{\ell rt}\,\forall r,\ell} \quad & \sqrt{\sum_{r=1}^{R}\sum_{\ell=1}^{R}(\rho_{\ell rt}-\hat{\rho}_{\ell rt})^2} \\
\text{s.t.} \quad & L_{rt} = (1+n_t)\sum_{r=1}^{R}\rho_{\ell rt}L_{\ell t-1} \qquad \forall r, \\
& \sum_{r=1}^{R}\rho_{\ell rt} = 1 \qquad \forall \ell, \\
& \rho_{\ell rt} \geq 0 \qquad \forall \ell, r.
\end{aligned} \tag{A.1}
$$

The second correction comes from the fact that data on bilateral migration flows $\hat{\rho}_{\ell rt}$ is only available from 1960 onward, as in previous Census waves there is no question regarding workers' region of previous residence. Therefore, we cannot retrieve $\rho_{\ell rt}$ for 1950 and 1940 by solving (A.1). Instead, we look for $\rho_{\ell r1950}$ that minimize the Euclidean distance with $\rho_{\ell r1960}$ (which is based on observed migration flows), and similarly for $\rho_{\ell r1940}$ with $\rho_{\ell r1950}$ as target; with constraints as in (A.1). Additionally, we refine our strategy using data on net migration in each region for both 1950 and 1940, such that $\rho_{\ell r1950}$ and $\rho_{\ell r1940}$ are consistent with the volume of internal migration

[28] This happens for two reasons. First, there may be possible measurement error in either the migration flows or the employment stocks. Second, regional employment growth in the model only depends on net internal migration and aggregate population growth. However, in the data it also depends on additional factors as for instance regional differences in the unemployment rate, in the fertility and mortality rates, or in the incidence of international migrations. Foreign immigration is Spain started in the 00's, which does not affect our period of study. However, a significant number of workers in rural areas migrated abroad between 1960 and 1973 (mostly to Germany, France, and Switzerland). The exact numbers are unclear, see Bover and Velilla (2005) for details.

in Spain in 1940 and 1950.[29] Specifically, we solve the following problem for $t = 1950$ and $t = 1940$

$$\begin{aligned}
\min_{\rho_{\ell r t}\,\forall r,\ell} \quad & \sqrt{\sum_{r=1}^{R}\sum_{\ell=1}^{R}(\rho_{\ell r t} - \hat{\rho}_{\ell r t+1})^2} \\
\text{s.t.} \quad & L_{rt} = (1 + n_t)\sum_{r=1}^{R}\rho_{\ell r t}L_{\ell t-1} && \forall r, \\
& \sum_{r=1}^{R}\rho_{\ell r t} = 1 && \forall \ell, \\
& \rho_{\ell r t} \geq 0 && \forall \ell, r, \\
& NM_{rt} = \sum_{\ell=1,\ell\neq r}^{R}\rho_{\ell r t}L_{\ell t-1} - (1 - \rho_{rr})L_{rt-1} && \forall r\,.
\end{aligned} \tag{A.2}$$

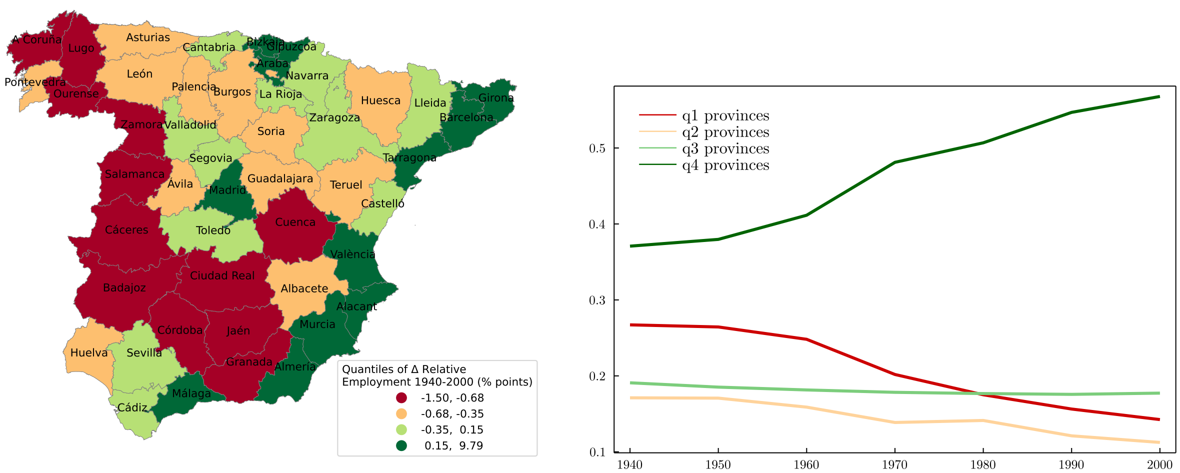
where $NM_{rt}$ is the net migration in each region in period $t$.

[29] Net migration is defined as the difference between total immigrants inflows and outflows. The data comes from the Census and is retrieved as a residual, given that for 1940 and 1950 we have information on births, deaths and population stocks for each province.

## Appendix B: Regional employment growth in Spain, 1940-2000

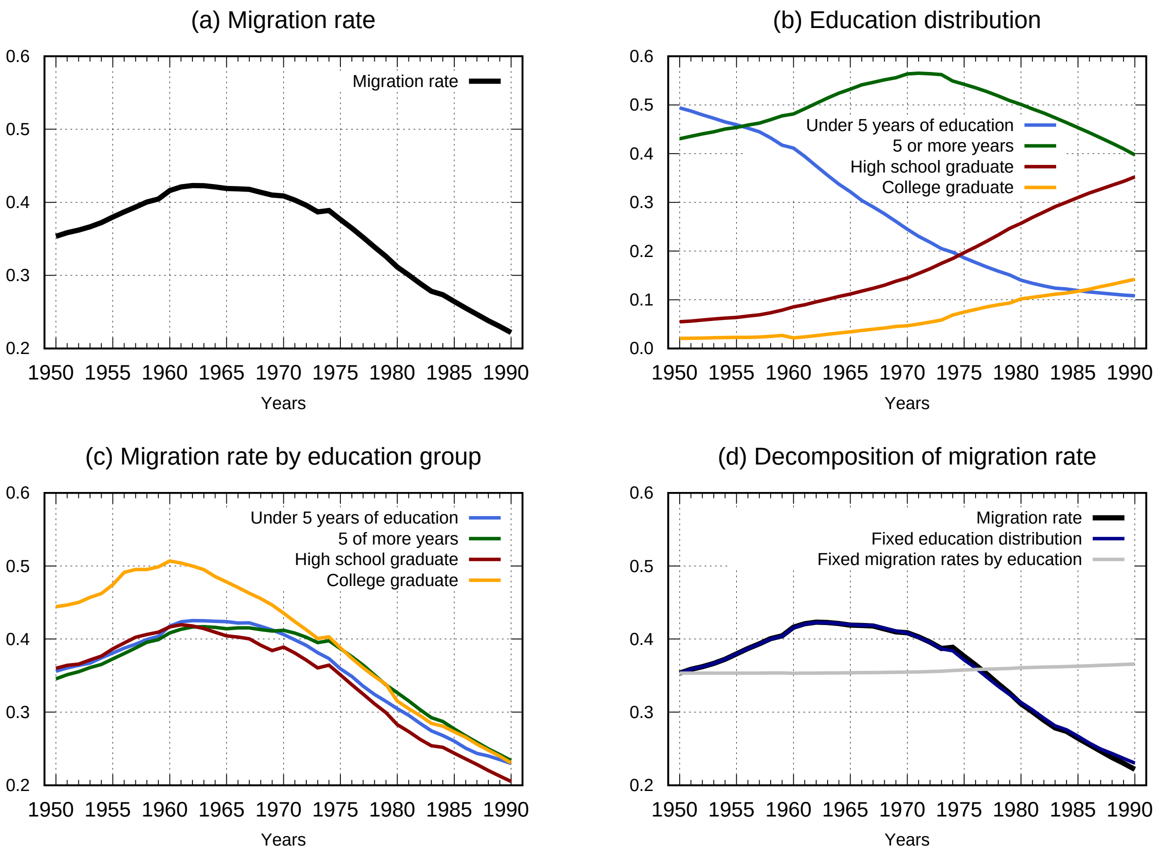
Throughout the paper we stress the heterogeneity in employment growth across provinces and how this relates to the sectoral composition of regional employment in 1940 (Figure 2). For many statistics, we classify provinces in four quartiles based on the change in their relative size within the country (measured as the province share of total employment) over the period 1940-2000. The provinces corresponding to each group are shown in the left panel of Figure B.1. The share of total employment represented by each group over time is shown in the right panel of Figure B.1.

Figure B.1: Regional employment growth 1940-2000




**Notes:** provinces in red are part of the first quartile of relative employment growth (q1), in yellow of the second quartile (q2), in light green of the third quartile (q3), and in dark green of the fourth quartile (q4).

Figure C.1: Migration and education, 1950-1979




**Notes:** Panels (a), (c), and (d) plot the fraction of people aged 18-25 year old in the given year that were observed as living in a province different from their birth province in some of the subsequent censuses (when they were aged between 26 and 56 years of age). Panel (b) plots the fraction of 18-25 year old in the given year that hold each education level, also according to the subsequent censuses.

## Appendix C: Education changes and migrations in Spain

In this Appendix we document the small role played by changes in education on the rural exodus in Spain. To do so, we use micro-data for the censuses of 1981, 1991, 2001, and 2011 (the only available ones).[30] From each census we select individuals that were aged 18 to 25 years old for each year between 1950 and 1979. We only use individuals that are observed in the census with ages between 26 and 56.[31] We classify these individuals as migrants if, when observed in the census, they reside in a different province from birth (and as non-migrants otherwise). The assumption is that most migrants moved in the age range 18 to 25 and hence the comparison of the two groups reveals

[30] The microdata comes from the IPUMS International Census Database. All cases corresponds to 5% samples of the census, with the exception of 2011 that is a 10% sample.

[31] We do not use older individuals to limit biases due to differential mortality across education or migration groups, and also to minimize the incidence of return migrations, which in Spain is typically linked to retirement.

differences between migrants and non-migrants. We collect education data in four categories: less than five years of education (50% of 18-25 year olds in 1950), 5 or more years of education but no high school degree (43%), high school or vocational school degree (5%), and college degree (2%).

In Panel (a) of Figure C.1 we plot the migration rate for 18-25 year old individuals between 1950 and 1979. We see how the migration rate increases until the early 1960s and declines afterward. In Panel (b) we observe a clear educational transition over the period. The share of 18-25 year old individuals with less than 5 years of schooling falls steadily, the share of 18-25 year old individuals with 5 or more years but no high school degree increases first and declines later, and the shares of high school graduates and college graduates increase. Additionally, we note that the share of 18-25 year old college educated individuals is still very small in 1980, around 10% of the population in that age group. In Panel (c) we see that over this period migration rates were only partially linked to education: while the migration rate of college graduates is substantially larger than for the other education groups, there is no clear education gradient of migration among the non-college educated (the vast majority of young population). In particular, individuals with less than 5 years of schooling migrated at least as much (if not more than) more educated individuals without college. As a result, the educational transition did not produce mechanically a large change in migration rates. In Panel (d) we decompose the migration rate for the 18-25 year old individuals by isolating changes coming from changes in education while keeping migration rates by education constant, and changes in migration rate by education while keeping the education distribution constant. We see that if we keep the migration rates by education at their 1950 level and only let the distribution of education change over time, there is no increase in migration over the period. Rather, the change in migration comes from the change in migration within education groups.

## Appendix D: International evidence on rural exodus

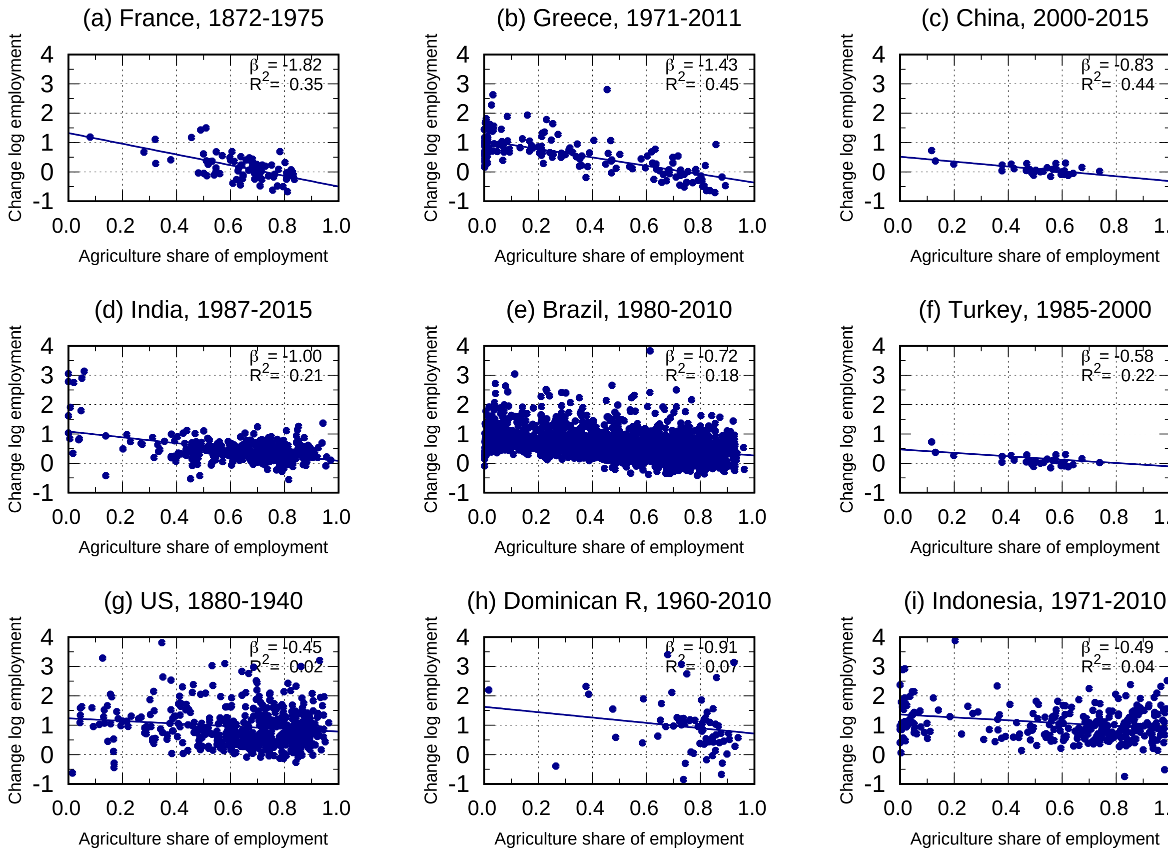
We use census data from several countries to explore the relationship between regional employment growth and the initial sectoral composition of regions. Our main data set is the IPUMS International Census Database. We complement these data with our own data for Spain, data from Hao et al. (2020) for China, data from García-Peñalosa and Bignon (2022) and Franck and Galor (2021) for France, data from Fan et al. (2021) for India, and data from Eckert and Peters (2022) for the US. We want to focus on development episodes, so we restrict the sample to countries such that (i) the time span between the first and the last year observed is larger than 10 years, (ii) the initial share of employment agriculture is at least 25% and (iii) the fall in the country-level agricultural share in the period considered is larger than 10 percentage points. This gives us 27 development episodes. Then, for every country we run the regression,

$$\Delta \log L_r = \alpha + \beta \frac{L_{ra}}{L_r} + \varepsilon_r$$

where $L_{ra}/L_r$ is the share of region $r$ employment in sector $a$ in the first year of the development episode and $\Delta \log L_r$ is total employment growth in region $r$ between the first and last year of observation.

Table D.1 reports the results. For most countries $\beta < 0$, which means that more agrarian regions tend to lose population in relation to the rest (the only exceptions being Haiti, 1982-2003, and Honduras, 1961-2001). However both the magnitudes of $\beta$ and the $R^2$ vary substantially across countries. The magnitude of $\beta$ is largest for Spain (1940-2000), with a value of -2.07, which means that a province with 10 percentage points higher share of employment in agriculture in 1940 experienced a 20% smaller population growth between 1940 and 2000. Remarkably in the case of Spain, 63% of the variance in employment growth across provinces is related to the initial share of employment in agriculture. Other development episodes where the initial share of agriculture across locations is strongly related to employment growth are France (1872-1975), Greece (1971-2011), Bangladesh (1991-2011), Senegal (1988-2013), and China (2000-2015), with slopes equal to -1.82, -1.43, -1.36, -1.27, and -0.83, and $R^2$ equal to 34%, 45%, 44%, 60%, and 44% respectively. Some intermediate cases are India (1987-2011), Brazil (1980-2010), and Turkey (1985-2000) with slopes equal to -1.00, -0.72, and -0.58, and $R^2$ equal to 21%, 18%, and 22% respectively. In contrast, in some salient development episodes like the US (1880-1940), Indonesia (1971-2010), Dominican Republic (1960-2010), or Costa Rica (1963-2011) migration flows are scarcely related to initial sectoral composition. For instance, the slope for these four countries is -0.45, -0.49, -0.91, and -0.70 respectively but the $R^2$ of the regression is only 2%, 6%, 3%, and 9% respectively.

Figure D.1: Rural exodus across development episodes




**Notes:** This figure plots the relative increase in employment (in logs) across sub-national divisions against the initial sectoral shares of employment in agriculture for several countries. Each panel also reports the slope of the relationship ($\beta$) and the share of variance in log employment growth explained by the corresponding x-axis variables ($R^2$).

Table D.1: Rural exodus across development episodes

| Country (1) | Period (2) | Ini Agr Sh (3) | $\Delta$ Agr Sh (4) | N (5) | $\hat{\beta}$ (6) | $R^2$ (7) |
|---|---|---|---|---|---|---|
| Bangladesh | 1991-2011 | 69.8 | 25.1 | 64 | $-1.36^{***}$ | 0.44 |
| Benin | 1979-2013 | 59.4 | 16.8 | 77 | $-0.05$ | 0.00 |
| Bolivia | 1976-2012 | 45.3 | 18.8 | 80 | $-0.83^{***}$ | 0.15 |
| Brazil | 1980-2010 | 29.1 | 15.1 | 2040 | $-0.72^{***}$ | 0.18 |
| Cambodia | 1998-2013 | 76.5 | 13.9 | 141 | $-0.40^{***}$ | 0.05 |
| China | 2000-2015 | 52.9 | 24.6 | 30 | $-0.83^{***}$ | 0.44 |
| Costa Rica | 1963-2011 | 48.1 | 35.1 | 55 | $-0.70^{*}$ | 0.09 |
| Dominican Rep. | 1960-2010 | 62.1 | 52.8 | 65 | $-0.91$ | 0.03 |
| Ecuador | 1962-2010 | 58.1 | 36.4 | 77 | $-1.74^{***}$ | 0.21 |
| El Salvador | 1992-2007 | 35.0 | 18.5 | 103 | $-0.54^{***}$ | 0.20 |
| France | 1872-1975 | 57.6 | 47.7 | 85 | $-1.82^{***}$ | 0.34 |
| Greece | 1971-2011 | 38.7 | 30.1 | 156 | $-1.43^{***}$ | 0.45 |
| Guatemala | 1964-2002 | 65.0 | 25.8 | 191 | $-1.27^{***}$ | 0.14 |
| Haiti | 1982-2003 | 33.7 | 14.0 | 19 | $0.84^{**}$ | 0.20 |
| Honduras | 1961-2001 | 66.2 | 24.6 | 96 | $1.20$ | 0.05 |
| India | 1987-2011 | 63.9 | 16.9 | 368 | $-1.00^{***}$ | 0.21 |
| Indonesia | 1971-2010 | 64.6 | 26.4 | 268 | $-0.49^{***}$ | 0.06 |
| Malaysia | 1970-2000 | 52.5 | 37.3 | 101 | $-0.72^{***}$ | 0.12 |
| Mali | 1987-2009 | 81.0 | 13.8 | 47 | $-0.40$ | 0.08 |
| Mexico | 1970-2015 | 40.6 | 30.4 | 2321 | $-1.17^{***}$ | 0.11 |
| Nicaragua | 1971-2005 | 48.1 | 14.4 | 68 | $-0.26$ | 0.01 |
| Panama | 1960-2010 | 46.3 | 34.5 | 35 | $-1.14^{*}$ | 0.11 |
| Paraguay | 1962-2002 | 54.1 | 28.1 | 60 | $-1.00^{**}$ | 0.05 |
| Senegal | 1988-2013 | 61.5 | 35.2 | 27 | $-1.27^{***}$ | 0.60 |
| Spain | 1940-2001 | 51.9 | 45.4 | 47 | $-2.07^{***}$ | 0.63 |
| Turkey | 1985-2000 | 55.8 | 12.3 | 114 | $-0.58^{***}$ | 0.22 |
| United States | 1880-1940 | 51.2 | 33.3 | 506 | $-0.45^{***}$ | 0.02 |

Notes: this table shows the relationship between the initial agricultural share of regional employment and subsequent regional employment growth for a group of selected countries. Selected countries meet the following criteria: (i) the time span between the first and the last year observed is larger than 10 years, (ii) the initial share of employment agriculture is at least 25% and (iii) the fall in the country-level agricultural share in the period considered is larger than 10 percentage points. For all countries except China, France, India, Spain and the US (see main text) data comes from IPUMS International Census Database. Regional-level employment is constructed by aggregating microdata on employed individuals between ages 20 and 59. The coefficient $\hat{\beta}$ reported in Column (6) is the point estimate of a regression of log employment growth at the regional level on the initial agricultural share of regional employment.

## Appendix E: Model details

### E.1 Analytic results

**Intersectoral and intrasectoral trade.** The goods market clearing condition, equation (9), can be rewritten as

$$P_{rj}Y_{rj} - \pi_{rrj}P_{rj}C_{rj} = \sum_{\ell \neq r}^{R} \pi_{r\ell j}P_{\ell j}C_{\ell j} \quad \forall r, j \tag{E.1}$$

giving us the gross exports of sector-j goods by region $r$. Note that the gross imports of sector-j goods by region $r$ is given by $(1-\pi_{rrj})\,P_{rj}C_{rj}$ (the fraction of sector-j expenditure sourced from other regions). We can define intrasectoral trade as the sum of sectoral gross imports or exports across regions: $\sum_r (1-\pi_{rrj})\,P_{rj}C_{rj}$ or $\sum_r [P_{rj}Y_{rj} - \pi_{rrj}P_{rj}C_{rj}]$, which are equal to each other because there is no international trade ($\sum_r P_{rj}C_{rj} = \sum_r P_{rj}Y_{rj}$). Then, net exports of sector-j goods by region $r$ is given by the difference of gross exports and gross imports of that sector,

$$NX_{rj} = \left[P_{rj}Y_{rj} - \pi_{rrj}P_{rj}C_{rj}\right] - \left[\,(1-\pi_{rrj})\,P_{rj}C_{rj}\right] = P_{rj}Y_{rj} - P_{rj}C_{rj}$$

which equals the difference between production and expenditure in that region-sector. We define intersectoral trade as the sum across regions of the positive sectoral net exports, which equals the sum of positive sectoral net imports because there is no international trade ($\sum_r P_{rj}C_{rj} = \sum_r P_{rj}Y_{rj}$). In particular, this would be $\sum_r \frac{1}{2}\left|P_{rj}Y_{rj} - P_{rj}C_{rj}\right|$

**Trade balance equation.** To derive the trade balance condition (13), note that the budget constraint of the households in equation (2) can be aggregated at the region level as

$$\sum_j P_{rj}C_{rj} = \sum_j w_r L_{rj} = \sum_j P_{rj}Y_{rj} \Rightarrow \sum_j \left[P_{rj}Y_{rj} - P_{rj}C_{rj}\right] = 0 \tag{E.2}$$

which says that sectoral net exports have to add up to zero at the regional level. Plugging the definition of net exports as the difference of gross exports and gross imports into equation (E.2), we obtain the equilibrium equation (13).

**Proposition 1.** *Holding prices fixed and assuming $R$ is large enough, a decline in population $L_r$ in region $r$ leads to (a) a fall in labor demand $L_{rj}$ in all sectors $j$, and (b) a reallocation of labor away from sectors $j$ where local purchases represent a larger fraction of value added.*

*Proof.* Holding prices fixed and assuming that $R$ is large enough to dismiss the effects of changes in demand coming from other regions, the derivative of $L_{rj}$ in equation (12) with respect to $L_r$ can be written as

$$\frac{\partial L_{rj}}{\partial L_r} = \frac{\pi_{rrj} P_{rj} C_{rj}}{P_r Y_r} > 0 \tag{E.3}$$

which proves (a) as both numerator and denominator are strictly positive. To prove (b) note that multiplying and dividing the right-hand-side of equation (E.3) by $P_{rj}Y_{rj}$ easily gives us equation (16). Equation (16) shows that the elasticity $\epsilon_{jr} \equiv \frac{\partial L_{rj}}{\partial L_r}\frac{L_r}{L_{rj}}$ of changes in the sector $j$ labor demand $L_{rj}$ with respect to changes in the population $L_r$ is given by the ratio of $\pi_{rrj}P_{rj}C_{rj}$ to $P_{rj}Y_{rj}$. Actually, note that labor reallocation can we written as

$$\frac{\partial}{\partial L_r}\left(\frac{L_{rj}}{L_{ra}+L_{rm}+L_{rs}}\right) \propto \epsilon_{jr} - \bar{\epsilon}_r$$

where $\bar{\epsilon}_r \equiv \sum_i \epsilon_{ir}\frac{L_{ri}}{L_r}$, such that sectors where $\epsilon_{jr}$ is larger than the average elasticity $\bar{\epsilon}_r$ increase their relative size when there is a population inflow. □

**Proposition 2.** *Holding prices fixed and assuming $R$ is large enough, a decline in population $L_r$ in region $r$ leads to (a) a fall in total labor demand $\sum_j L_{rj}$ that is (weakly) lower than the fall in population $L_r$, and (b) a (weak) increase in the local wage $w_r$.*

*Proof.* Add up equation (E.3) over all sectors $j$ to get,

$$\frac{\partial}{\partial L_r}\sum_j L_{rj} = \frac{\sum_j \pi_{rrj}P_{rj}C_{rj}}{P_r Y_r} \leq 1$$

where the weak inequality follows from the trade balance condition, $\sum_j P_{rj}C_{rj} = P_r Y_r$ and the amount of trade across regions, that is, $\pi_{rrj} \leq 1\ \forall j$. Indeed, note that without trade $\pi_{rrj} = 1\ \forall j$ and the decline in labor demand would be equal to the decline in labor supply, preventing changes in the equilibrium wage $w_r$. □

**Proposition 3.** *Holding expenditure shares constant, an increase in region $r$ wage $w_r$ decreases labor demand in all sectors $j$ in region $r$, more so in those that are more tradable, that is, those with lower trade costs.*

*Proof.* The fraction of sector $j$ goods that region $\ell$ buys from region $r$ is given by the $\pi_{r\ell j}$ in equation (8). An increase in region $r$ wage, $w_r$, will decrease this fraction $\forall \ell$ (including $\ell = r$) because region $r$ becomes less competitive (the prices it offers to all regions $\ell$ are larger). To see this note that we can rewrite $\pi_{r\ell j}$ as $\pi_{r\ell j} = [1 + A_{r\ell j}]^{-1}$ where

$$A_{r\ell j} \equiv \sum_{k\neq r}\left(\frac{w_r}{w_k}\frac{\tau_{r\ell j}}{\tau_{k\ell j}}\right)^{\theta_j}\frac{T_{kj}}{T_{rj}} \tag{E.4}$$

is the inverse of how competitive is region $r$ in selling sector $j$ goods to region $\ell$ in relation to all other regions (including $\ell$ itself). Now, we can show that

$$\frac{\partial \pi_{r\ell j}}{\partial w_r} = -\theta_j \left[1 + A_{r\ell j}\right]^{-2} A_{r\ell j} \frac{1}{w_r} < 0$$

In addition, we can show that the loss of market share is larger in sectors that are more tradable. To see this, let's first decompose $\tau_{r\ell j} = \tau_j \tau_{r\ell}$ into a sector-specific term $\tau_j$ and a bilateral term $\tau_{\ell r}$. Then,

$$\frac{\partial^2 \pi_{r\ell j}}{\partial w_r \partial \tau_j} = \frac{\partial^2 \pi_{r\ell j}}{\partial w_r \partial A_{r\ell j}} \frac{\partial A_{r\ell j}}{\partial \tau_j} \propto -\left(1 - A_{r\ell j}^2\right) \frac{\partial A_{r\ell j}}{\partial \tau_j}$$

Now, to sign this derivative we need two pieces. First, $\frac{\partial A_{r\ell j}}{\partial \tau_j} > 0$ because the effect of an increase in $\tau_j$ cancels in all ratios $\frac{\tau_{r\ell j}}{\tau_{k\ell j}}$ in equation (E.4) except for the case $k = l$ where $\frac{\tau_{r\ell j}}{\tau_{k\ell j}} = \tau_{r\ell j}$ because $\tau_{\ell\ell j} = 1$. That is, an increase in $\tau_j$ decreases the ability of region $r$ to sell sector $j$ goods in region $\ell$ because the competition with local goods $\ell$ is tougher. Second, we have that $A_{r\ell j} > 1$ whenever $\pi_{r\ell j} < 1/2$. Hence, with $R$ large enough this will be the case and hence we will have $\left(1 - A_{r\ell j}^2\right) < 0$. Therefore, whenever $\pi_{r\ell j} < 1/2$, the decline of $\pi_{r\ell j}$ with a wage increase is larger in sectors where trade costs are lower, that is, in sectors with lower $\tau_j$. □

### E.2 Solving for the equilibrium.

The problem of finding equilibrium prices and allocations can be simplified to finding the vector of regional wages $\{w_r\}_{r=1}^R$ that clears the regional labor markets in equation (10). Once the equilibrium wages are pinned down, the rest of equilibrium objects obtains easily. Note that labor demand in (12) only depends on wages $\{w_r\}_{r=1}^R$ and on the supply of workers in each region $\{L_r\}_{r=1}^R$ because $c_{\ell j}$, $P_{\ell j}$, and $\pi_{r\ell j}$ are all functions of wages given by equations (3), (7), and (8) respectively. The labor supply in equation (6) depends only on wages too as the $\rho_{r\ell j}$ are characterized by equation (5) and depend only on prices and wages.

## Appendix F: Quantitative analysis

### F.1 Engines of development

Our model features three main engines of development: the change in productivity, the change in trade costs, and the change in migration costs. Moreover, even if all these parameters remain constant, the economy features two other sources of dynamics. First, there will be some reallocation of employment across regions because the initial spatial distribution of population is not necessarily the one of the corresponding steady state. Second, the share parameters of the CES demand system $\omega_{jt}$ change over time,[32] reallocating expenditure across sectors and hence affecting the allocation of employment across regions and sectors.

To understand the importance of the different engines for the Spanish development experience, we run a number of counterfactual exercises in which we allow engines to change one by one. The main results of these exercises are presented in Table F.1.[33] First, we can see in Column (3) that population dynamics have almost no effect on real income per capita or the aggregate allocation of employment across sectors, despite implying substantial reallocation of employment across regions (the coefficient of variation of employment increases by 0.28 points compared to 0.63 in the benchmark). Next, in Column (4), we can see that changes in preferences generate substantial structural change out of agriculture and into services, due to the fact that $\omega_{jt}$ increases over time for services relative to agriculture, see Table H.1. This reallocation of employment across sectors combined with the spatial reallocations due to population dynamics increases real income per capita by 15 additional percentage points.

Productivity growth is the main engine of Spanish development between 1940 and 2000, see column (5). This follows from the fact that changes in productivity are the main driver of changes in relative prices and in real income, see Figure F.1. Indeed, productivity growth generates a larger increase in real income than the one of the benchmark economy (8.61 *vs.* 8.31). It generates as well most of the reallocation out of agriculture (-40 p.p.) and toward services (34 p.p.). Productivity growth is the main driver of the increase in the manufacturing employment share up to 1970, but cannot generate its decline afterwards. Despite decreasing the relative price of manufacturing

[32] Remember that we allow $\omega_{jt}$ to change over time in order to match the aggregate evolution of expenditure across sectors in Step 1.

[33] We proceed as follows. First, we solve for a counterfactual economy in which all parameters stay constant at their 1940 values, such that model dynamics only come from the reallocation of population in the transition towards the steady state (Column 3). Then, we allow for productivity, trade costs, and migration costs to vary over time one by one, and report how allocations change with respect to the economy in which parameters remain constant (Columns 4-6). We also solve for an economy that keeps productivity constant but allows for changes in both trade costs and migration costs, that is, changes in spatial frictions together (Column 7). Finally, the difference between the benchmark economy and the addition of results in Columns (3) to (6) gives us the strength of the interactions, see Column (8).

goods, which pushes employment away from manufacturing, these changes in productivity also generate income growth, which pushes employment towards manufacturing. Interestingly, changes in productivity generate stronger industrial growth in laggard than in leading regions, with employment growth in manufacturing being 21% and 8% of the initial local employment for regions in the q1 and q4 groups of employment growth, compared to 8% and 43% in the benchmark economy. Consistently, in this economy there is no significant relationship between the initial agricultural share and local employment growth, with $\hat{\beta} = -0.13$ and $R^2 = 0$. While productivity growth generates a more even structural transformation across regions, it also increases regional inequality, as the standard deviation of real income increases more between 1940 and 1960 and falls less after 1960 than in the benchmark economy.

The effects of changes in trade costs are shown in Column (6). Trade costs move employment across sectors in the directions observed in the data, except for the manufacturing share after 1970. However, their effect is limited, consistent with their limited impact on relative prices and real income, see Figure F.1. The changes in inequality generated by changes in trade costs go in opposite direction to the data, as they reduce inequality until 1960 and do not increase it afterwards. Finally, Column (7) reports the outcomes of an economy in which only migration costs change. While changes in migration costs have a limited impact on aggregate structural change and income growth, they are the main driver of the diverging patterns of industrialization across regions. In fact, changes in migration costs generate a relationship between the initial agricultural share and local employment growth even more negative than in the data. They also generate a reduction in regional inequality. Importantly, none of the engines on its own, nor combined (see columns 8-10), is able to generate the fall in the manufacturing share of employment after 1940: interactions among engines are necessary. Indeed, interactions also generate a substantial reduction in real income per capita.

Our model economy features three engines of development: the change in the sector- and province-specific productivities, the change in the matrices of sector-specific bilateral trade costs, and the change in the matrix of bilateral migration costs. In addition, even if parameters remain constant for the whole period, the economy experiences some changes over time due to the fact that the initial distribution of employment across provinces is not the steady state one. In this Section we quantify the role of each of these forces in the development experience of Spain between 1940 and 2000, reporting the main outcomes in Table F.1.[34]

[34]We proceed as follows. First, we solve for a counterfactual economy in which all parameters stay constant at their 1940 values, such that model dynamics only come from the reallocation of population in the transition towards the steady state (Column 3). Then, we allow for productivity, trade costs, and migration costs to vary over time one by one, and report how allocations change with respect to the economy in which parameters remain constant (Columns 4-6). We also solve for an economy that keeps productivity constant but allows for changes in both trade costs and migration costs, that is, changes in spatial frictions together (Column 7). Finally, the difference between the benchmark economy and the addition of results in Columns (3) to (6) gives us the strength of the interactions,

Table F.1: Counterfactual exercises (changes between 1940 and 2000)

| | ENGINES OF DEVELOPMENT | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Data (1) | Bench (2) | ΔPop. (3) | $\Delta\omega_{jt}$ (4) | $\Delta T_{rjt}$ (5) | $\Delta\tau_{r\ell jt}$ (6) | $\Delta\mu_{r\ell t}$ (7) | $\Delta\tau_{r\ell jt}, \mu_{r\ell t}$ (8) | $\Delta T_{rjt}, \tau_{r\ell jt}$ (9) | $\Delta T_{rjt}, \mu_{r\ell t}$ (10) |
| Real income pc | 8.16 | 8.31 | 1.04 | 1.15 | 8.61 | 1.31 | 1.12 | 1.4 | 8.79 | 9.19 |
| Agr share | −0.45 | −0.46 | −0.01 | −0.18 | −0.4 | −0.13 | −0.03 | −0.13 | −0.4 | −0.4 |
| Man share: 40-70 | 0.14 | 0.14 | 0.0 | 0.04 | 0.09 | 0.05 | 0.01 | 0.06 | 0.08 | 0.09 |
| Man share: 70-00 | −0.08 | −0.07 | 0.0 | 0.01 | −0.02 | 0.02 | 0.01 | 0.02 | 0.03 | −0.02 |
| Ser share | 0.39 | 0.39 | 0.0 | 0.14 | 0.34 | 0.06 | 0.01 | 0.05 | 0.29 | 0.34 |
| Man emp q1: 40-70 | 0.07 | 0.08 | 0.04 | 0.04 | 0.21 | 0.08 | −0.05 | 0.02 | 0.07 | 0.13 |
| Man emp q4: 40-70 | 0.44 | 0.43 | 0.11 | 0.06 | 0.08 | 0.02 | 0.08 | 0.11 | 0.12 | 0.19 |
| $CV(L_r)$ | 0.65 | 0.63 | 0.28 | −0.0 | −0.06 | −0.04 | 0.42 | 0.37 | −0.08 | 0.36 |
| $\hat{\beta}$ | −2.07 | −2.05 | −0.35 | −0.33 | −0.13 | −0.29 | −2.29 | −2.23 | −0.08 | −2.07 |
| $R^2$ | 0.63 | 0.6 | 0.03 | 0.03 | 0.0 | 0.02 | 0.65 | 0.6 | 0.0 | 0.63 |
| Sd($\log C_r$): 40-60 | 0.02 | −0.02 | 0.0 | 0.01 | 0.05 | −0.04 | −0.0 | −0.04 | −0.02 | 0.05 |
| Sd($\log C_r$): 60-00 | −0.21 | −0.24 | 0.0 | −0.07 | −0.16 | 0.0 | −0.01 | −0.01 | −0.21 | −0.18 |

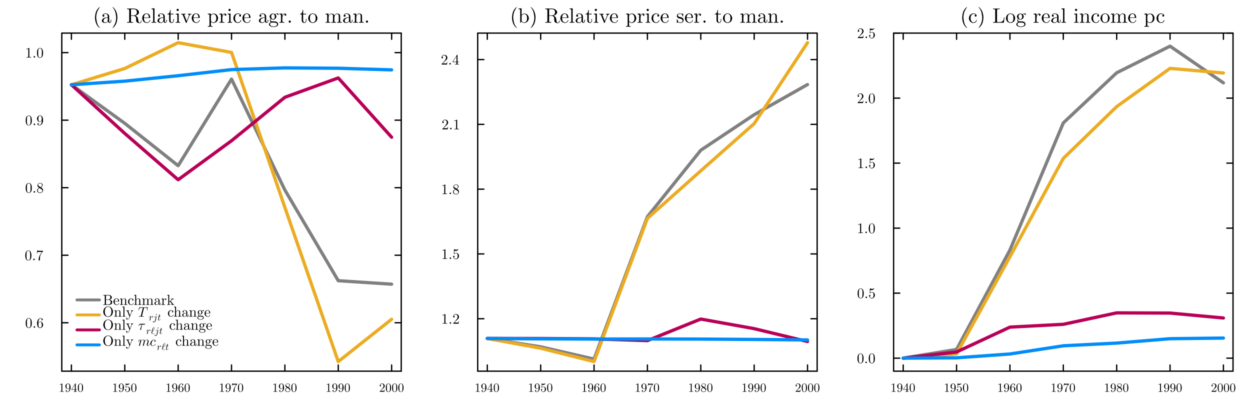
**Notes:** Each column reports level changes in the corresponding variable between 1940 and 2000, except other time frame indicated in the corresponding row. For real income per capita we report the ratio. "$\hat{\beta}$" and "$R^2$" are the coefficient and the share of explained variance of a regression of employment growth on initial agriculture share. "Man emp q1: 40-70" is employment change in manufacturing in q1 provinces relative to total employment of those provinces in 1940. Columns (1) and (2) are the data and the benchmark calibrated economy respectively. Column (3) is an economy in which all parameters stay put at their 1940 values. Columns (4)-(7) report the difference between economies in which only one set of parameters (CES demand shares, productivity, trade costs or migration costs) changes with respect to Column (3). Columns (8)-(10) reports the same differences for economies in which two engines change simultaneously.

Population dynamics due to the non-steady state distribution of population across space in 1940 play a minor role in the aggregate, despite generating some spatial reallocation. In particular, keeping all parameters constant, population moves from poorer to richer areas over time, increasing the dispersion of employment across provinces by 16 log points (compared to 40 in the benchmark economy) and with little sectoral bias, as the predictive effect of initial agricultural share on log employment changes is only -0.15 (-2.26 in the benchmark economy). This population movement has negligible effects in terms of aggregate output (a 3% increase overall) and structural change.

Productivity growth is the main engine of development in Spain between 1940 and 2000: it explains most of the growth in GDP (it increases output by a factor of 4.78, 5.3 in the benchmark economy) and most of the reallocation of economic activity across sectors (it produces a decline in the agriculture share of 39.4 percentage points, 45.3 in the benchmark economy, and a rise of services of 24.4 percentage points, 38.8 in the benchmark economy). Despite its asymmetries across regions, observed productivity growth contributes little to the overall reallocation of workers across space, as the standard deviation of employment across provinces only increases by 6 log points (compared to 40 in the benchmark economy). However, this predicted population movement goes in the same direction as in the data: the relationship between employment growth and the initial agricultural share is -1.03 (-2.26 in the calibrated model). Finally, it is worth noting that the

see Column (8).

Figure F.1: Forces of structural change



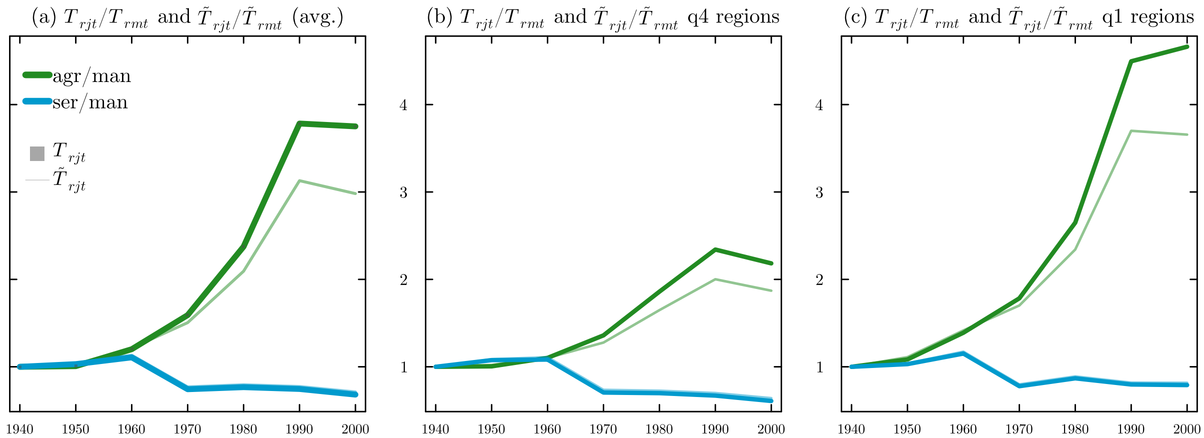

**Notes:** Relative price of agriculture with respect to manufacturing –Panel (a)–, of services with respect to manufacturing –Panel (b)–, and the real wage –Panel (c)– for our Benchmark economy and counterfactual economies in which only one engine of growth is allowed to change. Each line plots the population-weighted average across provinces.

divergence of productivity across regions in the period 1940-1960 and the convergence afterward is the main driver of the Kuznets curve of inequality. Indeed, the evolution of productivity generates a sharper inverted-U shape of income inequality over time than the one of the calibrated model: the standard deviation of productivity increases by 13.7 log points until 1960 (10 in the benchmark economy) and declines by 13 log points after that (11 in the benchmark economy).

The change in spatial frictions also plays a relevant role in the Spanish development episode. It adds sizable output growth (20 percentage points) and helps to reallocate employment across sectors in the same direction as observed in the data: it produces a decline in agriculture of 6.8 percentage points and a rise of services of 5.9 percentage points. These results are driven by the sectoral asymmetries in the fall of trade costs —which reduce the price of agriculture and manufacturing relative to services, see Panels (a) and (b) in Figure F.1— and from the increase in productivity (and hence income) generated by increased specialization, see Panel (c) in Figure F.1. The change in spatial frictions explains most of the rural exodus, with a slope of the regression of log employment change on the initial agricultural share of -1.56 (-2.26 in the benchmark) and an increase in the dispersion of employment of 11 log points (40 in the benchmark). These effects come mostly from the decline in migration costs towards the most prosperous regions between 1950 and 1980. Instead, the decline in trade costs has virtually no effect on the reallocation of workers across space. Finally, the change in spatial frictions partly offsets the increase in spatial income inequality generated by productivity growth between 1940 and 1960, as it induces a decline of 2.4 log points in the standard deviation of productivity until 1960.

To finish this decomposition, it is important to note that interactions between productivity growth and changes in spatial frictions are important for several outcomes. Among them, it stands out that the interactions are the sole driver of the de-industrialization of the country after the manufacturing peak in 1970, that is, the interaction between changes in productivity and spatial frictions generates a decline in the employment share of manufacturing of 12.6 percentage points, much larger than the 7.6 decline predicted by the calibrated model. To understand why, we note the following. First, if trade and migration costs vary over time but productivity stays constant, structural change is very limited as relative prices and income change little. As a consequence, the country is poorer, much more agrarian, and the manufacturing sector never thrives. Second, if productivity and trade costs change but migration costs remain at their 1940 values, industrial provinces are not able to attract enough workers to exploit their comparative advantage. This slows down industrialization, the manufacturing share peaks later in time and thus it does not show a hump before the year 2000. Lastly, if productivity and migration costs change but trade costs remain constant, the relative price of manufacturing with respect to services does not fall enough in the second half of the development process such that employment can shift from manufacturing to services (see Figure F.1), which prevents the manufacturing share from decreasing.

Figure F.2: Relative sectoral productivities




**Notes:** Panel (a) plots the average (across provinces) relative productivity of agriculture and services with respect to manufacturing. The thick line corresponds to the calibrated productivity parameters $T_{rjt}$ under the CRS assumption, while the thin line corresponds to the underlying productivity $\widetilde{T}_{rjt}$ when we allow for non-constant returns to scale. Panels (b) and (c) report the same averages for provinces within the q4 (highest population growth) and q1 (lowest population growth) groups respectively.

## F.2 Productivity paths with non-CRS production functions.

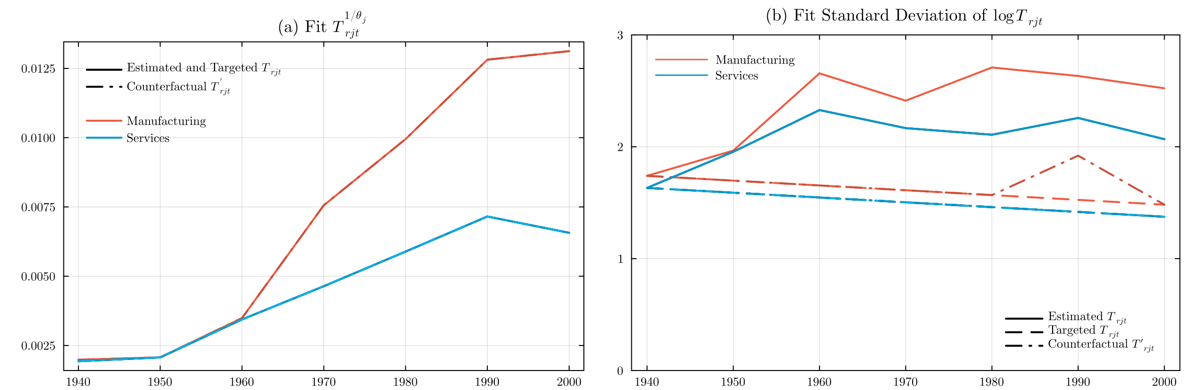
Here we compare the exogenous $\widetilde{T}_{rjt}$ with the calibrated productivity paths $T_{rjt}$. In Figure F.2, we plot the time paths of the relative sectoral productivity parameters for both cases, aggregated over all provinces (Panel a) and also aggregated over q4 and q1 provinces (Panels b and c), which are the 12 provinces with highest and lowest population growth respectively. We observe two patterns. First, the exogenous productivity of services relative to manufactures, $\widetilde{T}_{rst}/\widetilde{T}_{rmt}$, declines slightly more than the endogenous one, $T_{rst}/T_{rmt}$. This happens due to the increase in manufacturing employment, which raises $T_{rmt}$ over time. We also observe that this pattern is stronger in q4 than in q1 provinces due to the stronger industrialization in the provinces within the q4 group. Second, the endogenous productivity of agriculture relative to manufacturing, $T_{rat}/T_{rmt}$, grows more than the exogenous one, $\widetilde{T}_{rat}/\widetilde{T}_{rmt}$. This happens because the decline of employment in agriculture raises $T_{rat}$ more than the increase in employment in manufacturing raises $T_{rmt}$.[35] The pattern here is stronger for q1 provinces, which are the ones with a stronger decline in agriculture employment.

## F.3 Convergence of sectoral productivities.

Sectoral productivity diverged across provinces until 1960 and converged afterwards. A natural question is whether the patterns of migration could have been reverted had regional productivities started to converge since 1940. To explore this scenario, we want to generate alternative productiv-

[35] This result comes from the facts that (a) $|\alpha_a| > |\alpha_m|$ and (b) employment loss in agriculture is larger than the employment growth in manufacturing.

FIGURE F.3: Alternative Productivity Paths



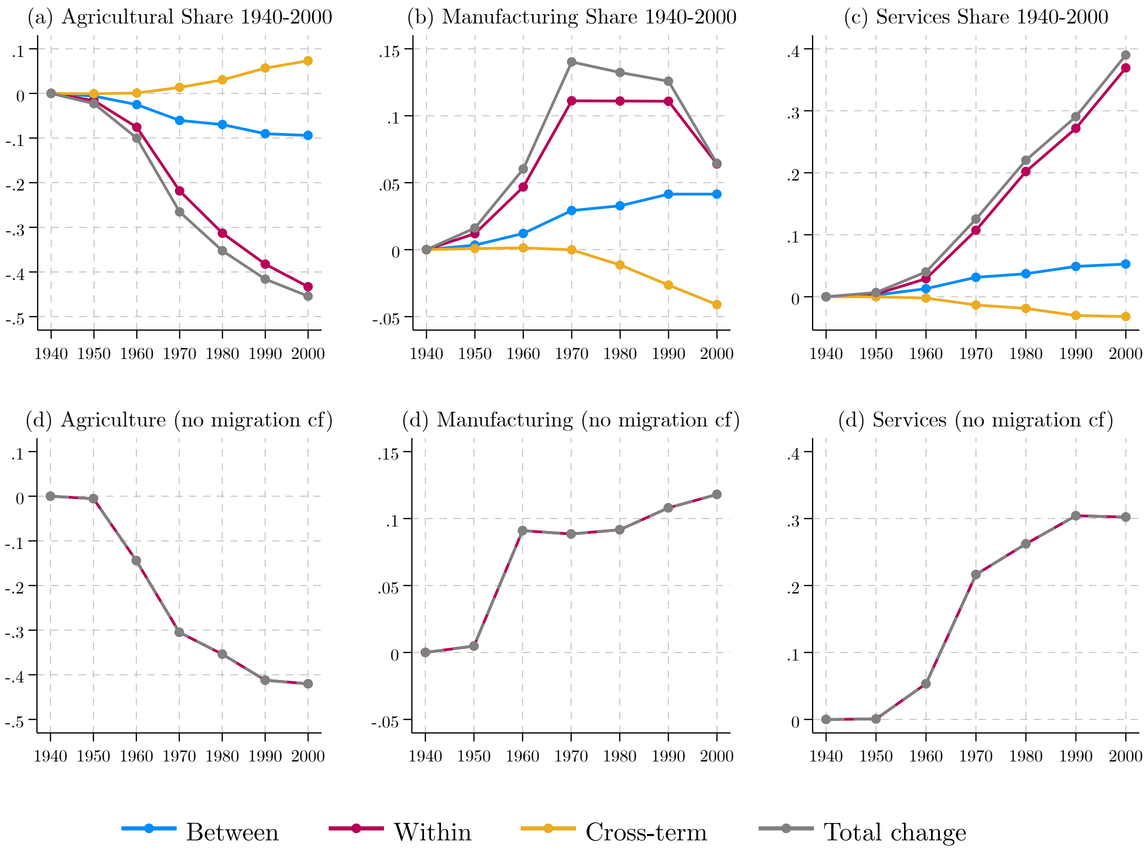

**Notes:** Left panel shows the standard deviation of $\log T_{rjt}$ in the calibrated economy (solid lines), the targeted standard dispersion of $\log T_{rjt}$ (dashed lines), and the dispersion of $\log T'_{rjt}$ (dotted line) obtained using equation (F.1) for manufacturing and services. Right panel shows the level of mean $\log T_{rjt}$ (solid lines) and the level of mean $\log T'_{rjt}$ (dotted line) for manufacturing and services.

ity paths $T'_{rjt}$ in each region $r$ and for each sector $j$ that preserve the evolution of average sectoral productivity across provinces as in the calibrated economy, but with a different evolution of dispersion across provinces. Of course, there are multiple ways of doing so. We proceed as follows. First, we parameterize the transformation from the calibrated $T_{rjt}$ to the counterfactual $T'_{rjt}$ with this simple function:

$$T'_{rjt+1} = T_{rjt+1}\,\lambda_{1jt}\left(\frac{T_{rjt}}{\bar{T}_{jt}}\right)^{\lambda_{2jt}} \tag{F.1}$$

where $\bar{T}_{jt} = \sum_{r=1}^{R} T_{rjt}/R$. With this formulation, $\lambda_{1jt} > 0$ controls the average of $T'_{rjt+1}$ and $\lambda_{2jt}$ the dispersion. With $\lambda_{1jt} = 1$ and $\lambda_{2jt} = 0$ we recover the original productivity paths $T'_{rjt+1} = T_{rjt+1}$. With $\lambda_{2jt} < 0$, $T'_{rjt+1}$ is relatively lower (higher) for high (low) productivity regions, which generates convergence across provinces and lowers dispersion. With $\lambda_{2jt} > 0$ the opposite is true. We then choose the time sequence of $\lambda_{1jt}$ and $\lambda_{2jt}$ for each sector $j$ to match the desired counterfactual evolution of the average and the dispersion of sectoral productivity paths across provinces. In particular, for sectors $j = m, s$, we target (i) the actual time evolution of the average of sector $j$ productivities across regions (solid lines in Figure F.3 Panel (b)); and (ii) the counterfactual time evolution of the dispersion of sector $j$ productivity across regions. This arises from starting with the actual dispersion in sector $j$ productivities in 1940 and letting it decline at the same average rate as the dispersion of services productivity $T_{rst}$ between 1970 and 2000 (dashed lines in Figure F.3 Panel (a)). The choice of $\lambda_{1jt}$ and $\lambda_{2jt}$ does an almost perfect match of the targets (see dotted lines in Figure F.3).

Figure G.1: Decomposition of sectoral employment shares




**Notes:** Panels (a), (b), and (c) plot the change in sectoral employment for agriculture, manufacturing, and services, plus the "between", "within", and "cross-term" corresponding to the three components in the right-hand side of equation (G.1). Panels (d), (e), and (f) plot the change in sectoral employment for the same sectors in the counterfactual economy without migrations.

## Appendix G: A simple decomposition of the time evolution of sectoral shares

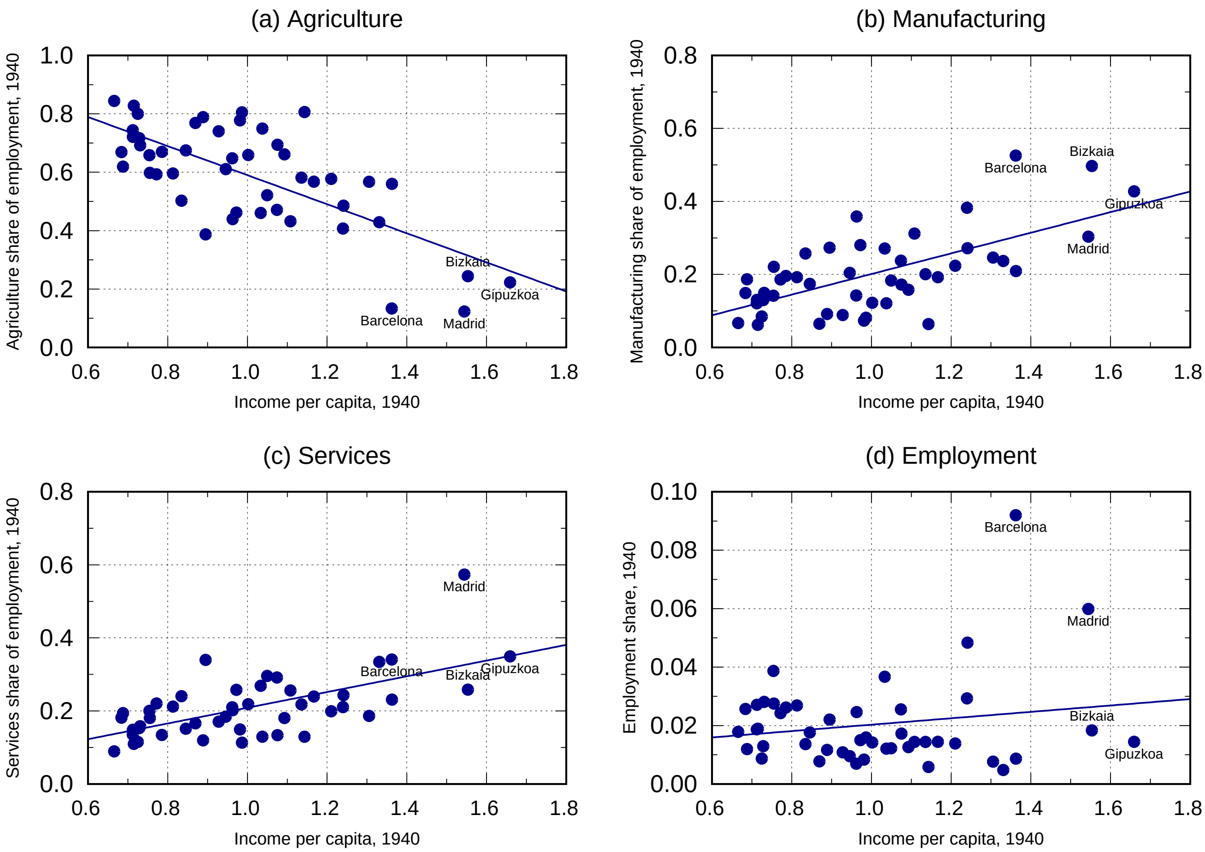
The time evolution of the employment share in a given sector, say agriculture, can be easily decomposed into a between-region term, a within-region term, and a cross-term:

$$l_{jt} - l_{j0} = \underbrace{\sum_{r=1}^{R} (l_{rt} - l_{r0})\, l_{rj0}}_{\text{between-region}} + \underbrace{\sum_{r=1}^{R} (l_{rjt} - l_{rj0})\, l_{r0}}_{\text{within-region}} + \underbrace{\sum_{r=1}^{R} (l_{rjt} - l_{rj0})\,(l_{rt} - l_{r0})}_{\text{cross-term}} \tag{G.1}$$

where $l_{rjt} = L_{rjt}/L_{rt}$ and $l_{rt} = L_{rt}/L_t$, see for instance Eckert and Peters (2022). The between-region term captures the change in the aggregate share of employment in a sector $j$ that comes from movements of population across regions of different sectoral composition. For instance, a rural exodus moving people from agrarian regions to industrial regions will mechanically generate an increase in aggregate industrial employment. The within-region term captures the change in the aggregate share of employment in sector $j$ that comes from changes in the sectoral composition of employment within each region $r$. It has been conventionally assumed that only the former term is related to migrations, and that a large within-region component dismisses the importance of migration for structural change. Yet, as we argue in this paper, the within-region term may also contain variation driven by migrations. For instance, the within-region term accounts for 20 percentage points in the fall of the share of employment in agriculture between 1940 and 1970, see Panel (a) in Figure G.1. Yet, in the counterfactual economy with no migration the share of employment in agriculture falls by 30 percentage points during the same period, see Panel (d) in Figure G.1. That is, the within-region component term of agriculture differs from the actual evolution of the share of agriculture in an economy without migrations by as much as one third.

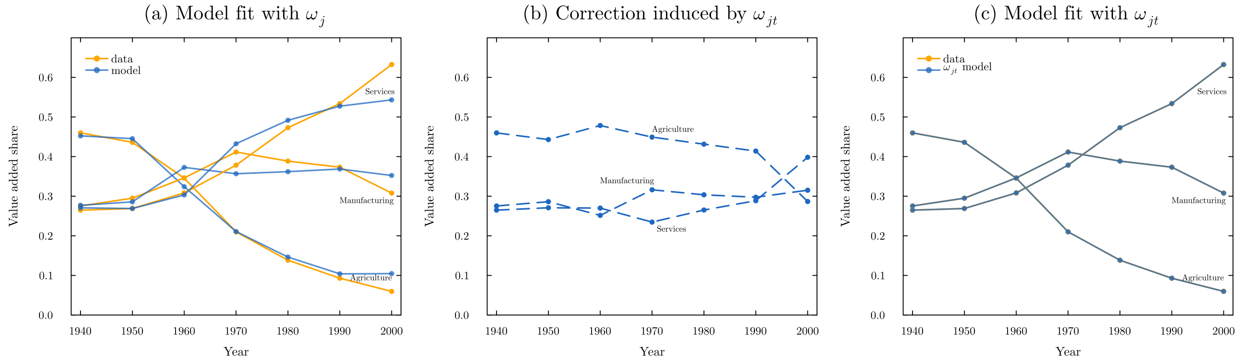
## Appendix H: Extra Figures and Tables

Figure H.1: Sectoral shares, 1940




**Notes:** Panels (a) to (c) plot the employment shares in Agriculture, Manufacturing, and Services against real income per capita in each province (relative to the country average). Panel (d) reports the relative size of each province (in terms of employment) against provincial income per capita. All panels for 1940.

## Figure H.2: Aggregate value added shares



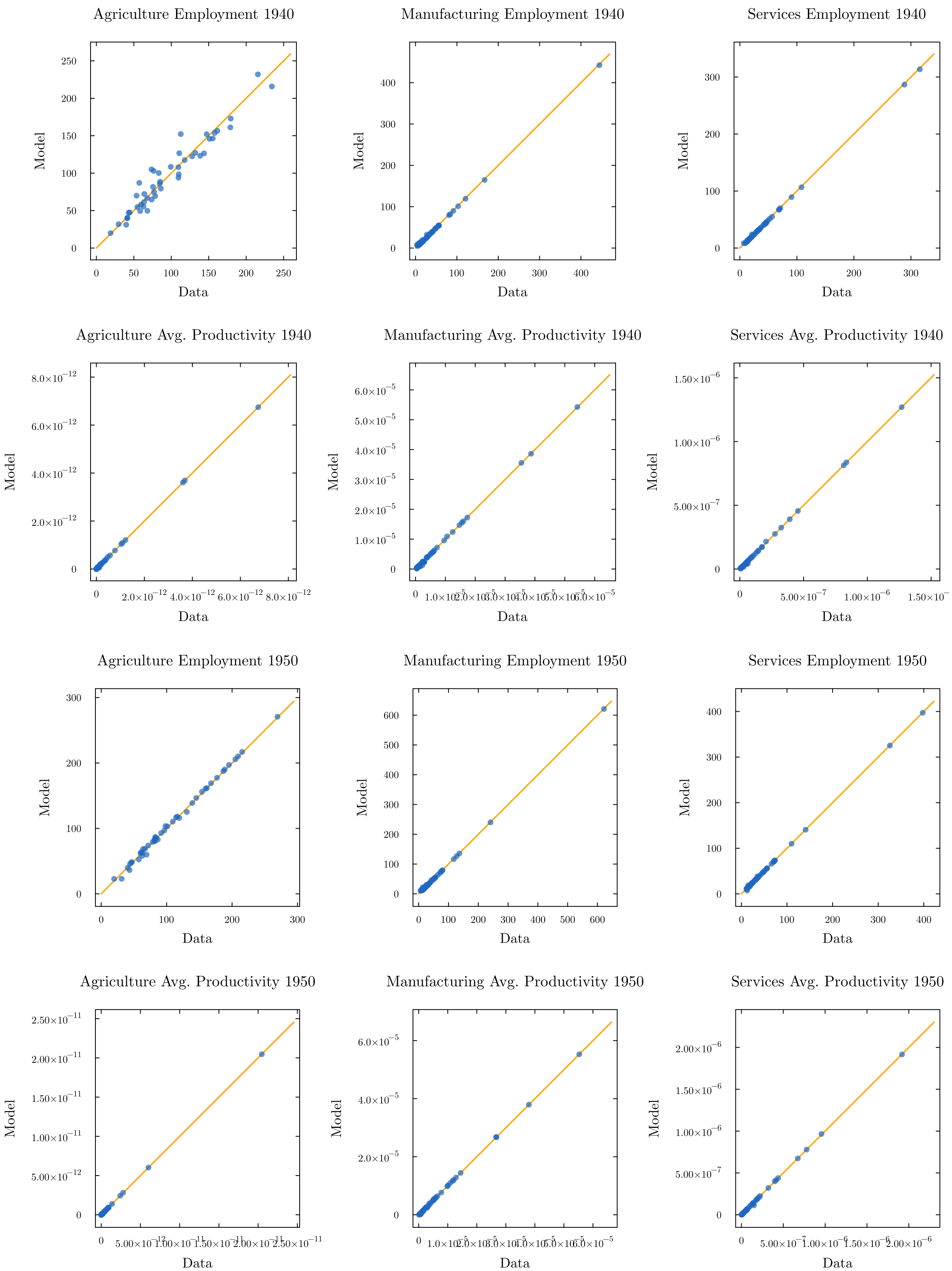

**Notes:** panel (a) plots the sectoral value added share in the data (yellow) and in the model (blue) using equation (17) with constant $\omega_j$ across time. Panel (b) plots the evolution of sectoral value added shares induced by the changes over time in $\omega_{jt}$, keeping prices and wages at their 1940 levels. Panel (c) plots the sectoral value added share in the data (yellow) and in the model (blue) using equation (17) with time-varying $\omega_{jt}$.

## Table H.1: Time-varying $\omega_{jt}$

| | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 | 2000 |
|---|---|---|---|---|---|---|---|
| $\omega_a$ | 0.024 | 0.023 | 0.026 | 0.022 | 0.022 | 0.021 | 0.014 |
| $\omega_m$ | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| $\omega_s$ | 0.376 | 0.372 | 0.415 | 0.295 | 0.345 | 0.383 | 0.501 |

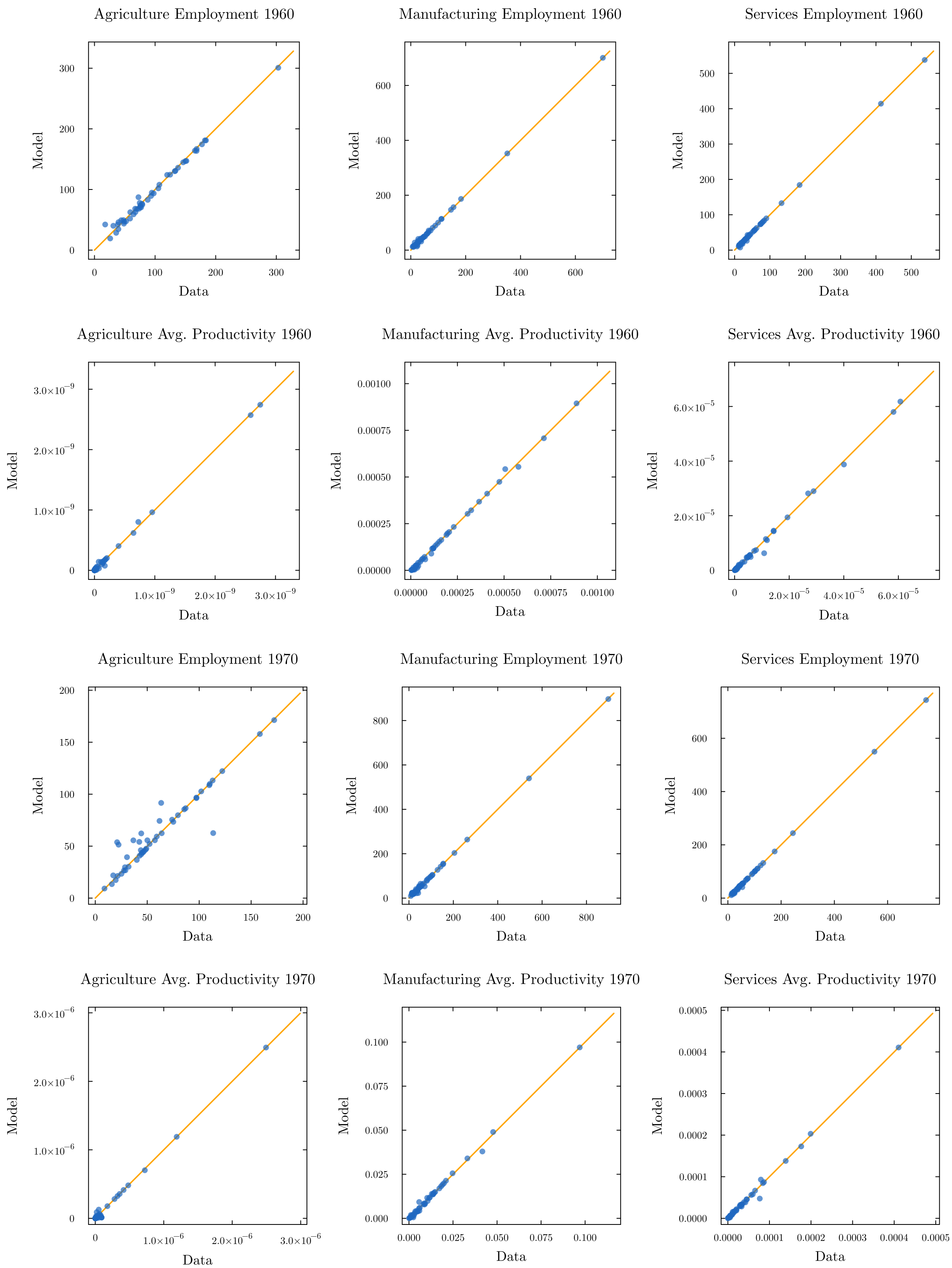
**Notes:** values for $\omega_{jt}$ obtained to perfectly match the data on aggregate sectoral value added shares over time.

## Figure H.3: Model fit employment and productivity: 1940 and 1950



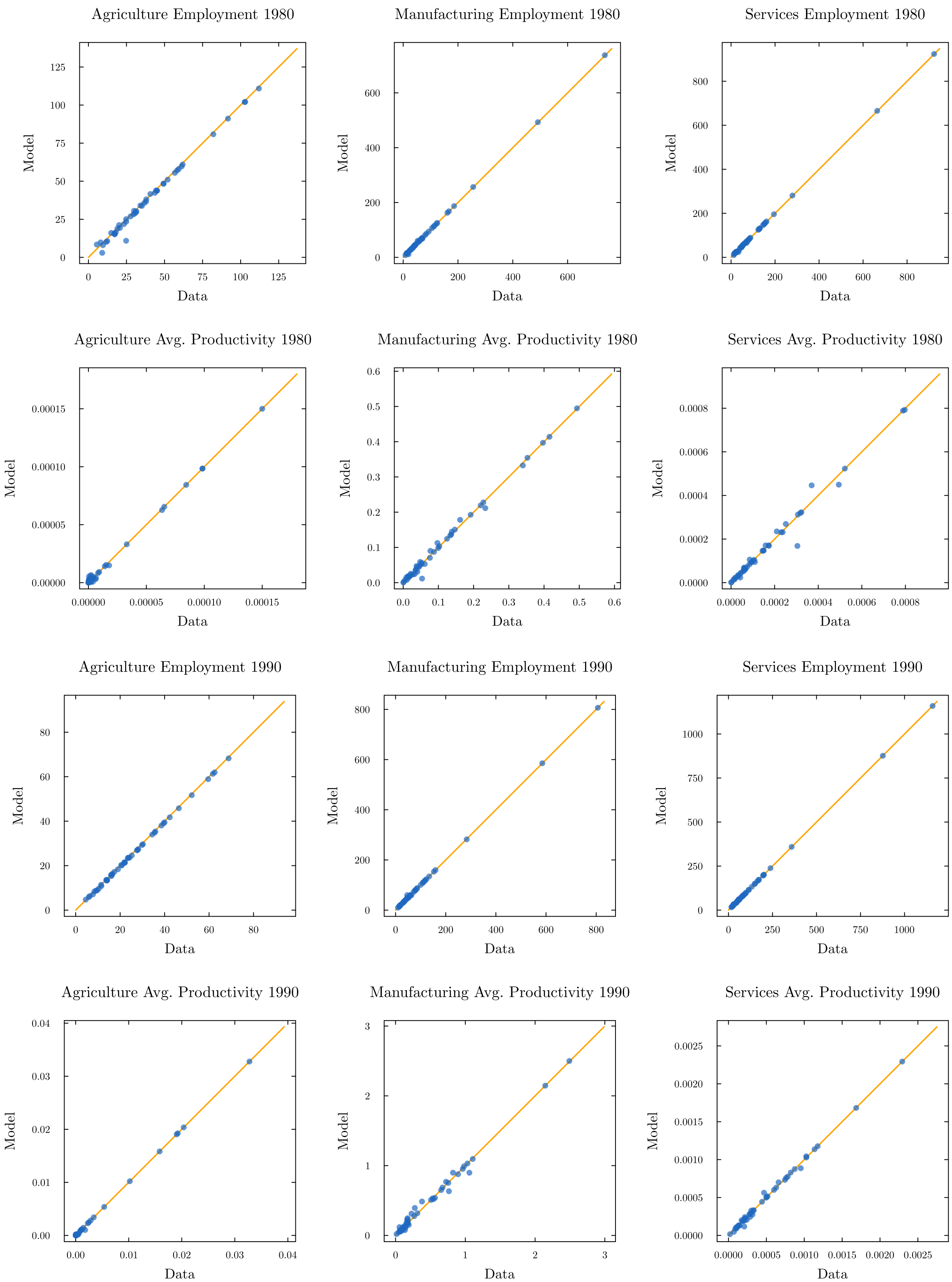

**Notes:** This figure confronts the model-predicted values for employment and producticity in each sector for the 1940 and 1950 decades. Each dot represents a province. Employment moments are measured in thousands of workers; value added moments are measured in billions of pesetas of 1990.

Figure H.4: Model fit employment and productivity: 1960 and 1970



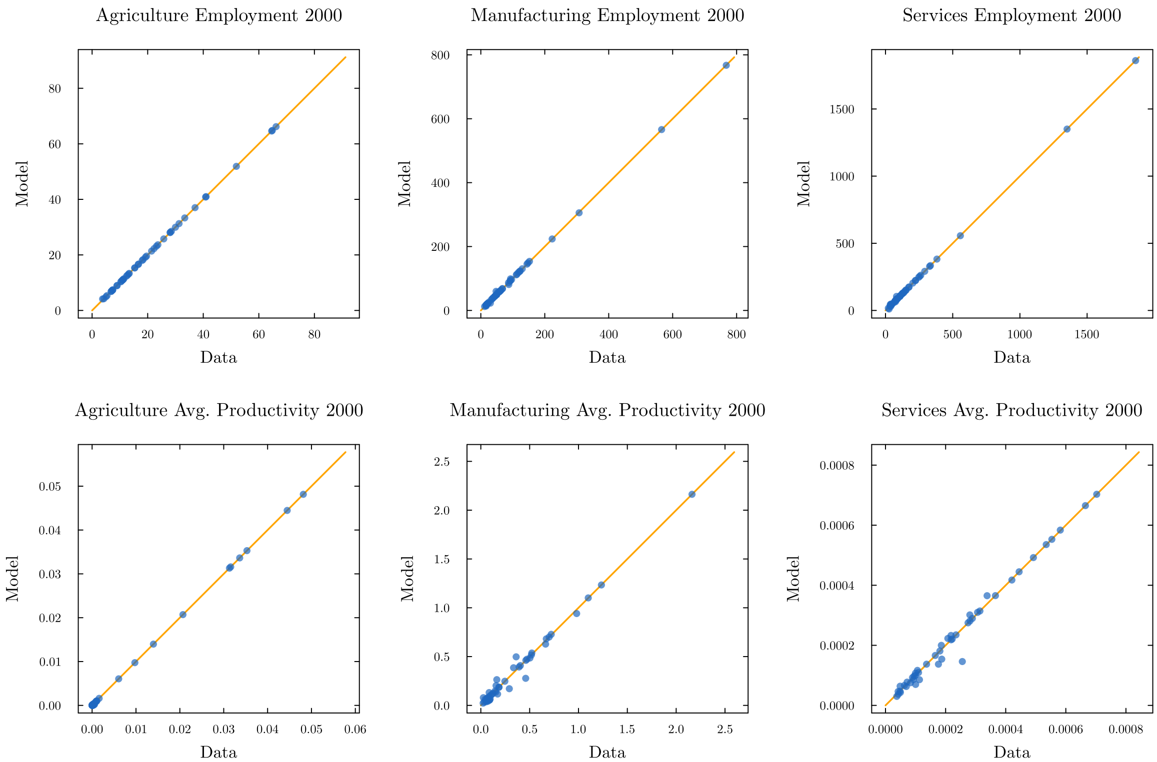

**Notes:** This figure confronts the model-predicted values for employment and producticity in each sector for the 1960 and 1970 decades. Each dot represents a province. Employment moments are measured in thousands of workers; value added moments are measured in billions of pesetas of 1990.

FIGURE H.5: Model fit employment and productivity: 1980 and 1990



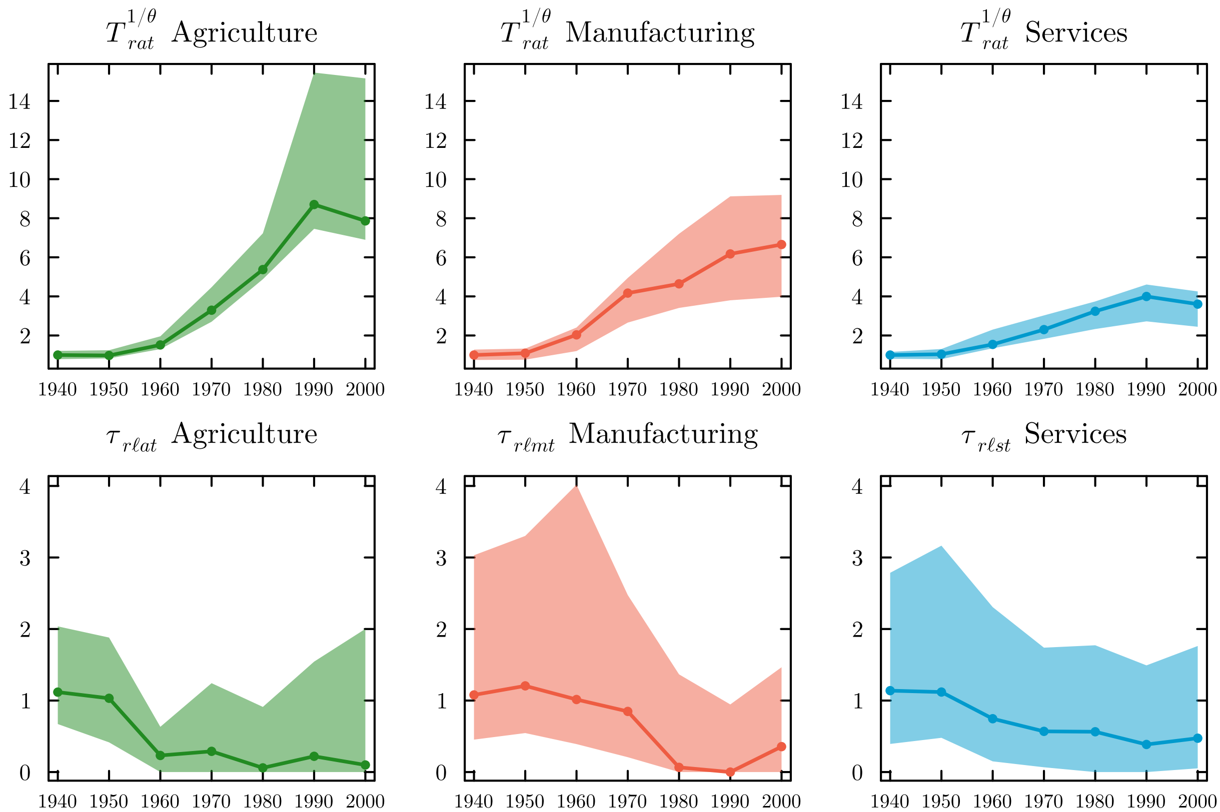

**Notes:** This figure confronts the model-predicted values for employment and producticity in each sector for the 1980 and 1990 decades. Each dot represents a province. Employment moments are measured in thousands of workers; value added moments are measured in billions of pesetas of 1990.

FIGURE H.6: Model fit employment and productivity: 2000




**Notes:** This figure confronts the model-predicted values for employment and producticity in each sector for the 2000 decade. Each dot represents a province. Employment moments are measured in thousands of workers; value added moments are measured in billions of pesetas of 1990.

Figure H.7: Evolution of productivity and trade costs over time



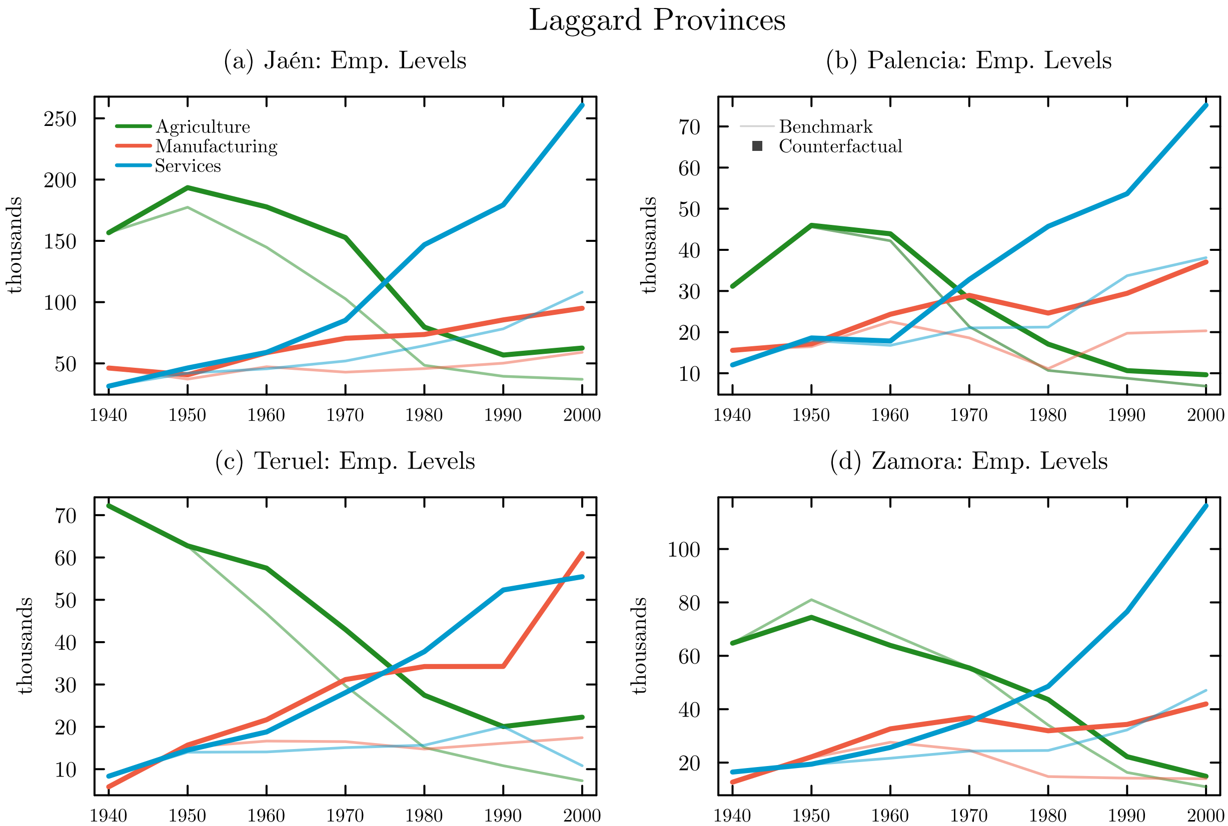

**Notes:** The thick dotted lines in each panel represent the evolution of the average across all provinces of the productivity (top row) and trade cost (bottom row) parameters. The bands represent the 25th to 75th percentiles of the distribution of the same parameters.

Table H.2: Correlation of Employment and Expenditure Shares over Time

| | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 | 2000 |
|---|---|---|---|---|---|---|---|
| Agriculture | 0.63 | 0.64 | 0.59 | 0.29 | 0.15 | -0.10 | -0.14 |
| Manufacturing | 0.56 | 0.61 | 0.58 | 0.31 | -0.32 | -0.61 | -0.50 |
| Services | 0.55 | 0.52 | 0.44 | 0.21 | 0.08 | 0.15 | 0.13 |

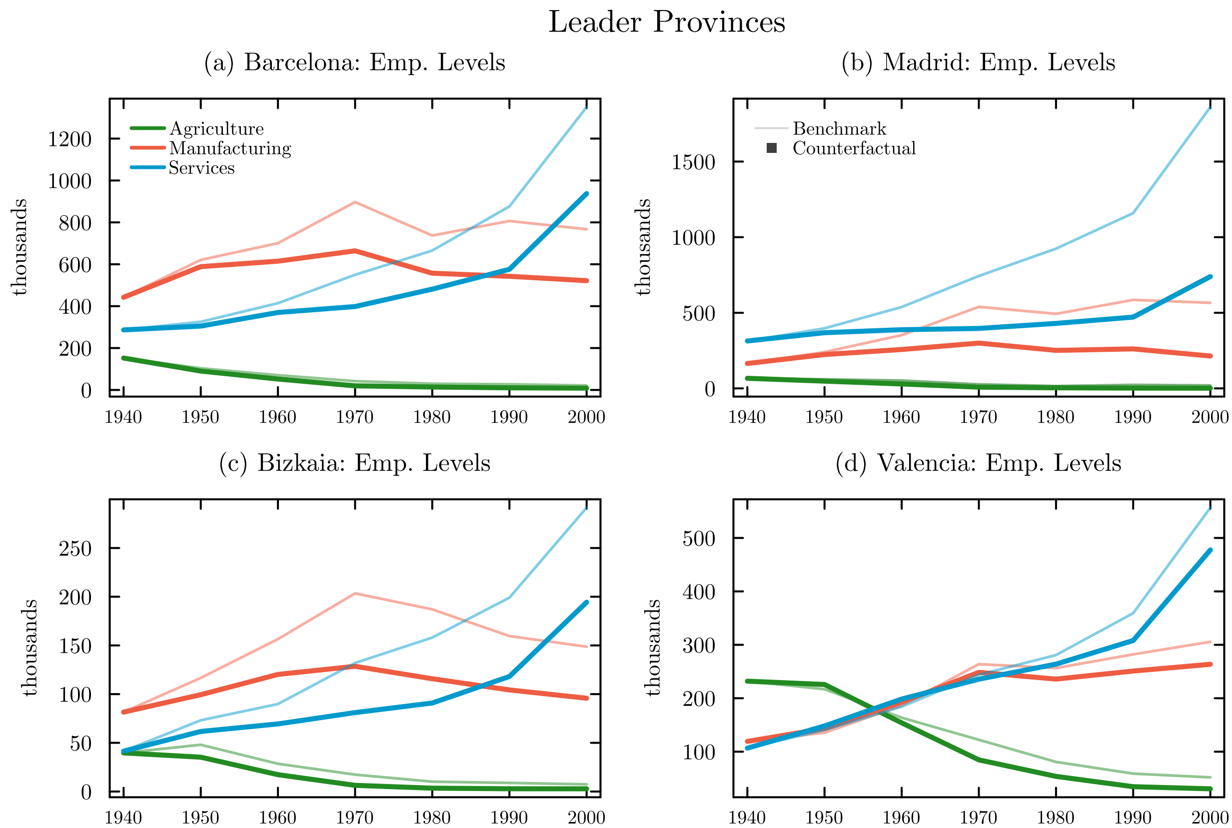
Notes: Each entry is the correlation of sectoral employment end expenditure shares across provinces for every year and sector.

Figure H.8: Counterfactual exercises: no migrations



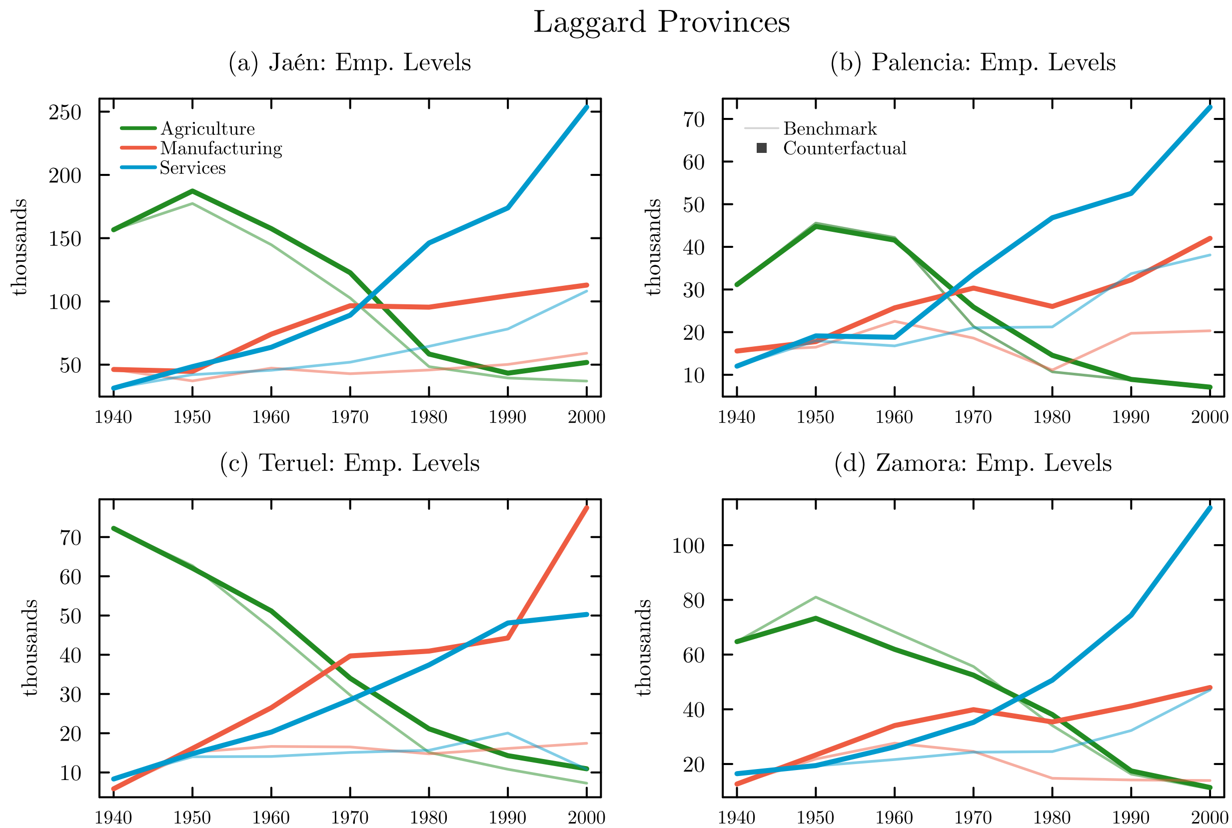

**Notes:** This Figure represents the time series evolution of employment levels in each sector for a group of selected provinces that lost population between 1940 and 2000. The thin lines correspond to the calibrated economy and the thick lines correspond to the counterfactual economy with no migration.

Figure H.9: Counterfactual exercises: no migrations



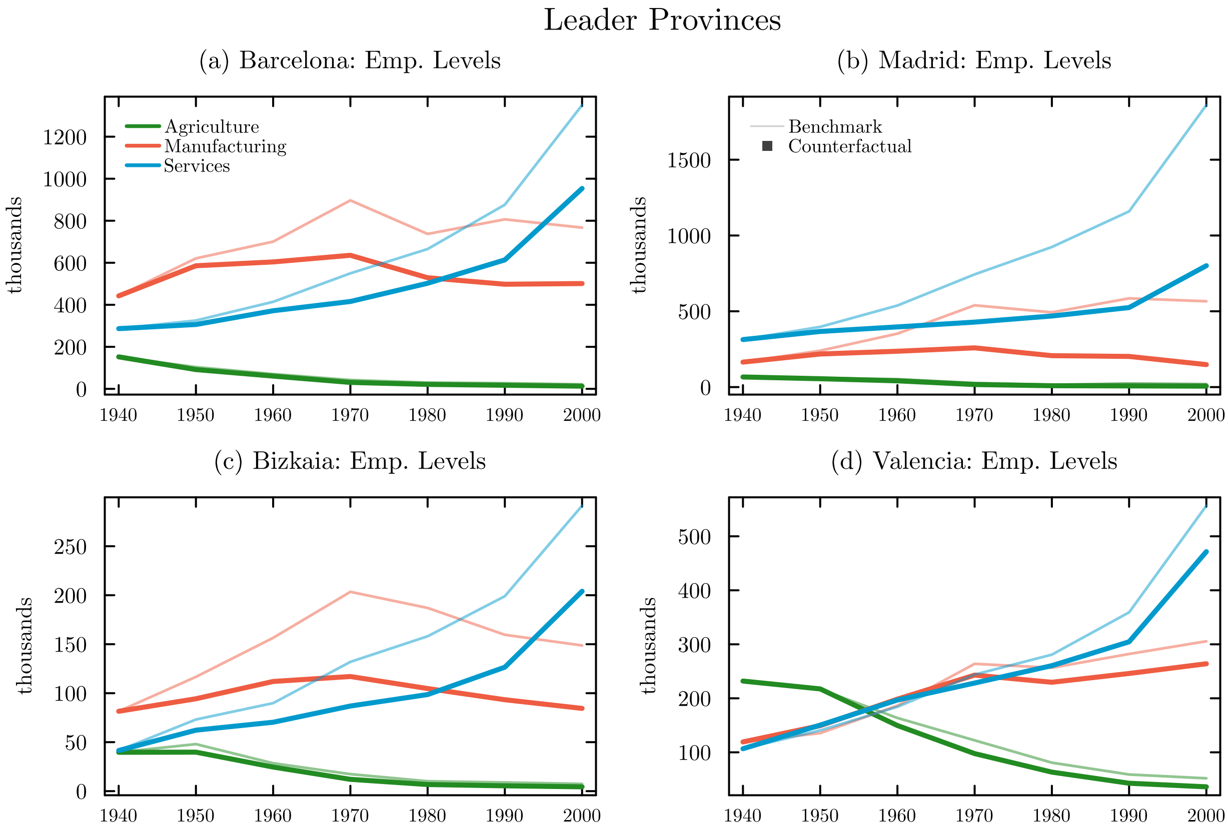

**Notes:** This Figure represents the time series evolution of employment levels in each sector for a group of selected provinces that gained population between 1940 and 2000. The thin lines correspond to the calibrated economy and the thick lines correspond to the counterfactual economy with no migration.

Figure H.10: Counterfactual exercises with agglomeration economies: no migrations




**Notes:** This Figure represents the time series evolution of employment levels in each sector for a group of selected provinces that lost population between 1940 and 2000. The thin lines correspond to the calibrated economy and the thick lines correspond to the counterfactual economy with no migration and agglomeration economies.

Figure H.11: Counterfactual exercises with agglomeration economies: no migrations




**Notes:** This Figure represents the time series evolution of employment levels in each sector for a group of selected provinces that gained population between 1940 and 2000. The thin lines correspond to the calibrated economy and the thick lines correspond to the counterfactual economy with no migration and agglomeration economies.