Routledge
Taylor & Francis Group

Historical Methods: A Journal of Quantitative and Interdisciplinary History

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/vhim20

The regional occupational structure in interwar England and Wales

Robin C. M. Philips, Matteo Calabrese, Robert Keenan & Bas van Leeuwen

To cite this article: Robin C. M. Philips, Matteo Calabrese, Robert Keenan & Bas van Leeuwen (2022): The regional occupational structure in interwar England and Wales, Historical Methods: A Journal of Quantitative and Interdisciplinary History, DOI: 10.1080/01615440.2022.2027303

To link to this article: https://doi.org/10.1080/01615440.2022.2027303

Published online: 28 Jan 2022.

HISTORICAL METHODS: A JOURNAL OF QUANTITATIVE AND INTERDISCIPLINARY HISTORY
https://doi.org/10.1080/01615440.2022.2027303

The regional occupational structure in interwar England and Wales

Robin C. M. Philips[a,b], Matteo Calabrese[c], Robert Keenan[d], and Bas van Leeuwen[b]

[a]Department of History and Art History, Utrecht University, Utrecht, Netherlands; [b]Research Department, International Institute of Social History, Amsterdam, Netherlands; [c]Luxembourg Centre for Contemporary and Digital History, Université du Luxembourg, Esch-sur-Alzette, Luxembourg; [d]Wageningen University & Research, Wageningen, Netherlands

ABSTRACT

A lack of regional data on the occupational structure in England and Wales during the interwar years has so far prevented extensive study of this time period. In the current paper, we fill this gap by reconstructing the occupational structure at the district level, based on a recently-digitized register for 1939 and by linking this dataset with the population censuses of 1911 and 1921. The resulting data reveals significant regional differences in the expansion of the tertiary sector, and the relative decline of agricultural and industrial activities. For industry, we find an increase in the level of geographical concentration during 1911–1921, to decline by 1939. The primary sector followed a similar pattern, whereas activities in the tertiary sector became less concentrated.

KEYWORDS

England; Wales; occupational structure; census; interwar

Introduction

As it was the first country to industrialize, the occupational structure of England has been studied extensively (e.g., Crafts and Mulatu 2005; Shaw-Taylor and Wrigley 2006; Shaw-Taylor 2009). However, the period covered in these studies often stops prior to World War I, more specifically at the 1911 census, largely due to the lack of data for the interwar period at a spatially disaggregated level. The 1931 population census was destroyed during the Second World War, and the conflict also prevented a census being taken in 1941. Hence, a data gap exists between the population censuses of 1921 and 1951, which has affected our understanding of the English labor structure during this period.

In an effort to fill this data gap, in the current paper we use an often-overlooked source: the 1939 National Register. This source allows us to reconstruct the labor structure in England and Wales at a district level for 1939. In addition, we standardize the published tables of the 1911 and 1921 population censuses, to make them suitable for a comparison with the estimates of the 1939 National Register, resulting in a complete dataset for the interwar period. The procedure is detailed in Section 2, where we expand on the issues and problems that this source presents, in particular the difficulties arising from classifying the occupations listed in the 1939 census. For this, we use the primary, secondary, tertiary (PST) classification system (Wrigley 2010), a classification system often applied in historical research to code occupations, and in particular used for the reconstruction of the occupational structure of England during 1379–1911[1]. In sections 3 and 4, we analyze the occupational structure of England and Wales and the geographical concentration and specialization of industries at the region, county, and district level for the interwar period (1911–1939). In line with Crafts and Mulatu (2005), we find evidence for a bell-shaped curve in industrial specialization during the period: with increasing specialization from 1911 to 1921, followed by a decrease from 1921 to 1939. We also find the same type of curve for geographical concentration at the region level, with the primary, secondary and tertiary sector experiencing an increase in concentration in 1921 followed by a decrease in 1939.

The sub-sectors with the highest level of industrial concentration at all the geographical levels are those related to specific geographical features (e.g., mining and agriculture) or highly benefitting from economies of scale and scope (e.g., the manufacturing of textiles and metals). The sub-sectors that instead consistently show the lowest levels of industrial concentration are those whose activities are traditionally more uniformly

CONTACT Robin C. M. Philips robin.philips@icloud.com Department of History and Art History,Utrecht University, Utrecht, Netherlands.

distributed, such as trades in the tertiary sector and the construction sector. However, our estimates also show that not all counties and districts followed the same patterns. For example, the Great Depression and the tariff wars during the interbellum resulted in a fall in international trade and – at best – ambiguous effects on the domestic market, which disproportionally hit the textile and metal industrial sectors. As a result, most notably the Lancashire cluster suffered. In addition, land-locked, centrally-located counties of the West Midlands and in the East, as well as the region of London, experienced the highest levels of national convergence due to the rapid expansion of vehicular transport. Instead, land-locked districts in Wales, the Northwest and Northeast England continued to specialize in agriculture as a response to their peripheral position. In Section 5, we put these conclusions in perspective, to position our results in the literature to date and outline some potential avenues for further research.

Sources and methods

The 1939 National Register

Due to the destruction of the 1931 census and the cancelation of the one planned for 1941, a gap in spatially disaggregated data exists for England during the interwar period, which can only be filled by using the 1939 register. The digitized 1939 National Register presents information collected from approximately 42 million people living in England and Wales. The original records are stored in the National Archives (Southport, Lancashire), whereas the records for Scotland, Northern Ireland, the Isle of Man and the Channel Islands have been excluded, as they remain un-digitized to date.[2] The 1939 National Register was modeled on the 1915 National Registration Act, which sought to accurately assess the manpower available to the British state during the First World War. Similarly, the purpose of the 1939 National Register was to gather relevant information about the population. As the British government became conscious of growing tensions in Europe, it wanted to efficiently plan the war effort.[3] The resulting registration was carried out by 65,000 enumerators, who ensured that on the night of the 29th of September, every household completed registration forms. These required the name, year of birth, gender, occupation, residence and marital status of each civilian, as well as information pertaining to whether or not they were a member of any reserve group for the armed forces. The collection of this data was subsequently used to issue identity cards to the civilian population, as it afterwards became a legal requirement to present these cards to the authorities upon request (until 1952).[4]

Although the National Register recorded similar information to that in the previously conducted censuses of 1801–1921, it cannot be formally described as a census. As such, its digitization and publication are not restricted by British privacy legislation, which limits access to censuses for the duration of 100 years[5]. Thus, in 2014–2015, the online genealogy service *Findmypast* was able to conserve, scan and digitize as many as 7,000 volumes and 1.2 million individual pages of the 1939 National Register. The *Findmypast* dataset of individual records was used for our reconstruction of the 1939 National Register[6]. For our reconstruction of the labor force over regions, we turned to three variables within the individual records: county, district and occupation title.

Classifying occupation titles in the 1939 National Register

Using the occupation titles in the National Register is, however, not without its problems. While easily defined categories such as year of birth, sex or marital status are reported in a fairly consistent fashion, occupational titles are reported with considerable variation. In different regions, workers in the same sector and with the same tasks are often reported with different job descriptions. For example: there are more than 2,400 uniquely stated occupational titles for general colliery miners, an occupation held by 6,003 people. Most of these occupational titles occur only once in the National Register, with occupation titles such as 'general laborer colliery' and 'laborer colliery heavy worker above ground'.

Given the not inconsiderable diversity of occupation titles in the 1939 National Register, we attributed the occupational titles in a systematic classification of occupations: the PST system, the latest update to which was devised by Wrigley (2010)[7]. This classification system is uniquely valuable, as it was used by the CAMPOP group to classify the British labor force for the 1379–1911 period, including the 1911 population census. Thus, by classifying the 1939 National Register in the PST system, we can not only easily link our estimates of 1939 with the 1911 population census, but also with regionally broken-down datasets extending back to the fourteenth century. In addition, its four-digit structure allows a relatively simple form of data insertion. This makes the possibility for errors due to the misattribution of occupational titles relatively small (Wrigley 2010, 7–8).

Notwithstanding its value, linking the 1939 occupation titles in the PST system did cause several issues to become apparent. First, we encountered many dual occupations in the 1939 National Register, such as 'Butcher & Farmer' and 'Motor Driver & Mechanic'. This is similar to nineteenth-century censuses, in which dual occupations are habitually reported as a single category (Wrigley 2010, 5). To resolve this inconsistency, we took the first stated occupation as the principal one. This should not significantly bias the resulting occupational structure, as many dual occupations with multiple instances are listed with a similar occurrence in both directions. For example, in addition to 'Butcher & Farmer', we find also titles such as 'Butcher and Farmer', 'Butcher & Dairy Farmer' and 'Assistant Butcher and Farmer' (for similar occupation titles, we find 468 different titles held by 1,344 people in the 1939 National Register). Likewise, we find a general occurrence of 'Farmer & Butcher', 'Farmer and Butcher', 'Dairy Farmer & Butcher', 'Farmer Butcher own account' and other similar titles (for which we find 338 different titles held by 1,000 people). Furthermore, as argued by Keibek (2017, III), by-employment has a higher occurrence among people employed in jobs related to agriculture. We should therefore expect a potential bias within this sector due to by-employment. Taking a look at the National Register, we find that 974,196 occupation titles contain the character '&', representative for 1,483,783 people in the entire dataset. Second, the stated occupational titles in the 1939 National Register are often too general to be attributed to a specific 4-digit PST sector[8]. For example, we categorized administrative occupations such as clerks and accountants in the PST sector of 'generic clerks' (PST code 5,31,0,40). Likewise, generic engineers are coded as 'engineers and other' (PST code 5,35,8,60), generic shop assistants as 'shop keepers' (PST code 4,0,0,3) and typists as 'other administrative officers' (PST code 5,31,2,60). Hands and apprentices with no specific job specification are classified as 'generic apprentices' (PST code 90,0,0,45). This resulted in the size of specific occupational groups at the 4-digit PST level (in which many people have a generic code) being underestimated.

Therefore, as a solution, we limited our reconstruction of the labor force in England and Wales to the PST 2-digit level, in which these allocations are leveled out. This level of aggregation holds consequences for our presented results. Through its 4-digit structure, the PST system allows a twofold approach to the analysis of these in-between cases: industry-driven and occupation-driven. By considering the first two digits only, our focus is on industry. The final two digits of the code, instead, shift the focus on the sole occupation. Consequently, by reporting our results on the 1-digit level and 2-digit level of the PST system, we restrict our reconstruction to a mainly industry-driven approach. This is also in part induced for reasons of comparison, as the separation of 'industry', i.e., what a worker is helping to produce, from 'occupation', that is, what a worker does, is not cleanly made in labor statistics until quite late in the 20th century. In particular, due to the population censuses of 1911 and 1921 being relatively highly-aggregated at the sectoral level, the 1-digit level and 2-digit level of the PST system was the highest possible level of disaggregation for our source (for a more detailed discussion, we refer to section 2.4.). But also, given the increase in service jobs within the primary and the secondary sectors – already taking place in the first part of the twentieth century in the US and in the UK (see, e.g., Goldin and Katz (2008)) –, this inhibits an important nuance for our analysis. Especially regarding the shift from secondary to tertiary occupations, as many tertiary occupations within a secondary industry will be classified as a part of the secondary sector, and not the tertiary sector. For example, consider an accountant working for a mining firm or a mining engineer. Considering the industry-driven approach in the first 2 digits of the PST system (see Wrigley: 19-20) and the occupation title which was often stated to great detail in the 1939 National Register (which makes it in most cases possible to link an aforementioned accountant to the mining sector), in our results, those occupations will hence be attributed to the mining sector.

Second, some of the listed occupations in the 1939 National Register were too general to be attributed to a PST 2-digit level sector. The occurrence of unspecified occupation titles such as 'laborer', 'worker' or 'general laborer' is a general problem with the coding of occupational titles. To account for this problem, the PST system even included an additional 'sector', titled 'sectorally unspecific occupations'. Due to the level of detail in the listed occupational titles in the 1939 National Register, these unspecified occupation titles represent a relatively small share of the listed occupations (see next section below). As a solution, we used the regional variation in the occupational structure of the district to attribute the unspecified laborers to a sector. In other words: we assumed that the occurrence of unspecified laborers was unbiased over sectors.

Third, as the National Register itself was designed to be enumerated exclusively on the 29th of September, two occupational groups are underrepresented. On the one hand: as only fishermen not at sea at the time were available to be surveyed, the National Register did not list fishermen on short and long-term voyages who were not docked in port at the time of enumeration. The national employment figure for fishermen in 1931, as documented by Mitchell (1988), is 40,000 employees. Yet, in our sample of the most occurring 13,359 occupations of the 1939 National Register, we count only 6,456 people in the fishing industry (for our sampling procedure for the 1939 Register, see the next section). For the most part, these occupations could be attributed to the primary sector, though with occupational titles such as 'seaman' (9 occupational titles held by 3,671 people) – an occupation that lends itself more to ambiguity – the attribution to a particular sector is sometimes more difficult to make. Nonetheless, and most importantly, an adjustment for the underestimated number of this group has to be made.

On the other hand, another group was also partially excluded from the National Register: armed forces personnel. Although civilians in military installations and military personnel on leave were surveyed, members of the armed forces personnel in barracks, naval stations and air force installations were not. Accordingly, our sample lists only 19,757 people occupied as army personnel, whereas this group (including civilians mobilized for the Army, Royal Air Force, Royal Marines and Royal Navy) at that time is assumed to have stood at 900,000 people (French 2000, 63). As the country was getting prepared for the impending war and mobilization was on the rise at the time that the National Register was conducted, as a result, the tertiary sub-sector "Services and professions" would strongly be undervalued without any adjustment.

Hence, we conducted an ex-post readjustment for both fishermen and armed forces personnel, based on the above-mentioned estimates in the secondary literature and under the assumption that both groups are proportionally distributed across the regions based on our sample of the National Register (see the next section). In other words: we take the geographical distribution of the fishermen and army personnel that was surveyed in the National Register and elevate the numbers for these groups to the national estimate in the secondary literature.

Sampling of the 1939 National Register

Given the large number of registered persons and occupational titles in the 1939 register, we used sampling to calculate the number of workers by sector in each district. To do so, we separated the 42,299,296 records in the National Register into two groups, based on the occurrence of the occupation titles.

The first group refers to 30,989,514 records contained in 13,359 unique occupational titles, each of which occurred a number of times greater or equal than 100: therefore, covering approximately 73.26 per cent of the total entries in the National Register. Looking at the 13,359 occupational titles, in a preliminary step, we removed the occupation titles of those who were not participating in the labor force or could not be attributed to an occupation title. For instance, the 1939 census includes 9.8 million records for civilians with 'unpaid domestic duties' and 6 million present 'at school', 3 million coded as 'baby' and 1.4 million listed with no occupation or classified as 'not recorded'. This resulted in 10,563 unique occupational titles, representing 8,511,486 people. By also including 681,870 people with 'sectoral unspecified occupations', 131,928 unemployed – those defined as 'unemployed' or as 'seeking work' – and after adjusting for the two groups partially excluded from the National Register (soldiers and fishermen), we count in our sample an active population of 10,040,071 people out of an estimated active population in 1939 of 19,750,000 (Broadberry and Howlett 2004), or approximately 50 per cent of the labor force. Those in the second group, containing the remaining 11,309,782 million records, were covered by 7,027,488 unique titles, encompassing not only occupations, but, just as in the first group, also titles describing people not active in the labor market.

As coding all titles into the PST classification system would have been an almost insurmountable task, we followed a two-stage sampling approach. First, we linked the 13,359 most frequent occupations to the PST classification system. However, it is likely that those frequently occurring occupations are non-randomly distributed among PST codes. To remove this bias, we used a second random sample gathered from the remaining 11.3 million occupation titles. To this end, we coded 20,000 additional occupational titles into the PST classification system, amounting to approximately 0.2 per cent of the 7,027,488 non-frequent unique titles. This second sample was subsequently used to correct the sample of 13,359 occupational titles. Based on our estimates (data available upon request), it appears that the 13,359

occupational titles resulted in *under* and *over* valuations of particular sectors, in which the highest deviations were found for specialized sectors such as the manufacturing of chemical products or metal working.

In order to assess the representativeness of our dual sampling method, we link all occupational titles into the PST classification system for the entire population in four districts: Penmaenmawr U.D. and Lleyn R.D. (Caernarvonshire), Winslow R.D. and Wolverton U.D. (Buckinghamshire). Comparing the results for these four districts of our dual-sampling method with a full sample of the entire population, we can get a view on the representativeness of our used sampling method. In Table 1, we compare our results at a subsectoral level for the primary, secondary, and tertiary sector. Given that the deviations are relatively small, with the sampling results not deviating from the full sample results by more than a percentage point in most cases (with an average error of 1.7 per cent), we argue that our estimates for the 1939 National Register using the dual approach are trustworthy.

Table 1. Assessment of the representativeness of our estimates of the 1939 National Register for four districts with complete data (%).

Penmaenmawr U.D. (Caernarvonshire)	two-stage sampling	Full data
Agriculture, estate work, forestry and fishing	7.35	5.34
Mining and quarrying	5.78	9.93
Food, beverages and tobacco	2.84	2.64
Textiles, wearing apparel and leather	2.84	2.77
Wood, furniture, paper products, etc.	1.47	1.19
Coke, chemical products, etc.	0.00	0.06
Basic metals, metal products, etc.	2.73	2.20
Other manufactured goods and repair	0.00	0.69
Construction and construction work	6.72	11.31
Trade	11.13	11.13
Services and Professions	51.37	42.30
Transport and Communication	7.77	10.43
Lleyn R.D. (Caernarvonshire)	**two-stage sampling**	**Full data**
Agriculture, estate work, forestry and fishing	46.15	42.97
Mining and quarrying	2.20	4.48
Food, beverages and tobacco	1.57	1.35
Textiles, wearing apparel and leather	1.89	1.76
Wood, furniture, paper products, etc.	1.67	1.16
Coke, chemical products, etc.	0.05	0.03
Basic metals, metal products, etc.	1.89	1.85
Other manufactured goods and repair	0.05	0.53
Construction and construction work	8.41	11.54
Trade	4.77	5.96
Services and Professions	25.23	21.94
Transport and Communication	6.16	6.42
Winslow R.D. (Buckinghamshire)	**two-stage sampling**	**Full data**
Agriculture, estate work, forestry and fishing	35.71	30.37
Mining and quarrying	0.17	0.22
Food, beverages and tobacco	2.89	2.84
Textiles, wearing apparel and leather	2.43	2.66
Wood, furniture, paper products, etc.	0.57	1.53
Coke, chemical products, etc.	0.17	0.18
Basic metals, metal products, etc.	2.66	3.13
Other manufactured goods and repair	0.11	0.29
Construction and construction work	14.77	16.44
Trade	4.81	5.46
Services and Professions	29.99	29.21
Transport and Communication	5.72	7.68
Wolverton U.D. (Buckinghamshire)	**two-stage sampling**	**Full data**
Agriculture, estate work, forestry and fishing	3.20	2.19
Mining and quarrying	0.13	0.38
Food, beverages and tobacco	2.89	1.83
Textiles, wearing apparel and leather	2.89	2.33
Wood, furniture, paper products, etc.	7.22	10.50
Coke, chemical products, etc.	0.66	0.95
Basic metals, metal products, etc.	23.20	26.52
Other manufactured goods and repair	0.13	0.27
Construction and construction work	8.67	14.51
Trade	8.01	6.42
Services and Professions	29.42	21.67
Transport and Communication	13.57	12.43

Adding the 1911 and 1921 censuses

In order to arrive at a complete set of estimates for the interwar period, we have to link our estimates of the labor force for 1939 to the estimates of the population censuses of 1911 and 1921. The 1911 census was retrieved from IPUMS International[9], based on individual level data, while the 1921 census was obtained from published census tables digitized by Visions of Britain, a resource for historical geographical and statistical data maintained by the Great Britain Historical GIS Project (2017) 'Great Britain Historical GIS' based in the University of Portsmouth[10].

First, to link the two censuses with occupational sectors, we matched the sectors in the 1911 and 1921 population censuses with the most-fitting PST sector (correspondence table available upon request). For 1911, Shaw-Taylor (2009) already linked the 1911 census with the PST system: we followed his methodology to link the 1911 occupational groups with the PST system at the 2-digit level. However, the data for the 1921 population census was only available at a highly-aggregated sectoral level with 31 different subsectors.

Considering the higher level of sectoral aggregation in the 1911 and 1921 censuses compared to the relative highly disaggregated sectoral level of the 1939 National Register on the one hand, and changes in the classification systems of the 1911 and 1921 censuses on the other hand, limitations arise as to an exact comparison between the 1911, 1921 and 1939 results, especially when turning to the more sectoral disaggregated results. A particular problem arises with a comparison between the 1911 and 1921 occupational coding, as the 1911 census followed a more industry-

driven classification system (with a breakdown in occupation-driven sub-categories) and the 1921 more an occupation-driven system. This produced some ambiguities for the attribution of certain aggregated job titles to specific sectors. For instance, public workers in the maintenance of the railway system (so, belonging to the sub-sector "Construction and construction work") and workers employed in factories for the production of railway components (in the sub-sector "Metal manufacture, machines, implements, vehicles, precious metals, etc.") were often put in the same category. Another sector where this becomes apparent is the sub-sector of the manufacture of "Food, beverages and tobacco". For this industry sector, the 1901 results present most likely an overvaluation, due to the inclusion of jobs such as bakers and butchers - which would be classified with the PST approach under retail trade, as tertiary jobs[11]. As such, it is no surprise that Mitchell's statistics reveal a sudden drop of ca. 5 percentage points for this sector between 1911 and 1921 – especially, considering that the conventions of the industrial categorization were changed between both these censuses (Mitchell 1988, 105).

Second, regarding the sectoral division, we can present our results on the sectoral level: for the primary sector (further subdivided in two sub-sectors, agriculture and mining[12]), the secondary sector, grouped in seven sub-sectors (food, beverages and tobacco; textiles, wearing apparel and leather; wood, furniture, paper products, etc.; coke, chemical products, etc.; metal manufacture, machines, implements, vehicles, precious metals, etc.; other manufactured goods and repair; construction and construction works) and the tertiary sector, grouped in three sub-sectors (trade; services and professions; transport and communication).

Occupational structure during the interwar period

National trends in the occupational structure

Table 2 presents the proportion of employment in the primary, secondary and tertiary sectors of the total labor force in England and Wales. At the national level, England and Wales changed from being a largely industry-based economy to a largely service-based one during the interwar period. Between 1901 and 1939, employment in the secondary sector as a proportion of the total labor force declined from 45.4 per cent to 30.4 per cent. This shift from industry to services occurred as a result of increased international competition for Britain's leading export manufacturers and increasing protectionism by many of its export markets, in particular on the European mainland (De Bromhead et al. 2017). Furthermore, with the impending war, mobilization drew laborers from agriculture and industry to the tertiary sector, as armed forces personnel are categorized in the tertiary sector in the PST classification system. At the same time, many unemployed workers previously occupied in the declining industries of the secondary sector (notably those formerly employed in the production of textiles) and the mining industry, reengaged in the labor market as daily wage-earners either in low-skilled tertiary jobs in the cities or as seasonal workers in agriculture (Caunce 2012). This does explain the (small) increase for agriculture during 1931–1939, despite the overall primary sector, including mining, had lost around 4 percentage points during 1911–1939 (see Table 2).

Over the period of 1901–1939, employment in the tertiary sector grew from 33.9 per cent to 56.8 per cent (see Table 2). This was mainly due to the growth in services and professions, and to a smaller extent the wholesale and retail sectors, during the interwar period (Eichengreen 1986). Trades constantly grew throughout the first interwar years, to decline due to the economic depression. But also because those employed in trades were the first responders to the call to arms that presented itself at the end of the interwar period (French 2000, 63–64), which in part explains the rise in services and professions during that same period. Overall, employment in transport remained stable during the interwar period, despite considerable technological change. During this time, motor vehicles began to replace rail as the preferred method of haulage: between 1920 and 1938 the number of commercial vehicles on British roads increased from approximately 100,000 to half a million (Thomas 2004), as railway operators failed to offer price competitiveness due to regulatory restrictions (Scott 2002).

For the primary sector, as with industry, the national proportion of employment also experienced a decline over the interwar period, despite an inversion in the declining trend for those employed in the agricultural sector, with a growth in the late thirties. Relative and absolute employment in mining constantly decreased throughout the interwar years, starting from a figure of 5.9 percent in 1911 – aligning with the estimations of Shaw-Taylor (2009, 20) – to eventually decrease to 3.7 percent in 1939[13].

Table 2 also shows that although employment in almost all industrial subsectors declined, this decline was not evenly spread across industry sectors. A

Table 2. Distribution of occupied population (male and female) in England and Wales by sector, 1901–1951, as a percentage of the labor force.

	1901 *Mitchell (1988)*	1911 *Thistext*	1921 *Thistext*	1931 *Mitchell (1988)*	1939 *Thistext*	1951 *Mitchell (1988)*
Primary	*14.7*	*16.1*	*14.3*	*11.7*	*12.3*	*8.5*
Agriculture, estate work, forestry and fishing	8.9	10.2	7.8	6.6	8.6	5.5
Mining and quarrying	5.7	5.9	6.5	5.1	3.7	2.9
Secondary	*45.4*	*38.4*	*38.0*	*28.7*	*30.4*	*29.8*
Food, beverages and tobacco	5.6	2.6	1.8	1.1	2.6	1.2
Textiles, wearing apparel and leather	16.4	13.7	11.6	9.2	7.7	6.3
Wood, furniture, paper products, etc.	3.8	4.9	5.1	3.7	3.2	3.4
Coke, chemical products, etc.	1.2	0.5	0.7	0.8	0.6	1.1
Basic metals, transport products, others, etc.	9.6	6.5	10.4	9.0	6.2	12.0
Construction and construction work	8.6	9.6	7.3	4.6	9.7	5.6
Other manufactured goods and repair	–	0.6	1.1	–	0.4	–
Tertiary	*33.9*	*44.8*	*47.2*	*49.4*	*56.8*	*51.5*
Trade	4.1	8.3	9.6	11.3	6.8	9.7
Services and professions	20.9	28.4	28.5	29.7	40.9	34.1
Transport and communication	8.8	8.1	9.1	8.6	9.1	7.6
Other occupied not attributed to a sector	5.8	–	–	10.0	–	10.1

Sources. For 1939: our own estimates based on the 1939 National Register. For 1901–1951 (with the exception of 1939): B. R. Mitchell, *British Historical Statistics,* (Cambridge, 1988, 104–106).

Notes: the definition of the Standard Industrial Classification did undergo some changes during the 1901-1951 period. This makes direct comparisons over time in principle difficult to make, even between large levels of aggregation of occupation groups as presented in this table. For more detailed information on these changes, we refer to Mitchell (1988, 104–106).

prime example of this can be seen for textile manufacturing. With Britain at the height of its colonial power prior to the First World War, textile production remained the most important employer in the industrial mix in 1911 (Mukherjee 2010). One in three industrial workers were employed in this industry. This declined strongly during the interwar period: falling from 16.4 percent in 1901 to 7.7 percent by 1939. A similar story can be found for the metal industry, where output had dropped by 1939 to 87 per cent of its level by 1930 (Broadberry 1997). Prior to the First World War, the British economy had facilitated the import of raw materials from abroad, before exporting processed products to the British possessions (most notably India and Australia), Europe and the USA. Following the war, textile and metal production suffered from the loss of comparative advantage in the international market, as the US and Japan rose as exporters over the course of the interwar period, influencing the decline of the relative proportion of employment in the secondary sector (Sunley 1992). Whereas the economic depression hit particularly hard export-oriented industries such as the shipbuilding industry, the rearmament boom at the end of the interwar period induced a more stabilizing effect on the metal and transport producing sectors during the studied period (Capie and Collins 1980, 45). At the same time, in contrast, the construction sector was one of the leading sectors during the interwar period (Capie and Collins 1980, 40), with an increase of around 2 percent points during 1921–1939.

Regional trends in the occupational structure

These national trends, however, tend to obscure the diversity at the regional level. Figure 1 illustrates the proportion of each sector in the labor force by region between 1911 and 1939.

Here, we find the decline of employment in agriculture and fishing to be occurring in nearly all of the regions in England and Wales. Nonetheless, there were some regional exceptions to this rule: so, for certain areas in the North and East of the country, a growth in the relative size of agriculture in the labor force occurred. See, for example, the county of Northumberland in North East England, where the percentage of the population employed in agriculture went from 7% to 12%, or the county of Norfolk in East of England where the relative weight of agriculture went from 26% to 28%. Even in the county of Lancashire and its immediate surroundings, areas with the lowest levels of employment in the agricultural sector in the country throughout the interwar period (due to the dominance of the industry sector in this region), the relative share of agriculture in the occupational structure remained constant, around 4%.

The tertiary sector saw a rise during the same period. Figure 1 shows this trend across the whole

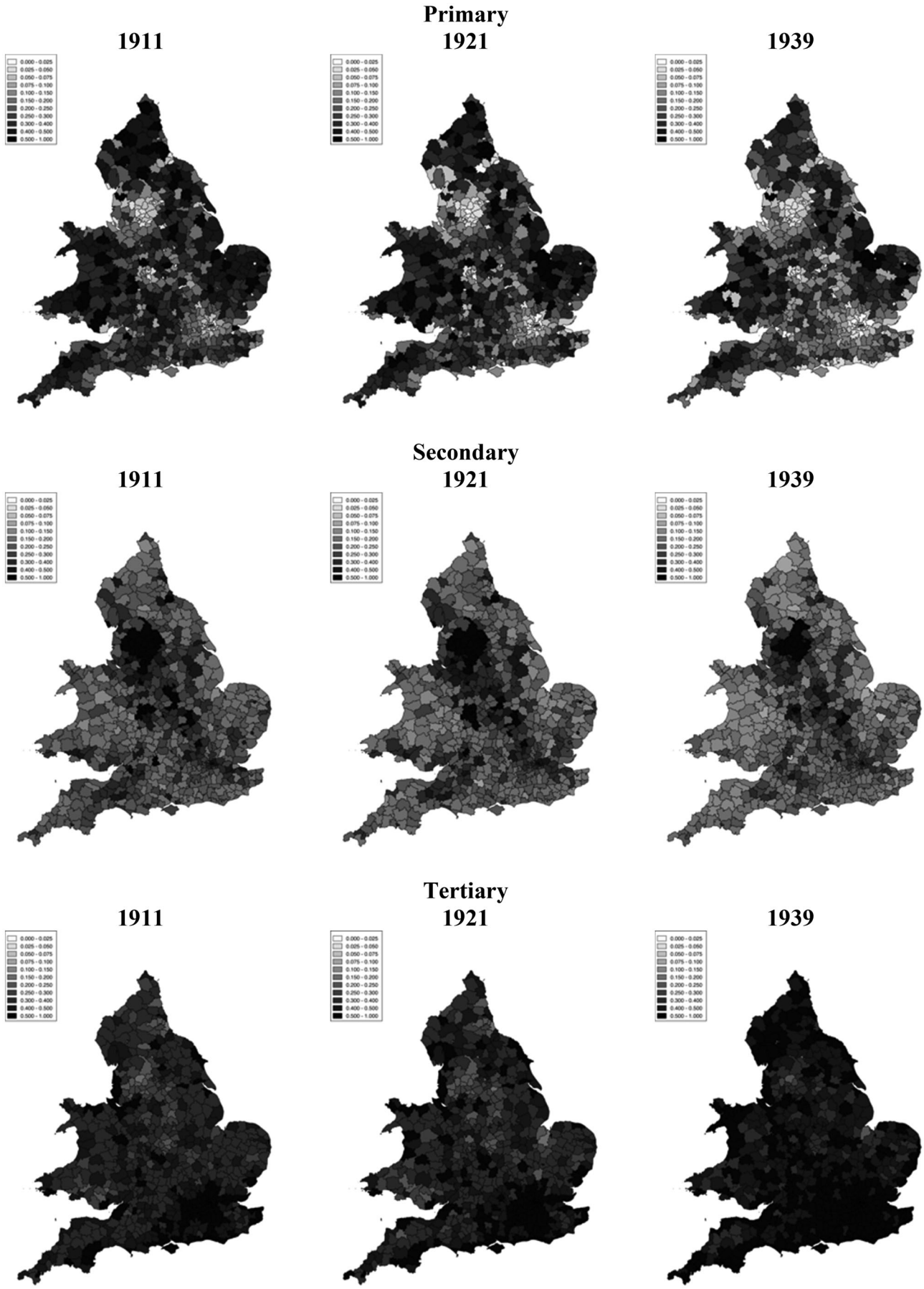


Figure 1. Regional distribution of occupied population by primary, secondary and tertiary sector, 1911–1939, as a share of the labor force.

country. This was particularly prevalent in the Northern regions of England and in Wales: for instance, the sector of services and professions grew for more than 10% in North East and North West England and around 9% in Wales. With these percentages, these regions grew more than the south-south-east regions of England with their traditional connection to the trading belt grew to serve the population of the sprawling London metropolis.

The greatest regional disparity, however, is found in the shift in the proportion of the workforce employed in secondary occupations. In particular, changes occurred in the industrial heartlands of the northwest. In 1911, patterns of industry were concentrated in outward reaching clusters from the northwest to the southeast. In 1939, it is clear that the manufacturing belt had withered in comparison with the growth of the tertiary sector. While the industrial hub in the northwest still existed in 1939, its proportion of employment contrasts starkly with its situation in 1911 and 1921. Other areas with high levels of employment in the secondary sector in 1911, most prominently the diagonal manufacturing belt stretching from Lancashire through the Midlands to the southeast and the southwest of England, also faced a decline of 20 to 30 per cent during the interwar period.

Regional trends in the location of industries

Many of these (relative) changes in the regional occupational structure can be explained by the development of specific industries. The atrophy of the secondary sector as a whole was largely due to the decline in the proportion of the labor force employed by the textile industry (see Table 2). By 1939, we find that textile production only remained a dominant industry in its traditional heartland, particularly a relatively small cluster in Northwest England. In this region, the proportion employed was still approximately 20 per cent for most districts, while in other traditional centers of textile manufacture – such as the districts in Dorset and Somerset in the southwest – it went through an average reduction of around 7%. A fall in employment in the metal industry also influenced the downsizing of the secondary sector before the rearmament boom of the late 1930s.

The declining employment in the textile industry does not account for the fact that employment in industry as a whole (in absolute numbers) did rise during the interwar period. Instead, it has been argued that whereas the first industrializing nation was a late participant in the technologies of the second industrial revolution (e.g., Mokyr 1999), protectionism, industrial mergers and anti-competitive pricing induced by cartel agreements ensured that Britain was able to sustain the growth of a chemical production and refinery industry that dominated this second revolution (e.g., Bowden and Higgins 2004). Figure 2 illustrates that for chemical production and refinery industries, clusters emerged throughout the interwar years, for example in Caernarfonshire and Cardiganshire in Wales, in Essex and Norfolk in the East of England, in Cheshire in the North West of England and Cornwall and Dorset in the South West of England. Another contributor to the rise of employment in industry in terms of absolute numbers was the growth of the metal manufacturing sector and the construction sector. Figure 2 shows the expansion of employment in these sectors during the interwar period, which benefited most noticeably the West Midlands, but also, for example, Newcastle, Exeter and Sheffield in respectively the far northeast, southwest and east of England.

Sectoral concentration and regional specialization

Geographical concentration in the occupational structure

There were thus obvious changes in the regional distributions of economic sectors between 1911 and 1939, which can be measured as geographical concentration (which measures the geographical distribution of a particular economic sector) and regional specialization (which indicates the sectoral distribution in a particular region). These are important variables for understanding the historical development of economies of scale. Starting with geographical concentration, in order to assess whether or not particular sectors became more concentrated in particular regions, we calculated the level of geographical concentration in the primary, secondary and tertiary sectors. We define this index of geographic concentration as $K_j^C = \sum_{i=1}^{n} \mid g_{ij}^C - g_i \mid$ where g_{ij}^C is the share of region i in the total national employment in sub-sector j and g_i is the share of employment of region i in national employment. The index of geographic concentration ranges from 0 (for sectors with a low level of localisation) to 2 (where the sector is localized in a limited number of regions), following Goschin et al. (2009, 98–99). In Table 3, we report the resulting index for England and Wales at the region level, for the nine regions in England in addition to the region

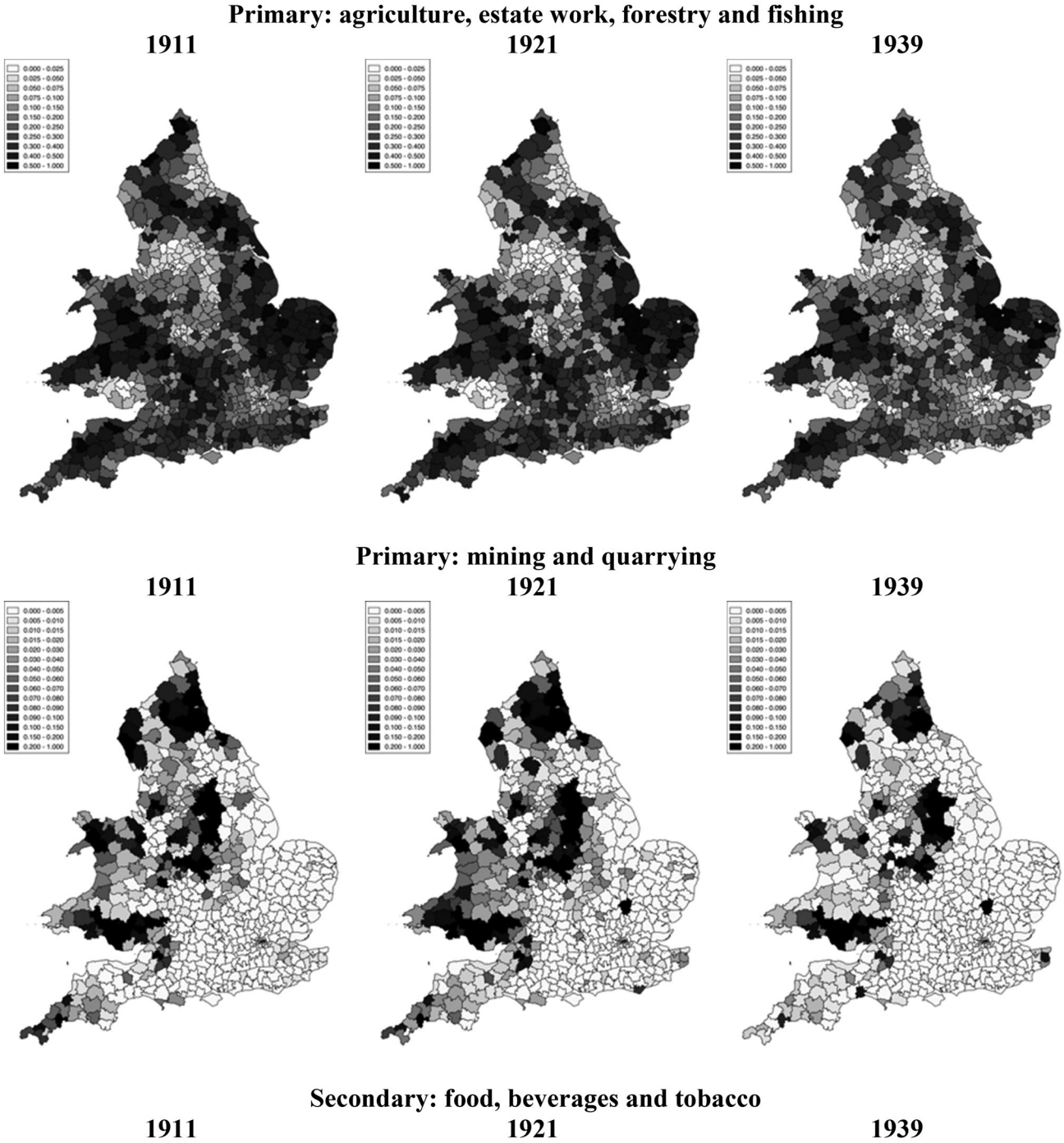


Figure 2. Regional distribution of occupied population by sub-sectors, 1911–1939, as a share of the labor force.

of Wales[14], but also to a large extent at a more deepened spatial scale: the levels of counties and districts[15]. Here, after an initial increase, we find that geographical concentration declined in agriculture (but not in mining) between 1911 and 1939. We find the same inverted U-shaped curve also for industry and services, with an increasing trend from 1911 to 1921 followed by a decreasing trend from 1921 to 1939.

Under certain assumptions, our results can be compared with those of Kim (1998) for the United States during the interwar period, as reported in Table 3[16]. In particular, our results seem in line with the relative low levels of concentration for the tertiary sector (at least, when compared to the primary sector and secondary sector), whereas the agriculture sector shows the highest levels of concentration in our results for England and Wales and the results of Kim (1998) for the United States.

Furthermore, we find some interesting variations between different levels of spatial aggregation in Table

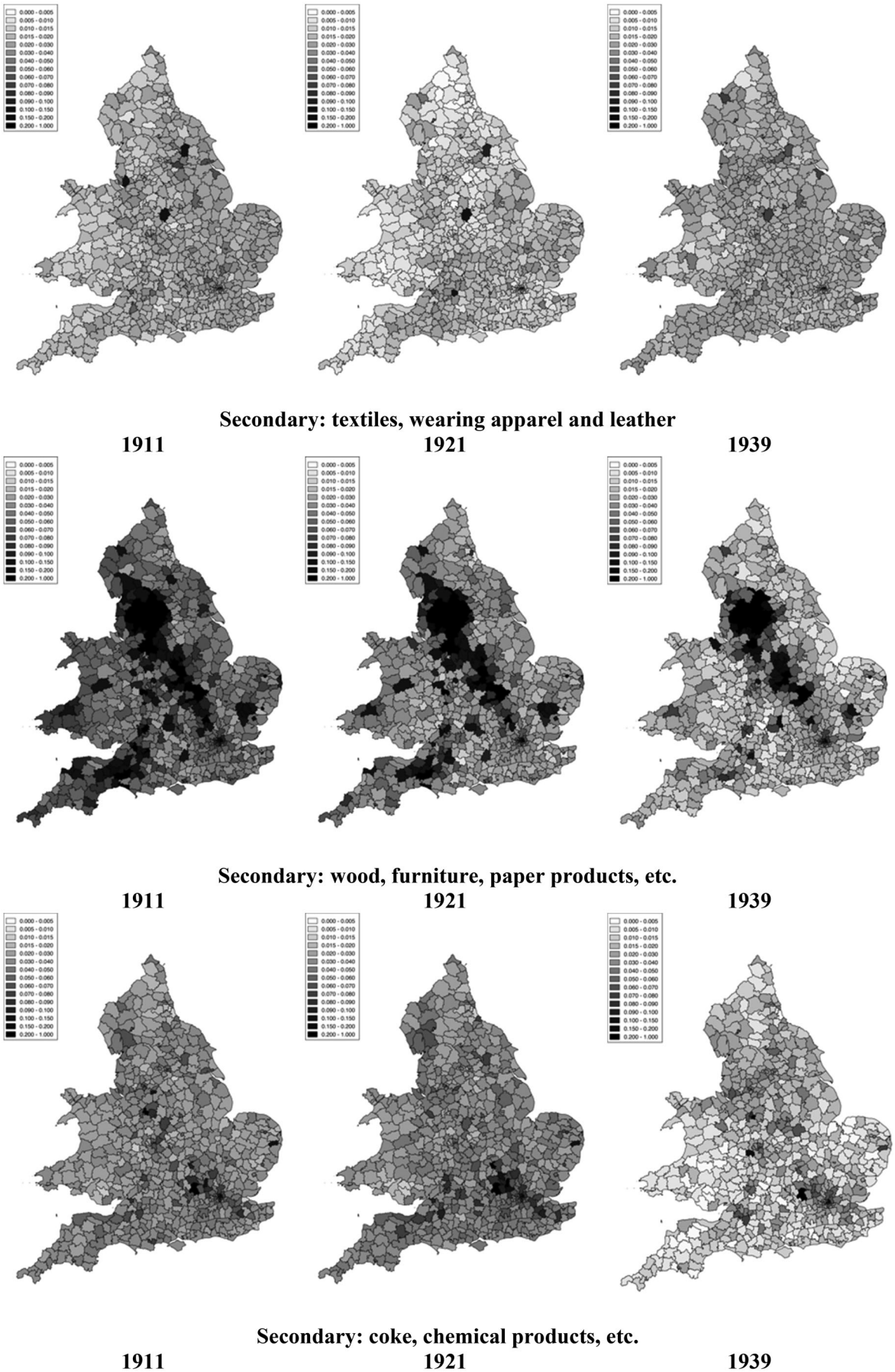
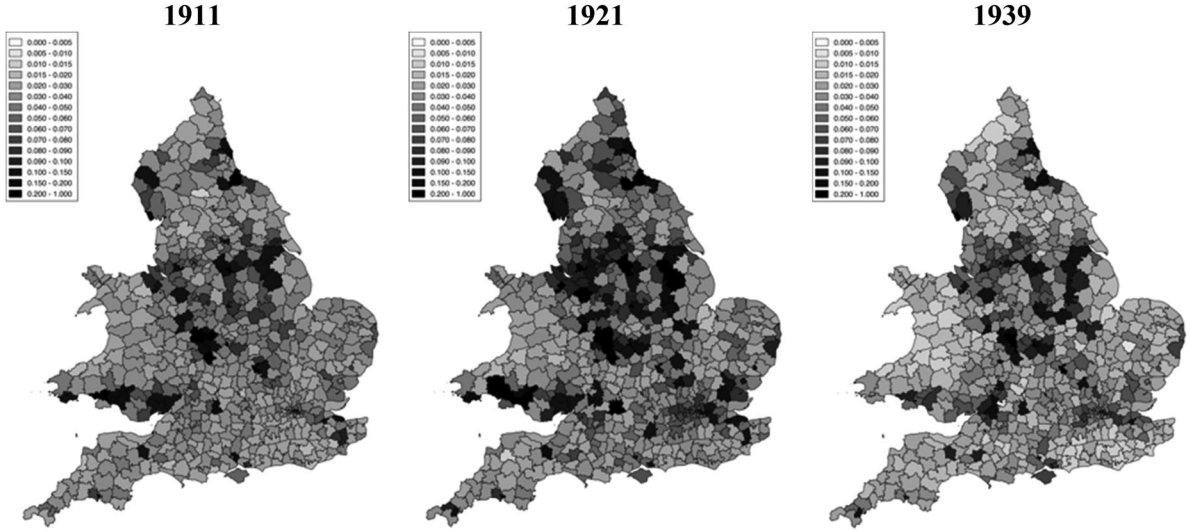
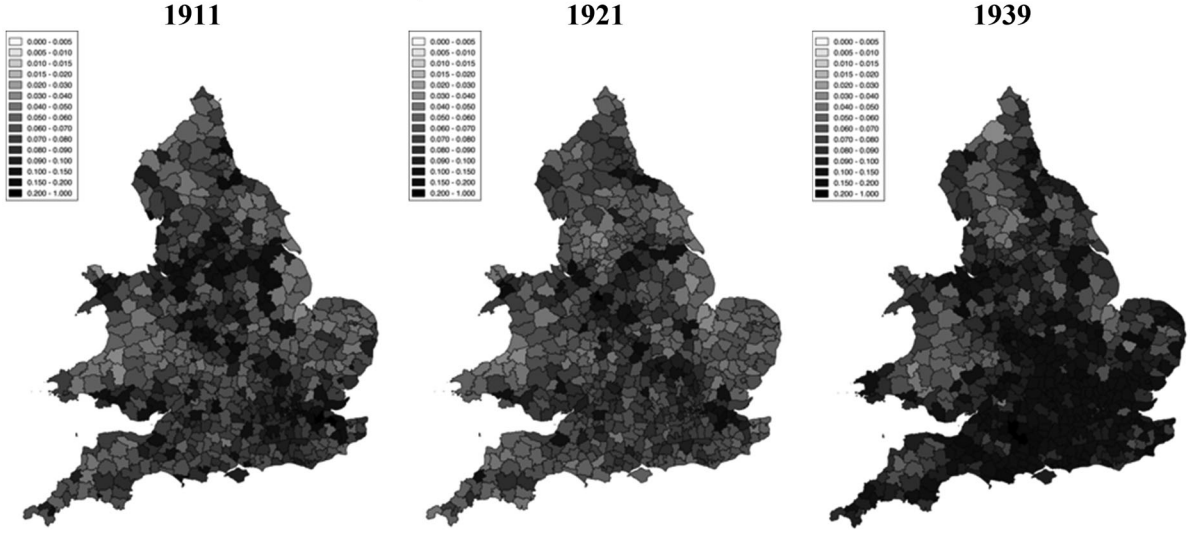


Figure 2. Continued.

Secondary: basic metals, transport products, others, etc.



Secondary: construction and construction work



Secondary: other manufactured goods and repair

Figure 2. Continued.

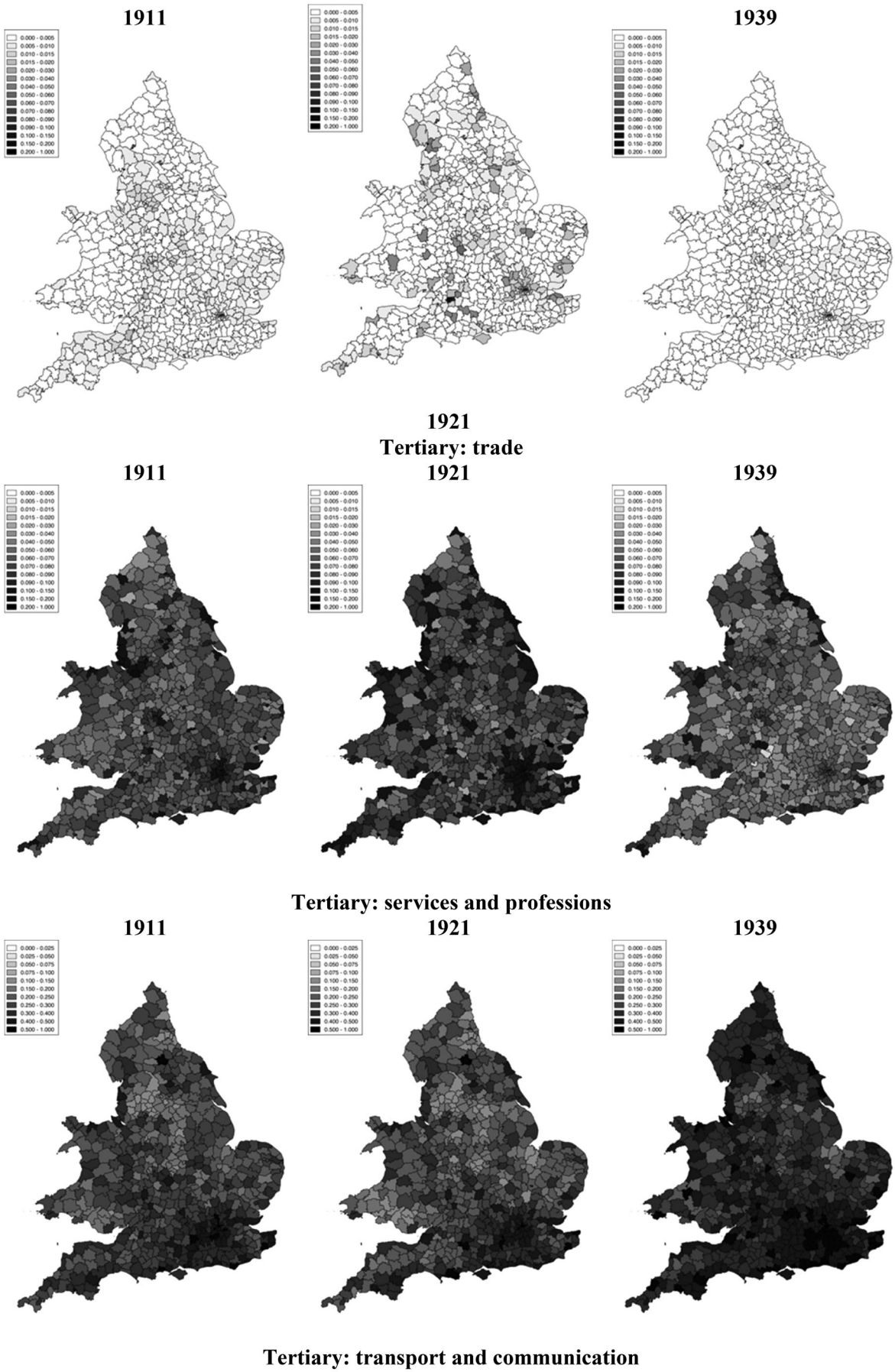


Figure 2. Continued.

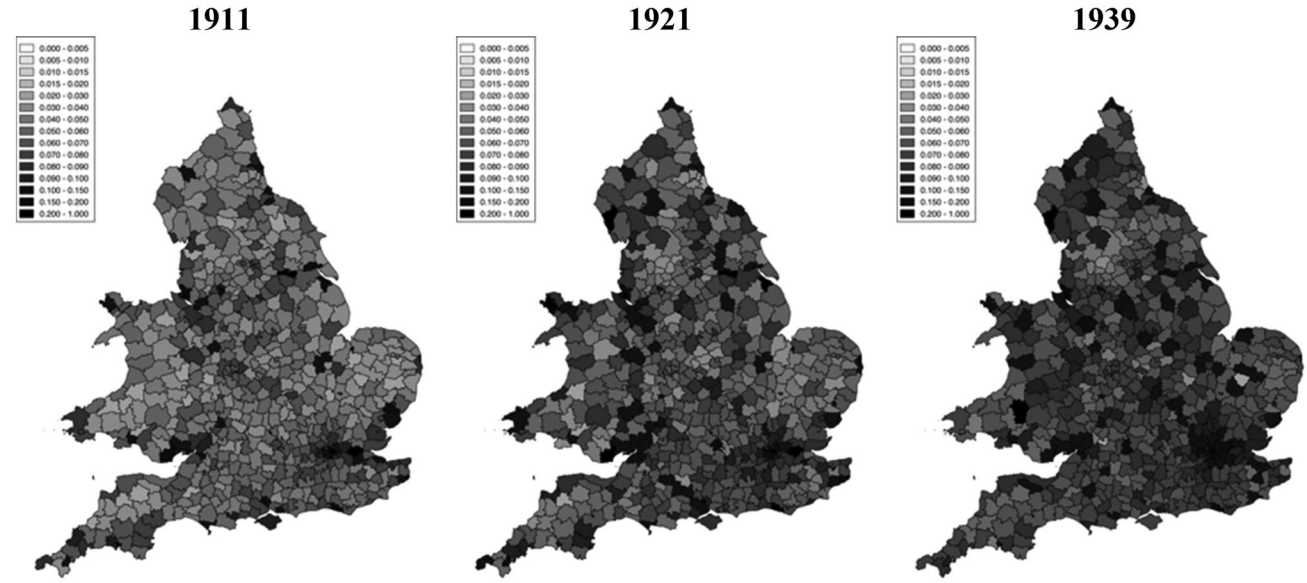


Figure 2. Continued.

Table 3. Geographical concentration index: comparison between our estimates (1910–1939) for England and Wales and Kim (1998) for US, for the primary, secondary, and tertiary sector.

	1900	1910	1920	1930	1939	1950
England and Wales (our results), Region level						
Primary		0.633	0.708		0.687	
Secondary		0.237	0.262		0.257	
Tertiary		0.151	0.171		0.107	
England and Wales (our results),County level						
Primary		0.404	0.544		0.465	
Secondary		0.164	0.188		0.187	
Tertiary		0.091	0.096		0.068	
England and Wales (our results),District level						
Primary		0.596	0.801		0.772	
Secondary		0.285	0.415		0.440	
Tertiary		0.186	0.260		0.250	
United States (Kim 1998), US state level*						
Agriculture	1.040		1.080			1.160
Secondary	0.750	0.890		0.890	0.870	0.810
Tertiary				0.180	0.180	0.150

Note: * For reasons of simplicity, the benchmark years presented in the table deviate to a small extent from the benchmark years for which Kim (1998) calculated the concentration index. Concentration in the agriculture sector was calculated for the benchmark years of 1900, 1920 and 1956; in the secondary sector for 1900, 1914, 1927, 1939 and 1947; and for the tertiary sector for 1929, 1939 and 1958.

4. We find that geographical concentration in agriculture decreased from 1911 to 1939 at the region level, whereas for the county and district levels, we find an overall increase between 1911 and 1939. Additionally, whereas we find an inverted U-shaped curve at the region level, for the secondary sector geographical concentration increased at the county and district level between 1911 and 1939. For the tertiary sector, instead, we find an inverted U-shaped trend at the region and county level, whereas at the district level there is an increase in concentration during the interwar years.

Accordingly, with the declining relative proportion of agriculture in the labor force and the increasing relative proportion of services in all regions, we find that the occupational structure became more similar over the larger regions, due to forces of market integration (Crafts and Mulatu 2005). However, with the growing levels of geographical concentration for agriculture and the secondary sector during 1911–1939 at the county and district levels, our estimates suggest that the occupational structure became more dissimilar at smaller geographical levels. These results indicate that the benefits of market integration were distributed unevenly across districts during the interwar period (a spatial scale where i.a. rural-urban differences come more into play), and thus increased the discrepancies between our estimates at different levels of spatial aggregation.

Table 4. Geographical concentration index for the primary, secondary and tertiary sector, 1911–1939.

Region level	1911	1921	1939	avg.
Primary	*0.633*	*0.708*	*0.687*	*0.676*
Agriculture, estate work, forestry and fishing	0.463	0.540	0.415	0.472
Mining and quarrying	0.803	0.877	0.959	0.879
Secondary	*0.237*	*0.262*	*0.257*	*0.252*
Food, beverages and tobacco	0.091	0.179	0.091	0.120
Textiles, wearing apparel and leather	0.540	0.541	0.615	0.565
Wood, furniture, paper products, etc.	0.226	0.212	0.200	0.212
Coke, chemical products, etc.	0.212	0.254	0.246	0.237
Basic metals, metal products, etc.	0.326	0.361	0.285	0.324
Other manufactured goods and repair	0.197	0.180	0.248	0.208
Construction and construction works	0.066	0.109	0.120	0.098
Tertiary	*0.151*	*0.171*	*0.107*	*0.143*
Trade	0.070	0.108	0.054	0.077
Services and professions	0.248	0.267	0.163	0.226
Transport and communication	0.134	0.140	0.105	0.126
County level	**1911**	**1921**	**1939**	**avg.**
Primary	*0.404*	*0.544*	*0.465*	*0.471*
Agriculture, estate work, forestry and fishing	0.355	0.507	0.481	0.447
Mining and quarrying	0.454	*0.582*	0.449	0.495
Secondary	*0.164*	*0.188*	*0.187*	*0.179*
Food, beverages and tobacco	0.141	0.142	0.091	0.124
Textiles, wearing apparel and leather	0.150	0.274	0.233	0.219
Wood, furniture, paper products, etc.	0.130	0.112	0.207	0.149
Coke, chemical products, etc.	0.365	0.239	0.211	0.271
Basic metals, metal products, etc.	0.171	0.166	0.327	0.221
Other manufactured goods and repair	0.110	0.283	0.192	0.195
Construction and construction works	0.069	0.106	0.054	0.076
Tertiary	*0.091*	*0.096*	*0.068*	*0.085*
Trade	0.062	*0.075*	0.055	0.064
Services and professions	0.078	*0.092*	0.071	0.080
Transport and communication	0.133	*0.122*	0.080	0.111
District level	**1911**	**1921**	**1939**	**avg.**
Primary	*0.596*	*0.801*	*0.772*	*0.723*
Agriculture, estate work, forestry and fishing	0.471	*0.738*	0.759	0.656
Mining and quarrying	0.722	*0.864*	0.785	0.790
Secondary	*0.285*	*0.415*	*0.440*	*0.379*
Food, beverages and tobacco	0.219	0.366	0.324	0.303
Textiles, wearing apparel and leather	0.287	0.431	0.496	0.404
Wood, furniture, paper products, etc.	0.245	0.289	0.520	0.351
Coke, chemical products, etc.	0.506	0.535	0.578	0.539
Basic metals, metal products, etc.	0.305	0.448	0.454	0.402
Other manufactured goods and repair	0.240	0.635	0.504	0.459
Construction and construction works	0.192	0.199	0.208	0.199
Tertiary	*0.186*	*0.269*	*0.250*	*0.235*
Trade	0.151	0.268	0.319	0.246
Services and professions	0.154	0.230	0.209	0.197
Transport and communication	0.253	0.309	0.224	0.262

Geographical concentration in the industrial structure

So far, we have looked at the differences in geographical concentration across the primary, secondary and tertiary sector over time. However, there might have also been differences in geographical concentration by industrial sub-sectors. After all, as noted in the previous sections, the overall development of industry was far from uniform across secondary sectors (see e.g., Section 3). Table 4 presents the level of geographical concentration of individual industries between 1911 and 1939 and includes an average geographical concentration level for the whole period. Our previous findings indicated at the region level an increase in geographical concentration from 1911 to 1921 and a decrease from 1921 to 1939. Our findings again indicate an inverse U-shaped pattern in the level of geographical concentration for most manufacture sub-sectors during 1911–1939, at the region level.

For the 1911–1939 period, besides agriculture and mining, the textiles and metal producing sectors present the highest values for geographical concentration for all the levels of spatial aggregation. These findings are in line with Henderson, Kuncoro, and Turner (1995), who argued that Marshallian externalities are highest for the 'traditional' manufacturing sectors, most notably the textile sector[17]. Similarly, as suggested by Crafts and Mulatu (2005, 513): whereas Britain was losing the comparative advantage deriving

Table 5. Krugman specialization index for the manufacturing sector in England and Wales compared to selected international studies.

	1900	1910	1920	1930	1939	1950
England and Wales (our results), Region level		0.391	0.406		0.344	
England and Wales (our results), County level		0.279	0.354		0.275	
England and Wales (our results), District level		0.385	0.497		0.404	
United States (Kim 1998)*	0.750	0.890		0.860	0.870	0.810
England and Scotland (Crafts and Mulatu 2005)* **	0.590	0.610	0.790	0.720		
Sweden (Berger, Enflo, and Henning 2012)*	0.280	0.270	0.290	0.320	0.330	0.300
Spain (Betran 2011)*		0.790		0.906		0.666

Notes: * Some of the exact years deviate to a small extent from the header: Spain (1913; 1929; 1955), US (1914; 1927; 1939; 1947) and England (1901, 1911, 1921, 1931). Additionally, as Berger, Enflo, and Henning (2012) calculated an index for regional specialization on multiple geographical levels and multiple indices, for this overview we used their estimates of the Theil index for the NUTS3 level.

** We use the same sectoral grouping for the manufacturing sector (with the same 16 sub-sectors of manufacturing) as Crafts and Mulatu (2005). The other studies mentioned in this table calculate the index based on another occupational classification and different grouping for the manufacturing sector. Kim (1998), for example, uses the Standard Industrial Classification (S.I.C.).

from its legacy as the first country to experience the industrial revolution, clustering dynamics were put in place as a reaction to the interwar protectionist crisis. This is particularly evident in the increase of the geographical concentration levels during 1921–1939 for the textiles sector, most notably in Lancashire and Northamptonshire. In 1921, textile and metal production shared a dominant position as the largest industries in Britain, but due to the decline in international trade and the loss of Britain's comparative advantage, both sectors showed a relative decline.

A similar trend but supported by different factors is found for the 'new' industries of the second industrial revolution: the production of chemicals, transport and other miscellaneous products. For chemical products, we find an increasing trend at the region and district levels. Chemical clusters increasingly emerged, for example in the Midlands (see Figure 2), with more counties and districts – e.g., in Lincolnshire, Northamptonshire, Shropshire, and Herefordshire – within the Midlands being able to attract chemical plants (Heim 1983). For transport manufacture, we again find an inverse U-shaped pattern at the region and district levels, but with a substantial reduction of the geographical concentration values for all the levels as of 1939. Technological innovation, a progressive reduction in transport costs (cf. Crafts and Mulatu 2005, 500), an oversupply crisis in shipbuilding – in the aftermath of the First World War – and increasing competition from American steel and ship manufacturers (Thomas 2004) left many traditional clustering centers, such as the northwest of England (Northumberland), in fiercer competition from the 1930s onwards (Murphy 2018) – all concurred to an increasing dispersion of employees in this sub-sector.

Regional specialization in industry sectors

Not only were there differences in the patterns of geographical concentration across individual industries, similarly, we can also expect regional differences of specialization. Hence, we turn to patterns in regional specialization in the industry sub-sectors, based on differences in the industrial structure. In order to assess the degree of regional specialization, we calculated in Table 5 a Krugman index of relative regional specialization in the manufacturing sectors – based on the same sectoral division as Crafts and Mulatu (2005, 507) – for all the three geographical levels (regions, counties, districts). We defined this index as: $KSI_r = \sum_i \mid s_{r,i} - s_i \mid$, where $s_{r,i}$ is the proportion of manufacturing sector $_i$ out of the total employment in region $_r$ and s_i is the proportion of manufacturing sector $_i$ in the country overall. The numerical value of the specialization index ranges from 0 (in cases where the regions have a sectoral structure identical to the national structure) to 2 (where the sector structure is completely different across regions)[18].

We find that the degree of regional specialization in manufacturing follows a common pattern at all the three reported geographical levels (regions, counties, districts): an increase during 1911–1921, followed by a decline during 1921–1939. Keeping in mind the differences in definitions of the industry sector in other studies, we can put our results in an international perspective: our results are similar to the inverse U-shaped trend in regional specialization of industry during the interwar period found in other countries (e.g., Kim 1998; Berger, Enflo, and Henning 2012; Betran 2011). In Britain, this is most likely to have been stimulated by increased international competition, as well as increased protectionism within European trade policy during the interwar period. In addition, we confirm the findings of Crafts and Mulatu (2005), who calculated the Krugman index at the region level for England. Despite the methodological differences between our calculation and that of Crafts and Mulatu (2005), both our results and the results of Crafts and Mulatu (2005) show the highest

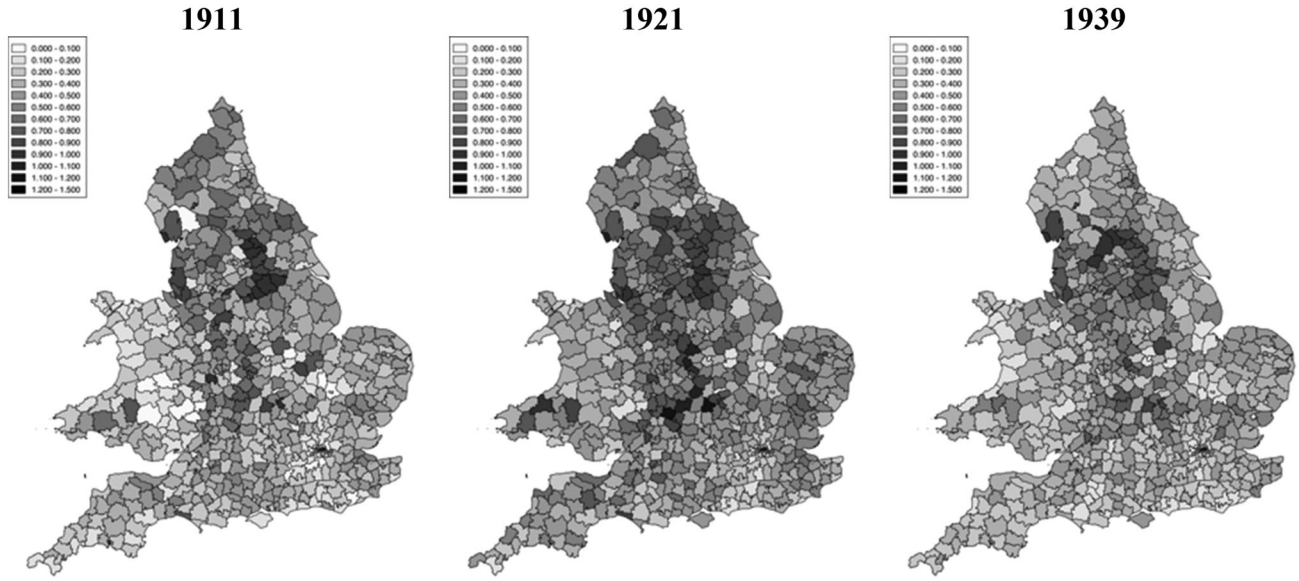


Figure 3. Krugman specialization index.

level of regional specialization in manufacturing in England and Wales in 1921[19].

In Figure 3, we calculate the Krugman regional specialization index by district in England and Wales. This allows us to pinpoint the centers of regional specialization in industry. In particular, for 1911, we find that most of the industrial specialization can be attributed to the above-outlined heartlands of metal, textiles and chemical production in the country. For example, in Lancashire, there is in general a high level of industrial specialization in 1911 and 1939 (see Figure 3). Basically, as a result of economies of scale, there was an increase in the levels of specialization for these highly-specialized districts during 1911–1921 (in line with the theories of e.g., Krugman Venables 1995 and the findings of Crafts and Mulatu 2005). Yet during 1921–1939, also led by the concurrent de-specialization dynamics of the growing tertiary sector, even at the district level, the industrial specialization declined. In particular, the spectacular growth of road transport as an alternative to rail or sea haulage for businesses, dramatically decreased transport costs. Consequently, the benefits of agglomeration and Marshallian externalities – as well as the benefits of specialized factor endowments – declined, for which most notably highly-specialized regions such as the districts in the Lancashire cluster suffered from this evolution[20].

Conclusion

In this paper, we present a recently constructed database on the occupational structure in England and Wales in 1939 at the district level, based on the 1939 National Register. We link this with two population censuses to bridge a gap in the literature for the interwar period. In particular, we standardized and linked our dataset with the population censuses of 1911 and 1921. By reconstructing the occupational structure at a spatially disaggregated level, we have been able to extend our knowledge of regional economic structures and patterns of regional specialization in England and Wales, beyond Crafts and Mulatu (2005) study for the period 1871–1931. This dataset was used to produce empirical evidence regarding the levels of geographical concentration in the agriculture, industry and service sectors.

Our empirical evidence lends support to a peak of regional industrial specialization during the interwar period, with regional industrial specialization increasing during the first decades of the twentieth century and starting to decrease prior to World War II, as found for other regions of Europe (e.g., Wolf 2007; Betran 2011; Berger, Enflo, and Henning 2012). We also found increasing geographical concentration in industry and agriculture at the county and district level during this period, together with decreasing concentration in the service sector at the region and county level. In 1939, the English manufacturing industry was highly focused in the northwest, accounting for around 46 per cent of the labor force in Lancashire. Increasing level of manufacture and agriculture concentration could have been the reaction to increased protectionism in European trade policy during the interwar period, as well as increased competition from US and Japanese manufacturers following the First World War. As exports to Europe

became more expensive and global markets became more competitive, British industry began to decline. The drop in the proportion of the labor force involved in manufacturing was also likely to have been due to automation in certain industries, for example food and textile production. It is an easy task to see the decline of the textile industry and its contribution to the overall regional decline of the distribution of employment in the secondary sector. Figures 1 and 2 show great similarities in this regard. It is, however, more difficult to see the national effects of the growth of the metal and chemical industries during the same period. The expansion of the tertiary sector by 1939 was so great as to outpace the growth in industrial employment in absolute numbers. While each region experienced its own variation of tertiary growth and industrial decline, the national image is striking. As Britain moved to become a service-oriented economy, the industries that could not innovate and compete fell behind. Metal and chemical production industries would remain important into the Second World War, but their employment figures were eclipsed by the growing numbers employed in professions, trade and transport.

The results presented in this paper open up possible avenues for future research. There are likely to be other factors involved that explain the regionally varied increases in specialization levels. A movement away from the technologies of the First Industrial Revolution toward those of the Second Industrial Revolution appears to have been taking place during this period. Future research could strive toward further exploiting this dataset, which has been used in the current work to bridge a previous empirical lacuna concerning the interwar period.

Notes

1. This was done by *The occupational structure of Britain 1379-1911* project, a research project (funded by the British Academy Research Project since 2007) directed by Leigh Shaw-Taylor, Amy Erickson, and Tony Wrigley aimed to reconstruct the evolution of the occupational structure of Britain from the late medieval period down to the early twentieth century. For more information, we refer to the CAMPOP project website: https://www.campop.geog.cam.ac.uk
2. The register lists information for the entire population of England and Wales, with two exceptions: military personnel present in barracks and fishermen at sea at the time of enumeration.. For more information, we refer to: http://www.nationalarchives.gov.uk/help-with-your-research/research-guides/1939-register/#4-how-the-register-was-compiled-and-arranged and https://www.findmypast.co.uk/1939register/what-is-the-1939-register
3. Ultimately, when the National Register Bill was passed in September 1939, it was intended to achieve three goals. First, to gather statistical information on the population. Second, to keep track of a population mobilizing for war and dislocated by evacuation, so as to plan a potential war-oriented economy. Third, to prepare for the possibility of rationing.
4. Following the war, the data from the National Register was also used in coordinating efforts of the National Health Service, which maintained the register during the 1948–1991 period. See: '1939 Register', nationalarchives.gov.uk.
5. Although it is not restricted by census publication rules, data protection legislation still applies to the National Register. Accordingly, public information is still not available for any individual recorded in the register who was born less than 100 years ago, or who died after 1991. However, this does not affect the metadata available for researchers.
6. Each record includes the information detailed, as well as a unique code made up of an enumeration district, a household or institution schedule number and an individual National Registry sub-number (Bringing the 1939 Register Online, findmypast.com; 1939 Register, nationalarchives.gov.uk).
7. For more information on the PST classification system - including working papers, definition tables, look-up tables, dictionaries, as well as previous versions of the PST classification system -, we can refer to the following page, developed by the CAMPOP project: https://www.campop.geog.cam.ac.uk/research/occupations/datasets/coding/
8. This is related to the nature of the PST classification system itself, which often categorizes similar occupations over a variety of PST codes at the 4-digit level. For instance, whereas the PST system subdivides drivers of vehicles into different PST codes based on the types of goods transported and the type of vehicles, we often had to categorize drivers of vehicles – due to a lack of information – as 'drivers of motorized vehicles' (PST code 6,2,0,1).
9. 1911 Census of England and Wales, IPUMS International, Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 7.2 [dataset]. Minneapolis, MN: IPUMS, 2019. https://doi.org/10.18128/D020.V7.2;
10. 1921 Census of England and Wales, VisionsOfBritain, Great Britain Historical GIS Project (2017) 'Great Britain Historical GIS'. University of Portsmouth.
11. See, for instance, the Census Report of the 1921 population census: "The returns of occupation in this census have been tabulated under a scheme differing so much from those in use previously as to preclude the possibility of an exact comparison with previous census results. This has come about because of failure in the past to maintain a clear distinction between occupation—the employment of the individual, and industry—the employment of the firm, or body of individuals organised under a common directing head." For more information, we refer to the Census

Report of the 1921 population census (also digitally available via https://www.visionofbritain.org.uk/census/EW1921GEN), under the sub-section 'Revised Classifications of Occupations and Industries'.

12. The mining sector is considered as part of the primary sector in the PST coding. Yet, the attribution of the mining sector is somewhat contested in literature: whereas it is grouped in the secondary sector by Shaw-Taylor (2009), it is not one of the 16 sub-sectors of the manufacturing sector in the study of Crafts and Mulatu (2005). In our reported results, we attributed the mining sector to the primary sector.
13. Again, we can attribute the growing value for mining of Mitchell in 1931 to different grouping conventions and to the fact that his results should be adjusted on a constant 10% of the labour force classified in his statistics as not attributed to a specific sector – see Tab. 2
14. We distinguish in England between the following nine regions: London, South East, East of England, South West, West Midlands, East Midlands, North West, Yorkshire and the Humber, North East.
15. We distinguish in England and Wales between 53 counties, following the list of administrative counties made in the aforementioned published census tables digitized by Visions of Britain. Whereas the number of regions and counties remained constant for our results in the three benchmark years, the number of districts varied over time: from 635 districts in 1911 to 1169 districts in 1921 and 1481 districts in 1939. Hence, all estimates for concentration and regional specialization at the district level in Table 3, 4 and 5 were estimated with a different number of spatial units, which might hold implications for a comparison over time at this spatial scale.
16. So, any comparison between our study and the study of Kim (1998) is difficult to make, due to different definitions in the level of spatial aggregation and the sectoral subdivision. On the one hand, we compare different sizes of regions: US states compared to regions/counties/districts in England and Wales. Not only are the US states greater in average surface, also they are more populous on average. On the other hand, differences in methodologies used to reconstruct the labour force exist between our study and the study of Kim (1998). For instance, the tertiary sector in the pre-World War II estimates of Kim (1998) only includes the retail trade. Also, differences exist in the definition of the secondary sector: for Kim (1998) this includes all manufacturing sectors and excludes mining and construction, whereas for our estimates we include all manufacturing sectors and the construction sector, but do not include the mining sector.
17. In contrast, as found by Berger, Enflo, and Henning (2012, 303) for Sweden, the textile industry incrementally increased its geographical dispersion during the 1930s, moving outwards from the heartland of the northwest and the southwest.
18. These Krugman-type indexes have a full set of characteristics that make them a reliable indicator of the variation of proportions. They respect the principle of 'anonymity', by ensuring that the resulting degree of specialisation is the same for different permutations of the same employment proportion, and the 'Pigou-Dalton Principle', by ensuring equity of allocation throughout the rankings. As a result, the index of geographic concentration of sectors should be considered as similar to the Krugman specialisation index, or as 'two sides of the same coin' (Goschin et al. 2009).
19. The deviation between our estimates and the estimates of Crafts and Mulatu (2005) is the result of different definitions and a different use of geographical units to calculate the Krugman specialization index. First, Crafts and Mulatu (2005) use a different source for the reconstruction of the labour force, the occupational statistics of Lee (1979), thereby presenting methodological differences with our reconstruction. Second, we based statistics on the attributed PST code to each occupation. As mentioned, the PST introduced different conventions for the attribution of certain categories throughout the sectors (e.g. the dealers involved in food or textile wholesale). Furthermore, we consistently follow an industry-driven approach (e.g. an engineer working in mining is coded in the mining sector) that was not consistently followed in the census conventions between 1911 and 1951 (cf. Mitchell 1988, 104–106). Third, Crafts and Mulatu estimate the specialization index for England, Wales and Scotland (whereas we restrict our reconstruction to England and Wales), and differentiate between 8 regions of England (whereas we differentiate between 9 regions). Nevertheless, in an attempt to make the comparison of our results as much as possible in line with those of Crafts and Mulatu (2005), we also subdivided the manufacturing sector into 16 sectors and we calculate the specialisation index exlusively based on the manufacturing sector (that is, including all manufacturing sectors, but excluding the mining sector and the construction sector).
20. By 1938, there were almost two million private motor vehicles on the roads in Britain, as well as an additional half a million motorcycles and approximately another half a million commercial vehicles (Thomas 2004).

Funding

The research leading to these results received funding from the European Research Council under the European Union's Horizon 2020 Programme / ERC-StG 637695 - HinDI, as part of the project "The historical dynamics of industrialization in Northwestern Europe and China ca. 1800-2010: a regional interpretation".

ORCID

Robin C. M. Philips http://orcid.org/0000-0002-1436-8573
Matteo Calabrese http://orcid.org/0000-0003-2835-3417
Robert Keenan http://orcid.org/0000-0002-7402-4131
Bas van Leeuwen http://orcid.org/0000-0001-7817-5298

References

Berger, T., K. Enflo, and M. Henning. 2012. Geographical location and urbanisation of the Swedish manufacturing industry, 1900–1960: Evidence from a new database. *Scandinavian Economic History Review* 60 (3):290–308. doi: 10.1080/03585522.2012.727764.

Betran, C. 2011. Regional specialisation and industry location in the long run: Spain in the US mirror (1856–2002). *Cliometrica* 5 (3):259–90. doi: 10.1007/s11698-010-0060-1.

Bowden, S., and D. Higgins. 2004. British industry in the interwar years. In *The Cambridge economic history of modern Britain*, ed. R. Floud and P. Johnson, 374–402. Cambridge: Cambridge University Press.

Broadberry, S. 1997. The Long Run Growth and Productivity Performance of the United Kingdom. *Scottish Journal of Political Economy* 44 (4):403–424. https://doi.org/10.1111/1467-9485.00067

Broadberry, S., and P. Howlett. 2004. Blood, sweat, and tears: British mobilization for World War II. In *A world at total war. Global conflict and the politics of destruction, 1937–1945*, ed. R. Chickering, S. Förster, and B. Greiner, 157–76. Cambridge: Cambridge University Press.

Capie, F., and M. Collins. 1980. The extent of British economic recovery in the 1930s. *Economy and History* 23 (1):40–60. doi: 10.1080/00708852.1980.10418969.

Caunce, S. A. 2012. The hiring fairs of Northern England, 1890—1930: A regional analysis of commercial and social networking in agriculture. *Past & Present* 217 (1):213–46. doi: 10.1093/pastj/gts024.

Crafts, N., and A. Mulatu. 2005. What explains the location of industry in Britain, 1871–1931? *Journal of Economic Geography* 5 (4):499–518. doi: 10.1093/jeg/lbh069.

De Bromhead, A., A. Fernihough, M. Lampe, and K. H. O'Rourke. 2017. When Britain turned inward: Protection and the shift towards empire in interwar Britain (discussion paper, no. 11835). Centre for Economic Policy Research, February.

Eichengreen, B. 1986. Unemployment in interwar Britain: New evidence from London. *Journal of Interdisciplinary History* 17 (2):335–58. doi: 10.2307/204769.

French, D. 2000. *Raising Churchill's Army: The British Army and the war against Germany 1919–1945*. Oxford: Oxford University Press.

Goldin, C., and L. F. Katz. 2008. *The race between education and technology*. Cambridge, MA: Harvard University.

Goschin, Z., D. L. Constantin, M. Roman, and B. Ileanu. 2009. Specialisation and concentration patterns in the Romanian economy. *Journal of Applied Quantitative Methods* 4 (1):95–111.

Great Britain Historical GIS Project. 2017. 'Great Britain Historical GIS'. University of Portsmouth. https://www.visionofbritain.org.uk/about/credits/4

Heim, C. E. 1983. Industrial organization and regional development in interwar Britain. *The Journal of Economic History* 43 (4):931–52. doi: 10.1017/S0022050700030825.

Henderson, V., A. Kuncoro, and M. Turner. 1995. Industrial development in cities. *Journal of Political Economy* 103 (5):1067–90. doi: 10.1086/262013.

Keibek, S. A. J. 2017. The male occupational structure of England and Wales, 1600–1850. Doctoral dissertation. University of Cambridge. https://www.repository.cam.ac.uk/handle/1810/263608.

Kim, S. 1998. Economic Integration and Convergence: U.S. Regions, 1840–1987. *The Journal of Economic History* 58 (3):659–83. doi: 10.1017/S0022050700021112.

Krugman, P., and A. J. Venables. 1995. The seamless world: A spatial model of international specialization (discussion paper, no. 1230). Centre for Economic Policy Research, August.

Lee, C. H. 1979. *British regional employment statistics, 1841–1971*. Cambridge: Cambridge University Press.

Mitchell, B. R. 1988. *British historical statistics*. Cambridge: Cambridge University Press.

Mokyr, J. 1999. The second industrial revolution, 1870 – 1914. In *Storia dell'economia Mondiale*, ed. V. Castronono, 219–45. Rome: Laterza Publishing

Mukherjee, A. 2010. Empire: How colonial India made modern Britain. *Economic and Political Weekly* 45 (50): 73–82.

Murphy, H. 2018. An anatomy of speculative failure: WM Doxford & Sons Ltd, Sunderland, and the Northumberland Shipbuilding Company of Howdon on Tyne, 1919–1945. *The Mariner's Mirror* 104 (1):58–72. doi: 10.1080/00253359.2018.1411101.

Scott, P. 2002. British railways and the challenge from road haulage: 1919–1939. *Twentieth Century British History* 13 (2):101–20. doi: 10.1093/tcbh/13.2.101.

Shaw-Taylor, L. 2009. The occupational structure of England and Wales, c.1750–1911. Paper prepared for the INCHOS Workshop Held in Cambridge, July 29th-31st 2009. Working paper. https://www.campop.geog.cam.ac.uk/research/projects/occupations/britain19c/papers/paper5.pdf)

Shaw-Taylor, L., and E. A. Wrigley. 2006. The occupational structure of England, c.1750-1851: A preliminary report. Paper presented at the XIV International Economic History Congress, Helsinki Finland.

Sunley, P. 1992. Marshallian industrial districts: The case of the Lancashire cotton industry in the inter-war years. *Transactions of the Institute of British Geographers* 17 (3): 306–20. doi: 10.2307/622882.

Thomas, M. 2004. The service sector. In *The Cambridge economic history of modern Britain, volume II: Economic maturity, 1860–1939*, ed. R. Floud and P. Johnson, 99–132. Cambridge: Cambridge University Press.

Wolf, N. 2007. Endowments vs. market potential: What explains the relocation of industry after the Polish reunification in 1918? *Explorations in Economic History* 44 (1): 22–42. doi: 10.1016/j.eeh.2005.08.003.

Wrigley, E. A. 2010. *The PST system of classifying occupations*. Unpublished paper, Cambridge Group for the History of Population and Social Structure, University of Cambridge. https://www.campop.geog.cam.ac.uk/research/projects/occupations/britain19c/papers/paper1.pdf