Online Appendix for "The Distributional Consequences of Trade: Evidence from the Grain Invasion" (Not for Publication)

Stephan Heblich[*]
University of Toronto and NBER

Stephen J. Redding[†]
Princeton University, NBER and CEPR

Yanos Zylberberg[‡]
University of Bristol and CEPR

September 2024

Table of Contents

[*]Dept. of Economics, Max Gluskin House, 150 St. George Street, Toronto, Ontario, M5S 3G7, Canada. Email: stephan.heblich@utoronto.ca.
[†]Dept. Economics and SPIA, JRR Building, Princeton, NJ 08544, USA. Email: reddings@princeton.edu.
[‡]The Priory Road Complex, 12 Priory Road, Bristol, BS8 1TU, UK. Email: yanos.zylberberg@bristol.ac.uk.

A Introduction

This Online Appendix contains theoretical derivations, supplementary empirical results, and further information on the data sources and definitions. Section B reports additional evidence that supplements the main empirical results from Section 4 of the paper. Section C includes the derivation of all theoretical results from Section 5 of the paper. Section D provides further details on the quantitative analysis of the model from Section 6 of the paper, including the parameterization of the model and our counterfactuals for the Grain Invasion. Section E presents extensions of our theoretical model. Section F provides further information on the data sources and definitions for our parish-level data. Section G gives further details on the matching of individuals across census years using our individual-level population census data.

B Reduced-Form Evidence

This section of the Online Appendix reports additional reduced-form evidence, supplementing the empirical results reported in Sections 2 and 4 of the paper.

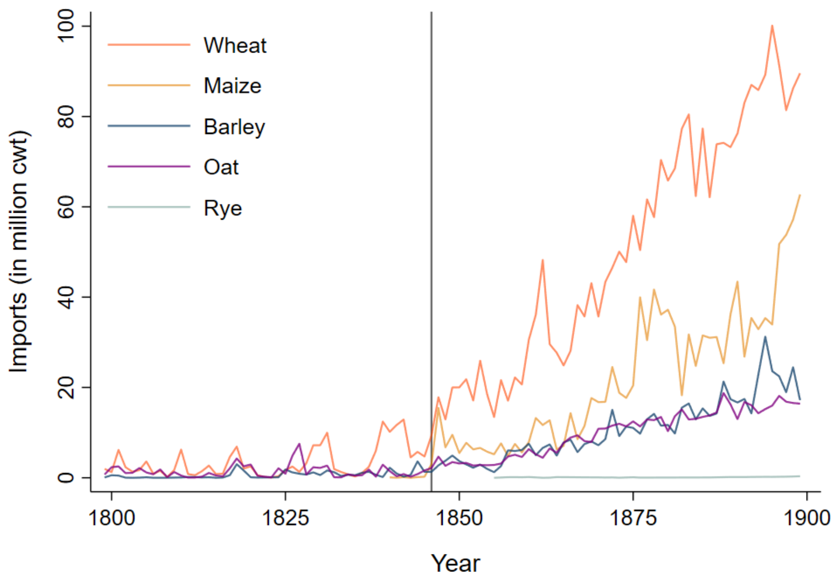
B1 The Grain Invasion

We begin by providing further evidence on the large-scale trade shock from the Repeal of the Corn Laws and the Grain Invasion, augmenting the empirical results from Section 2 of the paper.

The Repeal of the Corn Laws removed the protection that domestic producers of wheat, oats and barley enjoyed against imported grain. We provide descriptive statistics in Section 2 of the paper showing that the price of wheat gradually declined in the second half of the 19th century following the Repeal of the Corn Laws (see Figure 1 in the paper). The integration of wheat markets between the United States and Britain implies that, by 1913, there was no difference in the price of wheat between the Chicago market and the standard Gazette prices for England and Wales (O'Rourke 1997).

While we focus on wheat in our baseline analysis—as it represents the lion's share of domestic production affected by the Grain Invasion, Figure B.1 shows that there is a similar 900% growth in grain imports for wheat (from 10 million cwt in 1846 to about 90 million in 1899), maize (from 5-6 million cwt in 1846 to about 60 million in 1899), barley (from 2 million cwt in 1846 to about 20 million in 1899) and oat (from 2 million cwt in 1846 to about 20 million in 1899).

Figure B.1: Grain Imports, 1799–1899, Measured in Million Hundredweights (cwt)



Note: This Figure displays the evolution of grain imports between 1799 and 1899. As apparent, there is a marked increase in wheat and maize imports after 1846 and 1870.

B2 Wheat Suitability

In this subsection of the Online Appendix, we present additional evidence on the differences in agro-climatic suitability for the cultivation of wheat across parishes in England and Wales, supplementing the evidence reported in Section 4.1 of the paper.

B2.1 Grain and Grazing Regions

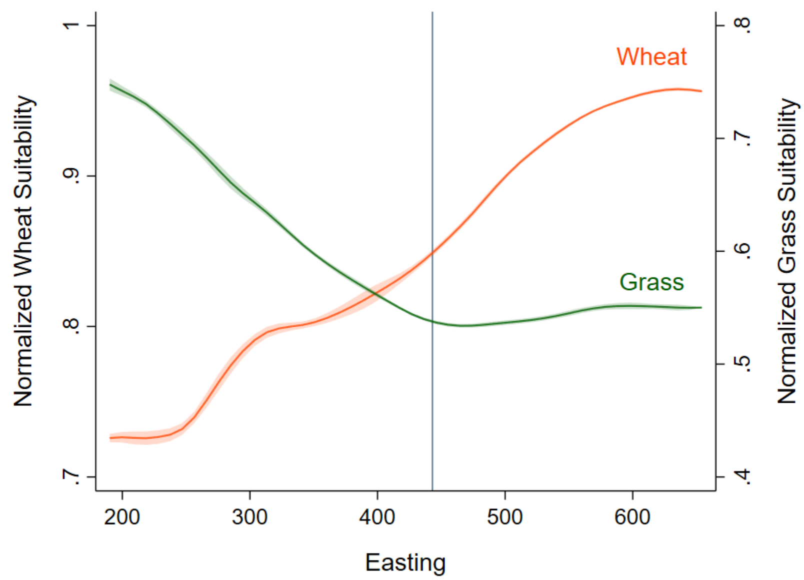
We use climatic conditions and soil suitability in England and Wales as an exogenous source of variation in a region's exposure to wheat imports. The western part of these countries is exposed to winds from the Atlantic and experiences more rain. This is more conducive for pastoral farming while the east, and especially the south-east, is the driest part in Britain and thus better suited for arable farming. This geographic divide was discussed by Cobbett (1830) in his popular book "Rural Rides" and Caird (1852) formally distinguished arable and pastoral farming counties on a map.

We illustrate this East-West gradient in suitability for arable/pastoral farming in Figure B.2, where we plot our measures of wheat and grass suitability (FAO-GAEZ), normalized to lie between zero and one, against parish eastings.[1] The eastern part of Britain is markedly more

[1] We collect a map of expected yield (dry matter per ha) in wheat farming and grass cultivation from the FAO-GAEZ database which we overlay with a map of consistent parishes across England and Wales. To reflect historical

suitable for wheat production and arable farming while the Western part is more suitable for pastoral farming—the *Caird line* (whose average easting is represented on Figure B.2 as a blue line) is relatively successful in dividing these two agro-climatic regions. Note that the hump in wheat suitability between eastings of 300 and 400 is mostly due to the agricultural suitability of Northern England, which is narrower than the Southern parts of England and Wales.

Figure B.2: East-West Pattern in Wheat Suitability



Note: this Figure shows kernel-weighted local polynomial regressions of a measure of normalized wheat suitability (left scale) and grass suitability (right panel) on the easting of a parish centroid (in km). The red line shows the average easting of points located every 10 kms along the Caird line that separates corn and grazing counties following Caird (1852).

B3 Wheat Suitability and Agricultural Production

In this subsection of the Online Appendix, we provide additional empirical validation for our measure of agro-climatic wheat suitability used in Section 4 of the paper, by showing that it is strongly related to measures of wheat production.

We collect data on actual agricultural production from the 1836 Tithe surveys. Tithe is an old English term meaning tenth and referred to an annual tax of one tenth on all agricultural produce. With the 1836 Tithe Commutation Act, Tithe payments were commuted into monetary payments, so called Tithe rents. The commutation affected parishes that were not extensively enclosed, and

growing conditions, we choose a specification where crops are rain-fed and grown with low input intensity. We normalize the measures by dividing them by the maximum yield in the sample of parishes. Eastings are projected coordinates from the British National Grid, here expressed in kilometers.

the payments were determined by the Tithe Survey conducted between 1837–1855 (90 percent was completed by 1845). This survey produced three documents for all Tithe districts: a map, an apportionment, and a file. Kain (1986) used this information to recover the pattern of land use and agricultural production. This provides the most complete picture of the rural landscape of England and Wales in mid-nineteenth century – a period of "High Farming" – before the fall in the relative price of wheat from the Grain Invasion.

There were 14,829 Tithe districts in England and Wales and, for 6,726 of these districts, surveyors' reports are still extant in the National Archives. These files were digitized by Kain (1986) and subsequently linked to the Cambridge Group for the History of Population and Social Structure's parish boundaries by Max Satchell.[2] Figure B.3 summarizes the information from the Tithe surveys on the share of acreage under arable farming for one third of our consistent parishes (3,642 parishes, panel a). While the Tithe districts are not randomly allocated, they are evenly distributed across space which allows us to compute an interpolated measure (see panel b).

Figure B.3: Share of Acreage under Arable (1836 Tithe Maps)

(a) Raw Data

(b) Interpolated Data

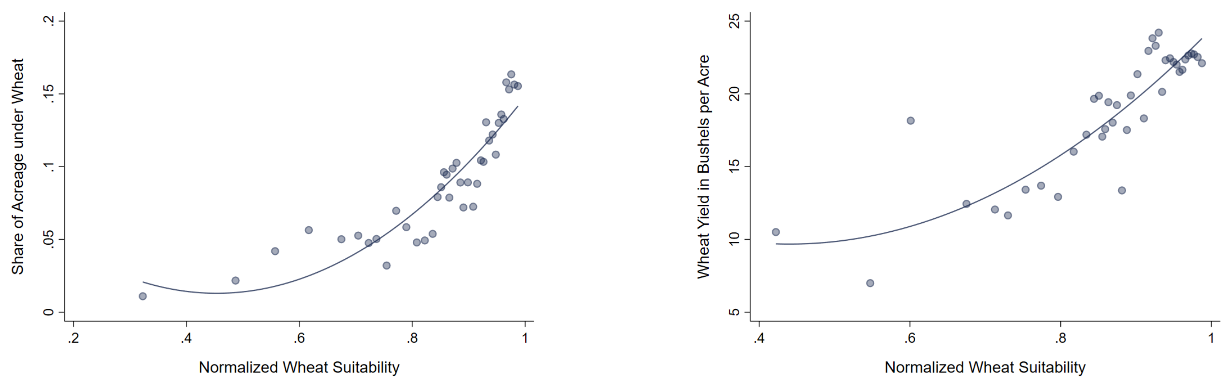
Note: Panel (a) displays the share of acreage under arable for 3,642 consistent parishes which can be mapped to existing Tithe reports from 1836; Panel (b) displays the interpolated values for England and Wales (using natural neighbor interpolation).

The disparity in arable and pastoral farming in England and Wales is mostly explained by exogenous agro-climatic conditions. Figure B.4 correlates the FAO suitability measure with the

[2] See https://www.campop.geog.cam.ac.uk/research/occupations/datasets/catalogues/documentation/ for a more detailed documentation of this geography.

share of acreage under wheat and the wheat yield in bushels of wheat per acre in tithe districts where we observe agricultural production. A quadratic function provides a good description of the relationship between wheat suitability and the share of acreage under wheat. It reflects the prominent role of the South-East as the wheat-growing region. The wheat yield (in bushels per acre) and suitability are best described by a linear relationship.

Figure B.4: Wheat Production, Acreage and Suitability



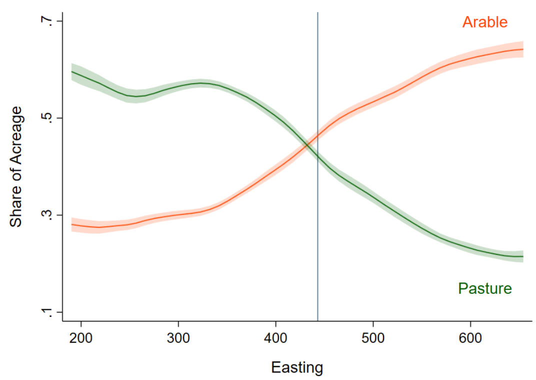
(a) Acreage Under Wheat

(b) Wheat Yield (Bushels/Acre)

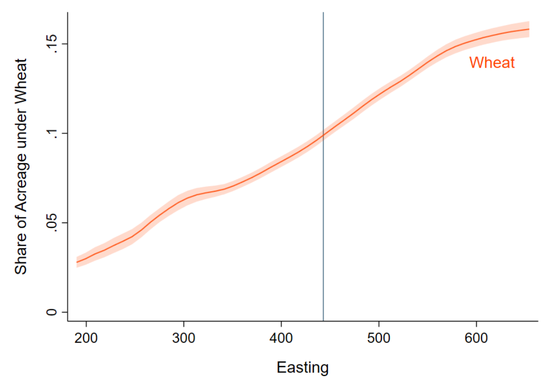
Note: This Figure shows bin scatter plots for the share of acreage under wheat (left panel) and the predicted yield for wheat (right panel) within the consistent parish units that overlap with the available 1836 Tithe districts. The local polynomial regressions are weighted by the parish area.

The East-West gradient in agro-climatic conditions thus translates into a strong, marked East-West gradient in arable and pastoral farming. We illustrate these patterns in Figure B.5. The share of acreage used for arable farming is increasing as we move to the East, while the share of pasture (measured as grassland) is decreasing. Upon closer inspection, this pattern is mainly driven by the shares of wheat and barley, which increase substantially as we move eastwards, while oats are more uniformly grown across England and Wales. This disparity in agricultural production induces a differential exposure to the Grain Invasion, as we document in the next section.

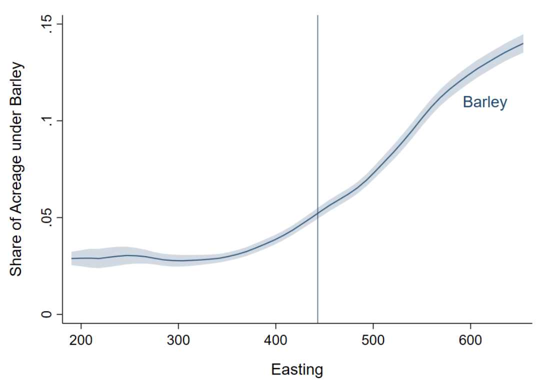
Figure B.5: East-West Gradient in Agricultural Production



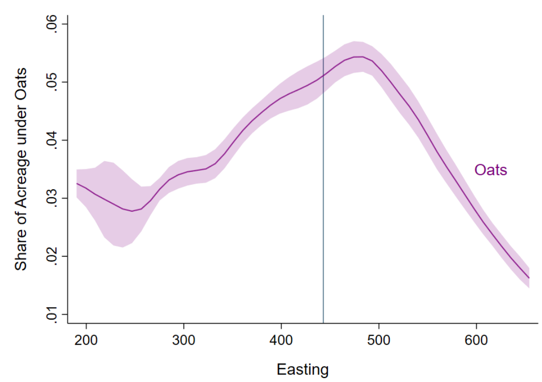
(a) Arable Versus Pasture



(b) Acreage Under Wheat



(c) Acreage Under Barley



(d) Acreage Under Oats

Note: This Figure shows the share of acreage under arable/pasture (top left), wheat (top right), barley (bottom left) and oats (bottom right) within the consistent parish units that overlap with the available 1836 Tithe districts.

B4 Structural Transformation

We next provide further evidence on structural transformation in areas of high and low-wheat suitability over time, supplementing the evidence reported in Section 4.2 of the paper.

B4.1 Agricultural employment after 1851

The East-West gradient in cropping patterns is reflected in the dynamics of agricultural employment after the Repeal of the Corn Laws and the Grain Invasion. The left panel of Figure B.6 displays the share of farmers and laborers (farm workers) in 1851, and its evolution over time. In the west, where pastoral farming is the dominant type of agriculture, we observe equal shares of laborers and farmers. As we move east, we see a slight decrease in the share of farmers and a steep increase in the share of laborers. This observation matches figures reported in Perry (1973), suggesting that the average farm in grazing countries was smaller and employed substantially fewer laborers.

Comparing 1851 and 1901, we see that the share of farmers is more or less constant while the share of laborers decreases as we move from West to East. This was a consequence of the structural change induced by the Grain Invasion that hit wheat growing regions particularly hard. Decreasing labor demand was making agricultural laborers' wages uncertain and precarious, while towns and mining areas allured with industrial wages that were 50 percent higher than agricultural wages (Bellerby 1953). In an attempt to escape their unattractive and dull country life, laborers often fell for "the attraction of city life, the continuous and increasing demand of urban industries, and the better wages and other enjoyments offered", leading to a report on *The Decline in The Agricultural Population of Great Britain* (Board of Agriculture and Fisheries 1906).

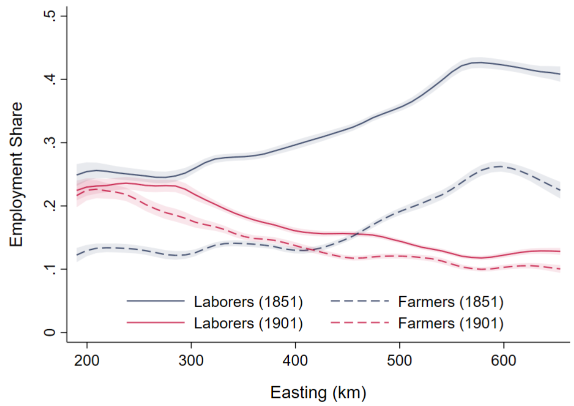
The right panel of Figure B.6 displays the ratio of laborers to farmers in 1851 and 1901. In both years, there is an upward-sloping West-East gradient. This gradient becomes noticeably flatter over time, consistent with the Grain Invasion having a greater impact on Eastern locations where agro-climatic conditions were more favorable for labor-intensive arable farming.

As a result of the Grain Invasion, the share of arable land shrank over the period from 1871–1901 by 29 percent from 8.2 to 5.9 million acres while the area of permanent pasture experienced a 36 percent increase from 11.4 to 15.4 million acres (Lord Ernle 1912). Looking at the distribution of land shares within arable farming, we observe that wheat production was often replaced by oats while the acreage share of barley and other crops remains largely unchanged (see Figure B.7).

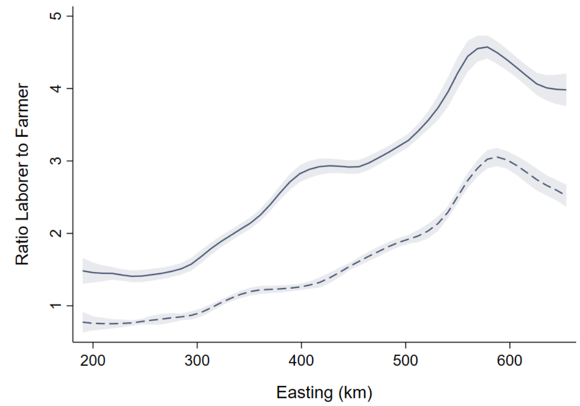
B4.2 Sectoral Employment Shares and Wheat Suitability

We now move to a more refined measure of exposure to the Grain Invasion based on agro-climatic conditions. Figure B.8 presents evidence on the evolution of sectoral employment shares over the

Figure B.6: East-West Gradient in Sectoral Employment Shares



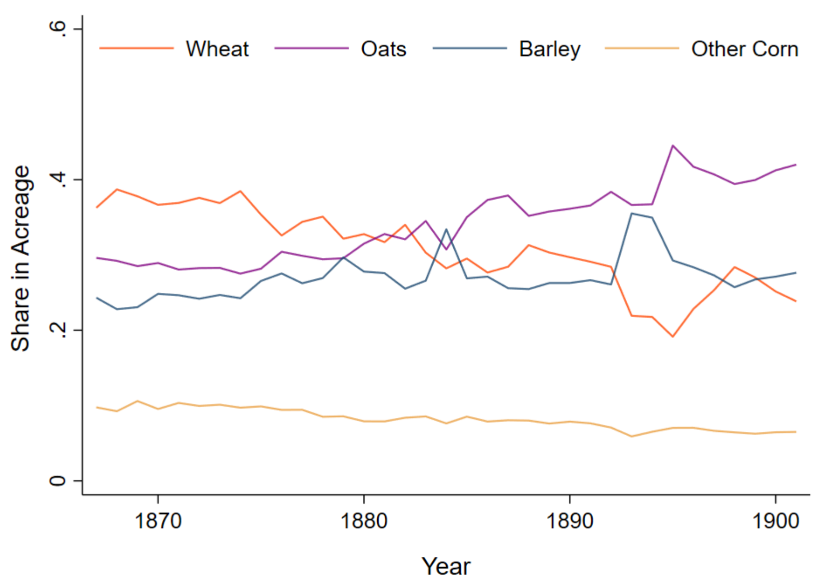
(a) Laborers and Farmers in 1851



(b) Laborers and Farmers Over Time

Note: The left panel shows kernel-weighted local polynomial regressions of the employment shares of laborers and farmers in 1851 and 1901 on eastings (in km). The right panel shows kernel-weighted local polynomial regressions of the ratio between laborers and farmers in 1851 and 1901 on eastings (in km).

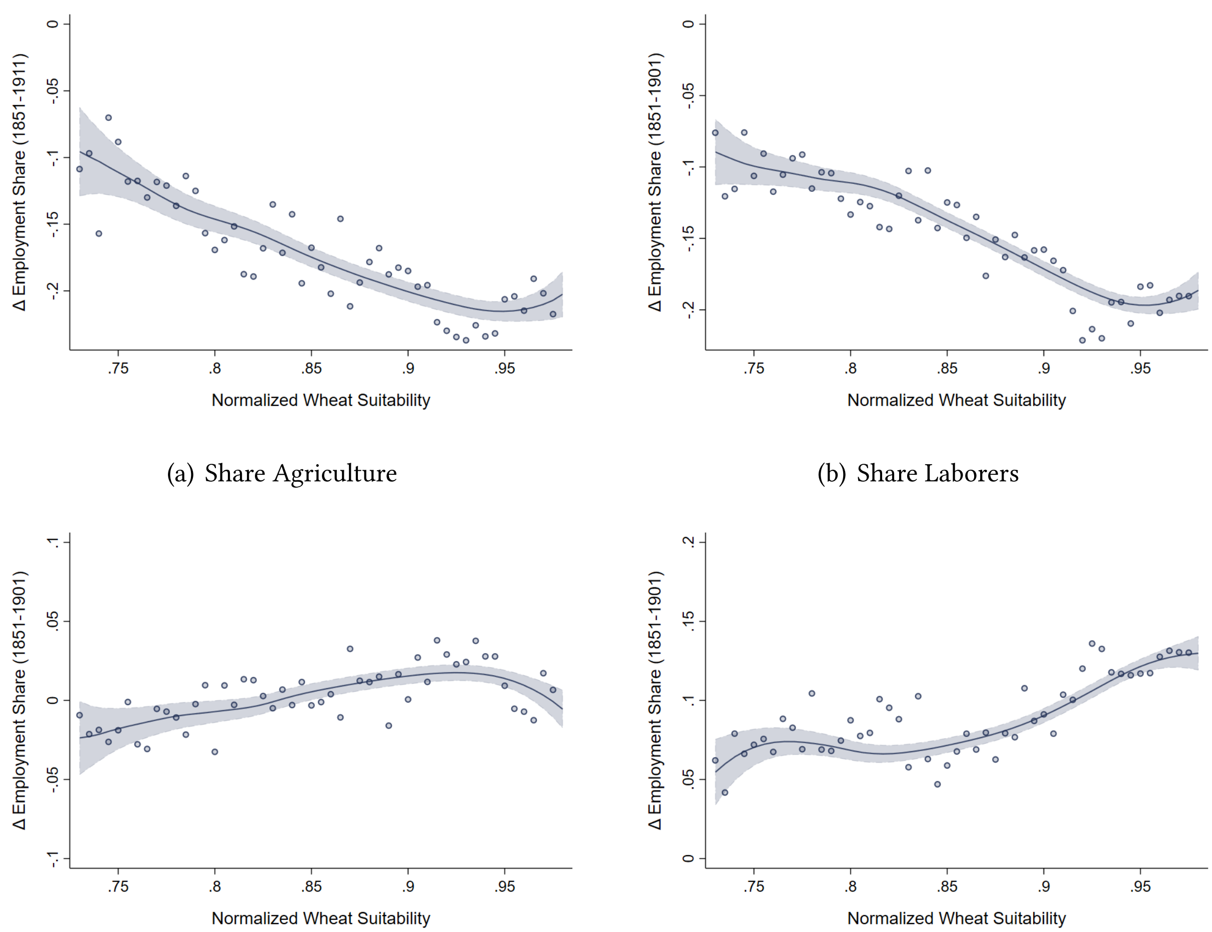
Figure B.7: Share of Acreage for Different Types of Cereal, 1867–1911



Note: This Figure uses data from Mitchell (1988) to show the division of land used for grain production into the share used for wheat, oats, barley and other crops.

period 1851–1911 for regions with different agro-climatic suitability to grow wheat—and thus different exposure to the Grain Invasion.

Figure B.8: Sectoral Employment Shares and Wheat Suitability over Time



(a) Share Agriculture

(b) Share Laborers

(c) Share Manufacturing

(d) Share Services

Note: The figure shows kernel-weighted local polynomial regressions of different sectoral and subsectoral employment shares or the ratio between agricultural laborers and farmers as outcome (bottom-right panel) on the normalized FAO measure of soil suitability for the years 1851 and 1901.

We find substantial structural change over this 50-year interval, with the employment share of the service sector rising at the the expense of both, agriculture and manufacturing. There is a stronger decline in the agricultural employment share in regions with greater wheat suitability. Within the agricultural sector, we find evidence that regions with greater wheat suitability experience a stronger decline in the share of laborers, which is in line with the notion that arable farming is more labor intensive.

B4.3 Farm-level Evidence

In this subsection, we provide additional evidence on the adjustment of farm production at the parish level to the Grain Invasion using individual-level census data from 1851–1881.

The Integrated Census Microdata Project (I-CeM) reports occupational codes which do not clearly disaggregate farming production. The occupational code however summarizes the occupation reported by respondents as a textual description, and this original textual description is quite rich for farmers in 1851, 1861 and 1881: for instance, the occupation of a land owner would be described as "FARMER OF 200 ACRES EMPLOYG 7 LABR 3 BOYS", "FARMER OF 56 ACRES EMP NO LAB" or "FARMER OF 98 AC EPL 1 LABR". We use a systematic textual analysis on this occupational description to extract consistent indicators of farm size (number of acres, and number of men/women/boys employed); generate indicators of average farm size (average number of acres, and average number of men/women/boys employed); generate indicators of land ownership concentration (Herfindahl indices based on acres and total labor).

There is a significant increase in the concentration of land over England and Wales, coupled with a decrease in the labor intensity of farming between 1851 and 1881: the median farm size in rural parishes (see Section F4 of this Online Appendix for a formal description) goes from 130 acres and 3.06 men employed in 1851, to 154 acres and 2.63 men employed in 1881. Farming is thus 32% less labor intensive in 1881 than in 1851. The concentration of land ownership increases markedly: the median Herfindahl indices at the parish level go from 0.09 in 1851, whether computed with acres or employed workers, to 0.18-0.19 in 1881.

We now examine the relationship between the change in farming production and wheat suitability using the following regression specification:

$$F_{jt} = \sum_{t=1851}^{t=1881} \beta_t \left(\mathbb{I}_t \times \text{Wheat}_j \right) + \sum_{t=1851}^{t=1881} X_j \delta_t + \eta_j + d_t + u_{jt}, \tag{B.1}$$

where the unit of observation is a rural parish j at time $t = 1851, \ldots, 1881$; $\mathbb{W}_j$ is an indicator variable that is one if a parish has above-median wheat suitability and zero otherwise; η_j are parish fixed effects; d_t are time fixed effects; and $X_j \delta_t$ are geographic controls interacted with time dummies, which include travel time to the nearest market town; travel time to the nearest coalfield; distance to London and distance to Manchester; Easting and Northing of the parish centroid; an indicator that is one for Wales; an indicator that is one for urban parishes based on the 1801 distribution of population densities; the coefficients β_t capture differential changes in farming production in high versus low-wheat-suitability parishes.

As shown in Table B.1, we find that farming becomes more concentrated among fewer land owners in the regions most affected by the Grain Invasion. First, we find that the population share of farm owners (column 1) and helpers within the same family (column 2) decreases by 1.8

percentage points between 1851 and 1861, and by an additional 1.4 percentage points between 1861 and 1881. In parallel, the share of agricultural laborers increases by 2.4 percentage points between 1851 and 1861, but decreases by about 0.8 percentage points between 1861 and 1881. Second, we focus on larger farms for which farm size is reported by Census surveyors: The average farm in parishes with high wheat-suitability is 74 acres larger in 1881 than in 1851, compared to parishes with low wheat-suitability (see column 4), and employ 4.4 additional agricultural laborers (see column 5). As a result, land ownership becomes much more concentrated in parishes with high wheat-suitability: the relative increase in Herfindahl indices is between .06 and .11—an effect between 30% and 50% of the standard deviation in Herfindahl indices across parishes at baseline (see columns 6 and 7).

Table B.1: Wheat Suitability and Farming Production

	Employment share			Large farms		Land concentration	
	Farmers	Helpers	Laborers	Acres	Employment	Acres	Employment
Wheat suitability × 1861	-.0032	-.0048	.0171	26.53	3.303	.0636	.0423
	(.0011)	(.0020)	(.0038)	(14.02)	(0.608)	(.0155)	(.0144)
Wheat suitability × 1881	-.0044	-.0154	.0152	40.06	2.704	.0708	.0480
	(.0014)	(.0025)	(.0030)	(14.78)	(0.622)	(.0163)	(.0142)
Observations	19,638	19,638	19,638	16,543	15,281	19,638	19,638

Note: Standard errors are reported between parentheses and are clustered at the parish-level. The unit of observation is a rural parish at a given time t. All specifications condition the analysis on parish fixed effects, time fixed effects and controls for observable parish characteristics interacted with time dummies—see Equation (B.1).

B5 Agricultural Bankruptcies

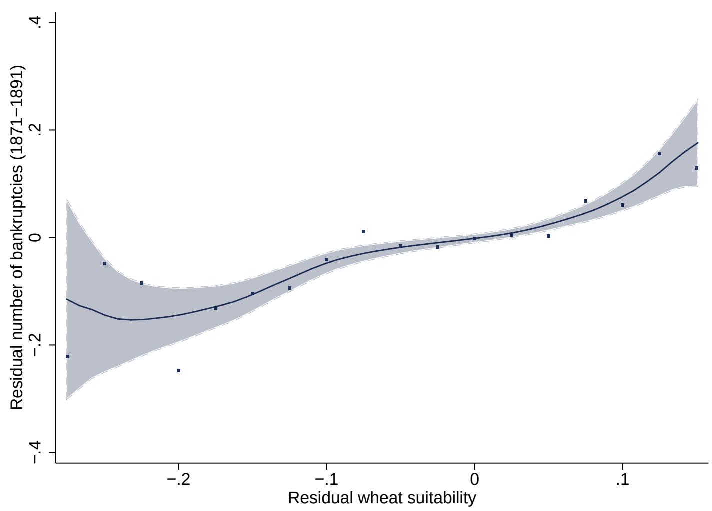
A large historical literature discusses the role of the Grain Invasion in the "great agricultural depression" in Britain after 1870. In this section of the Online Appendix, we use data on agricultural bankruptcies to provide further evidence in support of this role of the Grain Invasion in explaining the great agricultural depression.

Under the Bankruptcy Act of 1869, a notice of bankruptcy was required to be published in the *London Gazette*. This notice reported the occupation and residential address of the bankrupt, thereby allowing agricultural failures to be mapped geographically. We use the data collected by Perry (1972) and displayed in a sequence of maps for the periods 1871-3, 1881-3 and 1891-3. These maps plot individual bankruptcies as points within the boundaries of England and Wales. We first georeferenced these maps and geocoded each bankruptcy. We next assign bankruptcies to parishes based on their latitude and longitude coordinates. Finally, for each parish, we compute (i) The number of agricultural bankruptcies; (ii) The number of agricultural bankruptcies relative to geographical land area; (iii) The number of agricultural bankruptcies per farmer. We pool the data for all three time periods 1871-3, 1881-3 and 1891-3.

We use the Frisch-Waugh-Lovell Theorem to flexibly examine the relationship between agricultural bankruptcies and wheat suitability. First, we regress bankruptcies on a number of controls for parish characteristics and generate the residuals. Second, we regress wheat suitability on the same controls and generate the residuals. Third, we estimate a non-parametric regression of residual bankruptcies on residual wheat suitability.

We use the same set of controls for parish characteristics as for our event-study specification in Section 4.4 of the paper: (i) travel time to the nearest market town (to control for access to urban centers); (ii) travel time to the nearest coalfield (to capture access to coal as a natural resource); (iii) Easting and Northing of the parish centroid (to control for geographical location); (iv) an indicator that is one for Wales (to allow for differences in economic growth between England and Wales); (v) distance to London and distance to Manchester (to capture proximity to these two concentrations of urban population); (vi) an indicator that is one for urban parishes based on the 1801 distribution of population densities (to allow for differences in economic growth between urban and rural areas).

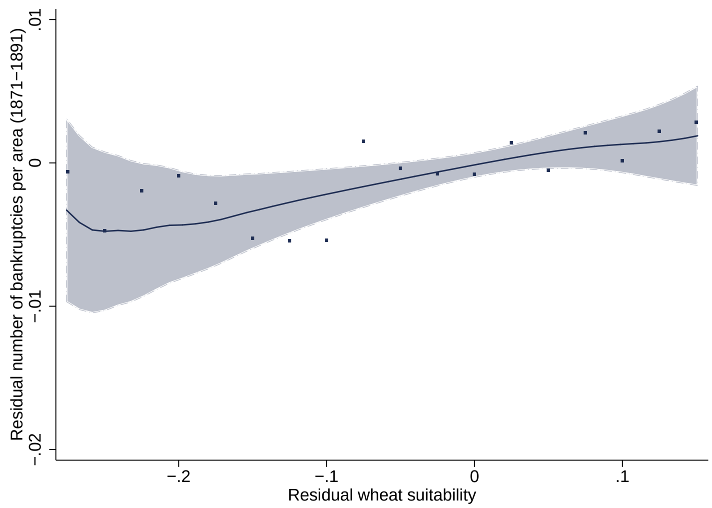
Figure B.9: Number of Agricultural Bankruptcies and Wheat Suitability



Note: The figure shows kernel-weighted local polynomial regressions of the residual number of agricultural bankruptcies (during 1871-3, 1881-3 and 1891-3) on residual wheat suitability; residuals control for (i) travel time to the nearest market town; (ii) travel time to the nearest coalfield; (iii) distance to London and distance to Manchester; (iv) Easting and Northing of the parish centroid; (v) an indicator that is one for Wales; (vi) an indicator that is one for urban parishes based on the 1801 distribution of population densities.

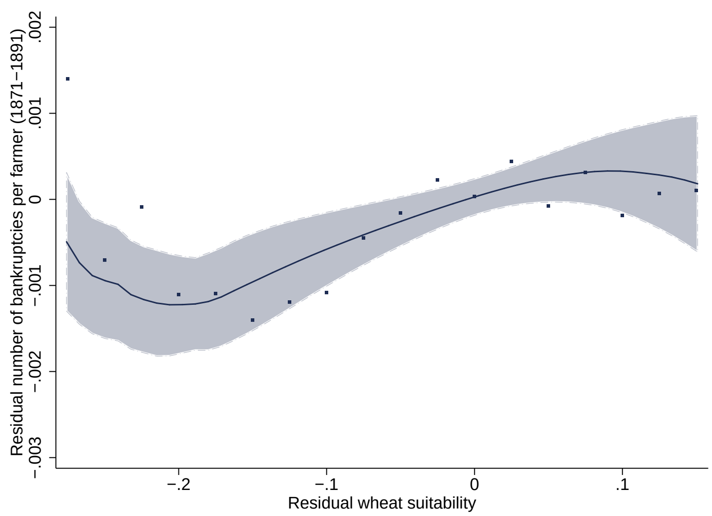
In Figure B.9, we display a kernel-weighted local polynomial regression of the residual number of bankruptcies on residual wheat suitability. In Figure B.10, we show the corresponding regression of the residual number of bankruptcies relative to geographical land area on residual

Figure B.10: Number of Agricultural Bankruptcies Relative to Land Area and Wheat Suitability



Note: The figure shows kernel-weighted local polynomial regressions of the residual number of agricultural bankruptcies relative to land area (during 1871-3, 1881-3 and 1891-3) on residual wheat suitability; residuals control for (i) travel time to the nearest market town; (ii) travel time to the nearest coalfield; (iii) distance to London and distance to Manchester; (iv) Easting and Northing of the parish centroid; (v) an indicator that is one for Wales; (vi) an indicator that is one for urban parishes based on the 1801 distribution of population densities.

Figure B.11: Number of Agricultural Bankruptcies per Farmer and Wheat Suitability



Note: The figure shows kernel-weighted local polynomial regressions of the residual number of agricultural bankruptcies per farmer (during 1871-3, 1881-3 and 1891-3) on residual wheat suitability; residuals control for (i) travel time to the nearest market town; (ii) travel time to the nearest coalfield; (iii) distance to London and distance to Manchester; (iv) Easting and Northing of the parish centroid; (v) an indicator that is one for Wales; (vi) an indicator that is one for urban parishes based on the 1801 distribution of population densities.

wheat suitability. In Figure B.11, we display the analogous regression of the residual number of bankruptcies per farmer on residual wheat suitability. Using all three bankruptcy measures, we find more agricultural bankruptcies in parishes with greater wheat suitability. We find that this relationship holds both conditional on our full set of controls for parish characteristics (as shown in the figures) and unconditionally.

Taken together, these findings provide further support of the idea that the Grain Invasion led to an agricultural depression that was more severe in the parts of England and Wales with agroclimatic conditions more favorable for wheat cultivation.

B6 Event-Study Robustness

In this subsection of the Online Appendix, we provide additional event-study evidence on the impact of the Grain Invasion, supplementing the results reported in Subsection 4.4 of the paper.

B6.1 Robustness to Different Sets of Controls

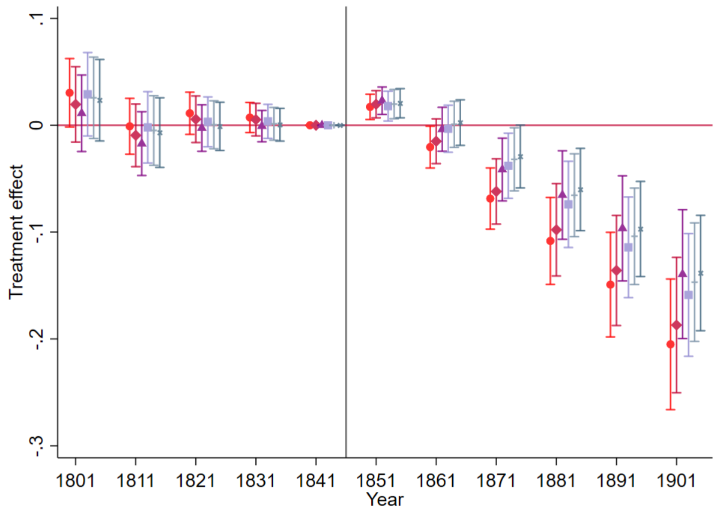
Figure 7 in the paper reports our baseline event-study specification, including all our controls for observable parish characteristics (X_j) interacted with census year dummies. Figure B.12 demonstrates the robustness of these empirical findings to different sets of controls. Across each specification, we find the same pattern of results. In the early decades of the 19th century, high and low-wheat-suitability locations have similar population trends, with estimated coefficients that are close to zero and statistically insignificant. After the Repeal of the Corn Laws and the Grain Invasion in the second half of the 19th century, we observe a decline in population in locations with high wheat suitability relative to those with low wheat suitability.

B6.2 Robustness to Alternative Exposure Measures

An advantage of our empirical setting is that we have an exogenous measure of exposure to the trade shock based on agro-climatic suitability for the cultivation of wheat. In our baseline event-study specification in Figure 7 in the paper, we use an indicator variable for above and below-median wheat suitability. In this section of the Online Appendix, we demonstrate the robustness of our results to the use of alternative specifications.

First, we re-estimate our event-study specification using indicators for low-wheat suitability (bottom tercile) and high-wheat suitability (top tercile) interacted with census-year dummies, where the excluded category is the middle tercile. As shown in Figure B.13, in the second half of the 19th century, low-wheat-suitability locations experience a statistically significant *increase* in population relative to those with medium-wheat suitability, while high-wheat-suitability locations experience a statistically significant *decrease* in population relative to those with medium-

Figure B.12: Estimated Treatment Effects of Wheat Suitability for Log Population (Robustness to Different Sets of Controls)

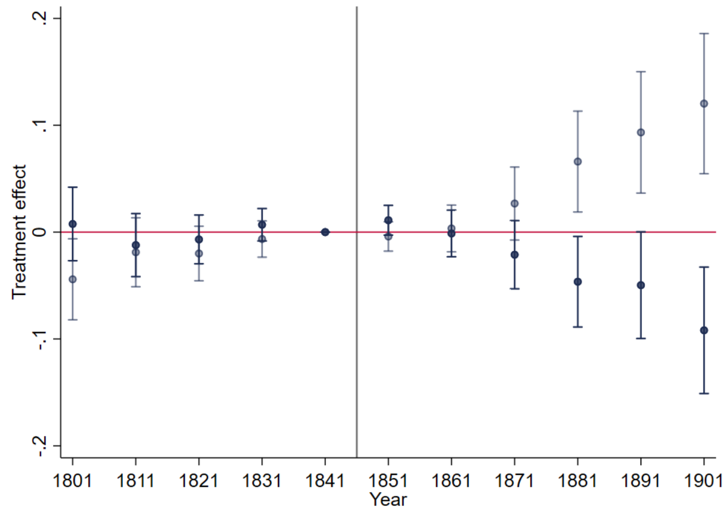


Note: This Figure demonstrates the robustness of our results in Figure 7 in the paper to the inclusion of different sets of time-varying controls. Panel (a) displays the main estimates for our population regression, while panel (b) shows the estimates for rateable values. We consider six different specifications: (i) the first specification includes parish fixed effects as well as travel time to the nearest market town and distance to London and distance to Manchester, all interacted with year fixed effects (red circle); (ii) the second specification adds an indicator that is one for Wales and an indicator that is one for urban parishes based on the 1801 distribution of population densities, all interacted with year fixed effects (dark-red diamond); (iii) the third specification adds travel time to the nearest coalfield interacted with year fixed effects (purple triangle); (iv) the fourth specification adds Easting and Northing of the parish centroid interacted with year fixed effects (lavender square, corresponding to our baseline); (v) the fifth specification adds road density in 1830, railways density in 1851 and waterways density in 1817, all interacted with year fixed effects (light blue tick); and (vi) the last specification adds soil characteristics (share of clay and share of heavy soil) interacted with year fixed effects (darker blue cross).

wheat suitability.

This consistent pattern of results at the top and the bottom of the distribution for wheat suitability provides further support for the idea that our results capture a systematic impact of the Grain Invasion across locations with different levels of wheat suitability.

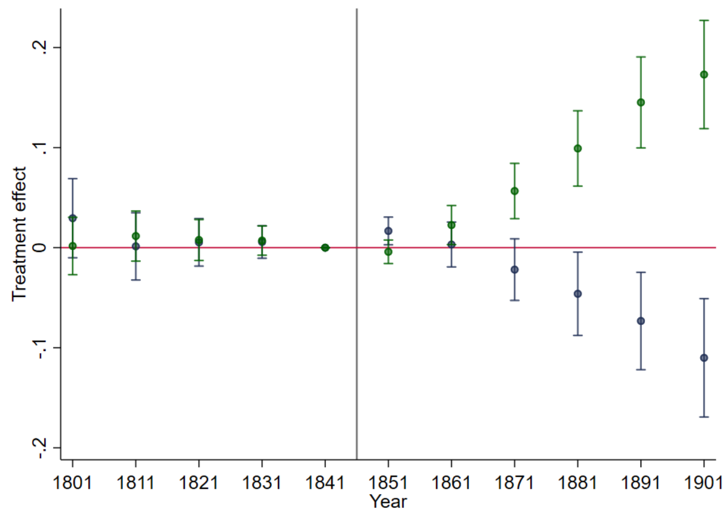
Figure B.13: Estimated Treatment Effects of Wheat Suitability for Log Population (Terciles)



Note: Estimated treatment effects from our event-study specification in equation (1) in the paper using log population, including indicators for parishes in the bottom and top terciles of wheat suitability interacted with census-year dummies (the excluded category is the middle tercile); dark blue circles show the estimated treatment effects for the top tercile of wheat suitability; light blue circles show the estimated treatment effects for the bottom tercile of wheat suitability; vertical lines show 95 percent confidence intervals based on standard errors clustered by registration district. The specification conditions on parish and year fixed effects, and interactions between year and (i) travel time to the nearest market town; (ii) travel time to the nearest coalfield; (iii) distance to London and distance to Manchester; (iv) Easting and Northing of the parish centroid; (v) an indicator that is one for Wales; (vi) an indicator that is one for urban parishes based on the 1801 distribution of population densities.

Second, we re-estimate our event-study specification using above-median wheat suitability, augmenting it with above-median grass suitability interacted with census-year dummies. In Figure B.14, we show the estimated coefficients for wheat suitability (blue) and grass suitability (green). Consistent with our results capturing the Grain Invasion, we find a similar pattern of estimated coefficients for wheat suitability, with negative and statistically significant treatment effects of around the same magnitude as before. The estimated coefficients for grass suitability are positive (rather than negative), which is consistent with a reallocation from arable to pastoral farming, and the systematic differences in agroclimatic conditions between the Western and Eastern parts of England and Wales shown in Figure 3 in the paper. As the Grain Invasion in the second half of the 19th century depressed economic activity in Eastern-corn areas, this increased relative levels of economic activity in Western-grass areas.

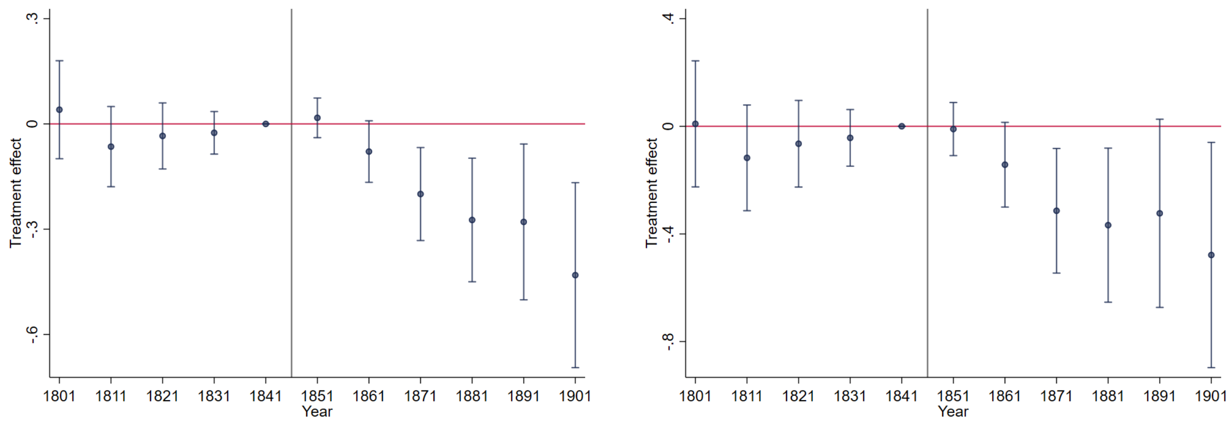
Figure B.14: Estimated Wheat and Grass Suitability Treatment Effects for Log Population (Robustness)



Note: Estimated treatment effects from our event-study specification in equation (1) in the paper using log population, including indicators for above-median wheat and grass suitability interacted with census-year dummies; blue circles show the estimated treatment effects for wheat suitability; green circles show the estimated treatment effects for grass suitability; vertical lines show 95 percent confidence intervals based on standard errors clustered by registration district. The specification conditions on parish and year fixed effects, and interactions between year and (i) travel time to the nearest market town; (ii) travel time to the nearest coalfield; (iii) distance to London and distance to Manchester; (iv) Easting and Northing of the parish centroid; (v) an indicator that is one for Wales; (vi) an indicator that is one for urban parishes based on the 1801 distribution of population densities.

Third, we report a number of further robustness tests for our event-study specification using above-median wheat suitability. Panel (a) of Figure B.15 uses a continuous measure of wheat suitability, instead of an indicator variable for above and below-median wheat suitability. Panel (b) of Figure B.15 measures exposure to the Grain Invasion using the arable land share in the 1836 tithe surveys instrumented with wheat suitability. In both cases, we find a similar pattern of results as in our baseline specification in the paper. In the first half of the 19th century, we find no evidence of differences pre-trends. In contrast, following the Repeal of the Corn Laws and the Grain Invasion in the second half of the 19th century, we find a decline in the relative population of high-wheat-suitability locations.

Figure B.15: Estimated Treatment Effects Using Alternative Exposure Measures



(a) Continuous Wheat Suitability

(b) IV Specification

Note: Panel (a) reports estimated treatment effects from an event-study specification for log population using a continuous measure of wheat suitability ($\mathbb{S}_j$); panel (b) reports estimated treatment effects from an event-study specification in which we measure exposure to the Grain Invasion using the (endogenous) share of arable production in 1836 ($\mathbb{A}_j$) instrumented with our (exogenous) measures of wheat suitability ($\mathbb{W}_j$). In both cases, we include our controls for observable parish characteristics that could affect economic growth interacted with year dummies. We include among these controls: (i) travel time to the nearest market town; (ii) travel time to the nearest coalfield; (iii) distance to London and distance to Manchester; (iv) Easting and Northing of the parish centroid; (v) an indicator that is one for Wales; (vi) an indicator that is one for urban parishes based on the 1801 distribution of population densities. Standard errors are clustered by registration district.

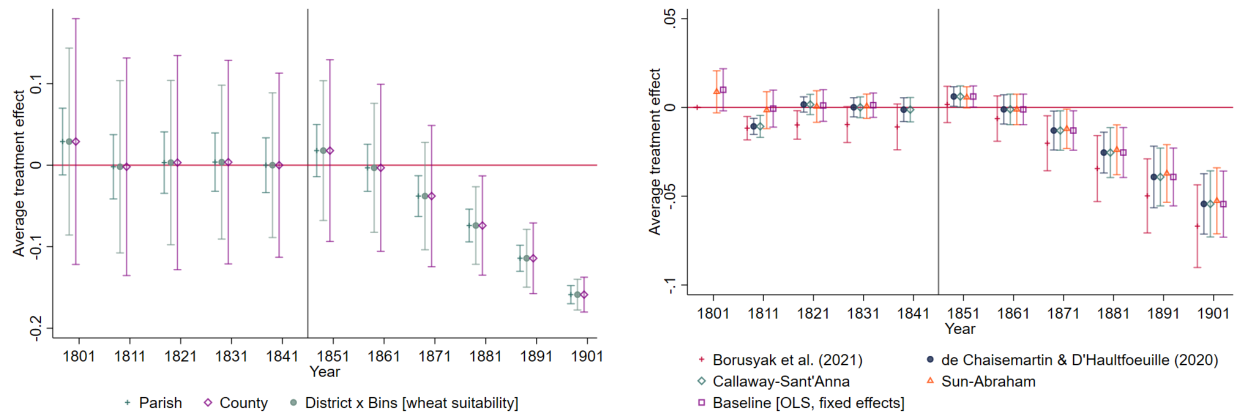
B6.3 Robustness to Alternative Difference-in-Differences Estimators

We next demonstrate the robustness of our results to alternative approaches to computing standard errors and to the use of alternative difference-in-differences estimators.

In our baseline event-study specification in Figure 7 in the paper, we report standard errors clustered by registration district, which allows the regression error to be serially correlated across parishes within registration districts and over time. Panel (a) of Figure B.16 re-estimates this specification using alternative approaches to computing the standard errors. First, we cluster the standard errors by parish (darker green), which only allows the regression error to be serially correlated over time. Second, we double cluster the standard errors by registration district and 100 bins of wheat suitability (lighter green), which allows for serial correlation across parishes within registration districts, over time, and across parishes with similar wheat suitability (following a related approach to Adão, Morales and Kolesár 2019). In each case, we find a similar pattern of results, with a statistically significant decline in the relative population of high-wheat-suitability locations in the second half of the 19th century.

In our baseline specification, we report results using the conventional two-way fixed effects estimator. However, a recent empirical literature has highlighted that the interpretation of this

Figure B.16: Alternative Clustering and Difference-in-Differences Estimators for Log Population



(a) Alternative Clustering

(b) Alternative Estimators

Note: Alternative clustering procedures and alternative difference-in-differences estimators for log population. In panel (a), we report our main results with standard errors clustered by parish (darker green), double clustered by registration district and 100 bins of wheat suitability (lighter green), and clustered by county (purple). In panel (b), we report estimates using the difference-in-differences estimators of Chasemartin and D'Haultfœuille (2020), Callaway and Sant'Anna (2021), Sun and Abraham (2021), and Borusyak, Jaravel and (2024) as well as our baseline estimates using the two-way fixed effects estimator. In all specifications, we condition on controls for observable parish characteristics interacted with census-year dummies. We include among these controls: (i) travel time to the nearest market town; (ii) travel time to the nearest coalfield; (iii) distance to London and distance to Manchester; (iv) Easting and Northing of the parish centroid; (v) an indicator that is one for Wales; (vi) an indicator that is one for urban parishes based on the 1801 distribution of population densities.

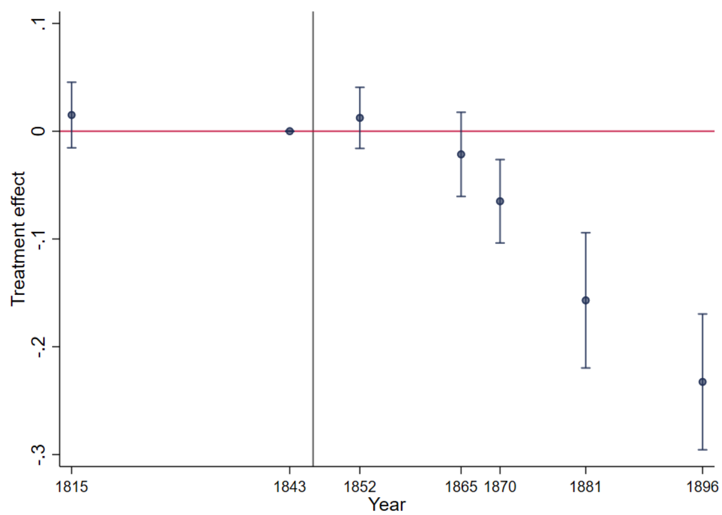
two-way fixed effects estimator can be problematic in the presence of treatment heterogeneity and a variable timing of the treatment. In our empirical application, we have a common timing of the treatment (the Repeal of the Corn Laws and the Grain Invasion). Nevertheless, as a robustness check, Panel (b) of Figure B.16 re-estimates our baseline specification using the alternative event-study estimators of Chaisemartin and D'Haultfœuille (2020), Callaway and Sant'Anna (2021), Sun and Abraham (2021), and Borusyak, Jaravel and (2024). We find the same pattern of results using all of these different estimators as with the two-way fixed effects estimator. In the first half of the 19th century, we find no evidence of differences in pre-trends. Following the Repeal of the Corn Laws and the Grain Invasion in the second half of the 19th century, we find a decline in the relative population of high-wheat-suitability locations.

B6.4 Property Values and Poor Law Relief

We next re-estimate our baseline event-study specification in equation (1) in the paper using property values instead of population. Our property values data measure the market rental value of land and buildings for tax purposes and capture the relative economic value of locations. They are available for the years of 1815, 1843, 1852 and 1881. We choose 1843 as the excluded category, as the last period before the Repeal of the Corn Laws. As in our baseline specification for population in the paper, we include parish fixed effects and year dummies, as well as interactions between our observable parish characteristics and year dummies, to control for other determinants of economic growth.

As shown in Figure B.17, we find a similar pattern of results for property values as for population, with no evidence of pre-trends in the first half of the 19th century, and a decline in relative property values in high-wheat-suitability locations in the second half of the 19th century. We find a somewhat larger estimated treatment effect for property values of around 24 percent by the end of the 19th century. This estimated treatment effect captures the impact of the Grain Invasion on the value of existing buildings and land, as well as on the construction of new buildings. It also captures both the direct effects of this trade shock on property values and its indirect or general equilibrium effects through the reallocation of population across locations.

Figure B.17: Estimated Wheat Suitability Treatment Effects for Log Property Values

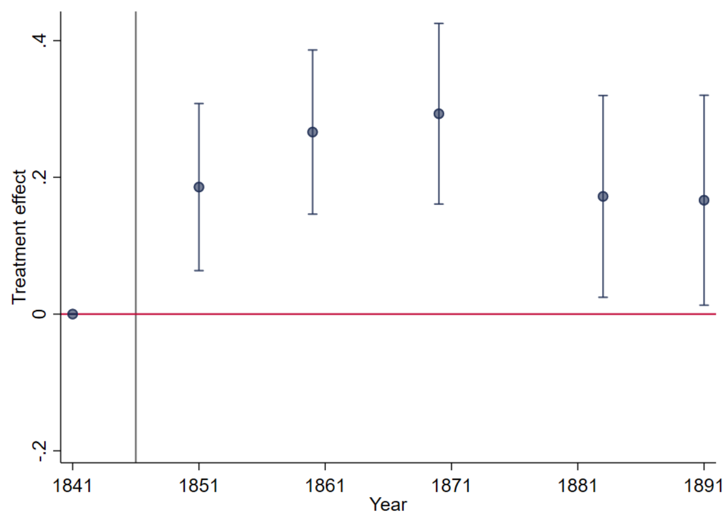


Note: Estimated treatment effects from the differences-in-differences specification in equation (1) in the paper using the log of property values (rateable values) for 1815, 1843 (excluded category), 1852, 1865, 1870, 1881, and 1896 and above-median wheat suitability interacted with census-year dummies; vertical lines show 95 percent confidence intervals based on standard errors clustered by registration district. The specification conditions on parish and year fixed effects, and interactions between year and (i) travel time to the nearest market town; (ii) travel time to the nearest coalfield; (iii) distance to London and distance to Manchester; (iv) Easting and Northing of the parish centroid; (v) an indicator that is one for Wales; (vi) an indicator that is one for urban parishes based on the 1801 distribution of population densities.

Finally, we estimate our baseline event-study specification in equation (1) in the paper using our data on poor law relief instead of population. We use the log share of the population receiving poor law relief as a measure of welfare transfers that captures local economic distress. These poor law relief data are available for the years of 1841, 1851, 1860, 1870, 1881 and 1891 at the level of registration districts. We choose 1843 as the excluded category, as the last period before the Repeal of the Corn Laws. We match parishes to registration districts and cluster the standard errors by registration district to take account of the fact that poor law relief is measured at a more aggregate level.

Figure B.18 displays our estimated treatment effects using poor law relief. Following the Repeal of the Corn Laws and the Grain Invasion in the second half of the 19th century, we find a statistically significant increase in welfare relief in high-wheat-suitability locations relative to low-wheat-suitability locations. This pattern of results is consistent with an increase in local economic distress in response to this trade shock driving the population reallocation away from these locations established above. We find that the estimated treatment effect for poor law relief declines somewhat in absolute magnitude by the last two decades of the 19th century, which is consistent with this population reallocation helping to dissipate local economic distress in high-wheat-suitability locations.

Figure B.18: Estimated Wheat Suitability Treatment Effects for Log Share of the Population Receiving Poor Law Relief



Notes: Estimated treatment effects (β_t) from the differences-in-differences specification in equation (1) in the paper using the log share of the population receiving poor law relief and above-median wheat suitability interacted with census-year dummies; vertical lines show 95 percent confidence intervals based on standard errors clustered by registration district. The specification conditions on parish and year fixed effects, and interactions between year and (i) travel time to the nearest market town; (ii) travel time to the nearest coalfield; (iii) distance to London and distance to Manchester; (iv) Easting and Northing of the parish centroid; (v) an indicator that is one for Wales; (vi) an indicator that is one for urban parishes based on the 1801 distribution of population densities.

C Theoretical Appendix

In this section of the Online Appendix, we report the derivations of all theoretical results from Section 5 of the paper. The title of each the following subsections is the same as the title of the corresponding subsection in Section 5 of the paper.

C1 Preferences

No further derivations required for Section 5.1 of the paper.

C2 Prices and Expenditure Shares

No further derivations required for Section 5.2 of the paper.

C3 Production Technology

No further derivations required for Section 5.3 of the paper.

C4 Production Within Agriculture

In this subsection, we report the derivations for Subsection 5.4 of the paper.

C4.1 Production of the Disaggregated Agricultural Goods

The firm problem for the disaggregated agricultural goods is:

$$\max_{n_{gi}(\varphi)} \left\{ \pi_{gi}\left(\varphi\right) = P_{gi} q_{gi}\left(\varphi\right) - w_i n_{gi}\left(\varphi\right) - r_{gi}\left(\varphi\right) \right\}.$$

From the first-order condition for profit maximization and zero profits, we obtain the following closed-form solutions for equilibrium employment, the wage-rental ratio and land rents for the disaggregated agricultural goods:

$$n_{gi}\left(\varphi\right) = z_{Ai}\left(\varphi\right) a_{gi}\left(\varphi\right) \left(\frac{\left(1-\alpha_g\right) P_{gi} \vartheta_g}{w_i} \right)^{\frac{1}{\alpha_g}}. \tag{C.1}$$

Substituting for equilibrium employment in the zero-profit condition, the bid rent satisfies:

$$r_{gi}\left(\varphi\right) = \alpha_g \left(\frac{1-\alpha_g}{w_i} \right)^{\frac{1-\alpha_g}{\alpha_g}} z_{Ai}\left(\varphi\right) a_{gi}\left(\varphi\right) \left(P_{gi} \vartheta_g\right)^{\frac{1}{\alpha_g}}. \tag{C.2}$$

We define the wage-rental ratio for disaggregated agricultural good g as:

$$\omega_{gi}\left(\varphi\right) \equiv \frac{w_i}{r_{gi}\left(\varphi\right)}. \tag{C.3}$$

Using the bid rent (C.2) and the definition of $\vartheta_g \equiv \alpha_g^{-\alpha_g} (1-\alpha_g)^{-(1-\alpha_g)}$ in equation (C.3), the wage-rental ratio for disaggregated agricultural good g satisfies:

$$\omega_{gi}(\varphi) = \frac{1}{z_{Ai}(\varphi)\, a_{gi}(\varphi)} \left(\frac{w_i}{P_{gi}}\right)^{\frac{1}{\alpha_g}}. \tag{C.4}$$

Under our assumption of a Cobb-Douglas production technology for each disaggregated agricultural good, payments to labor and land are constant shares of revenue, which implies that equilibrium employment per unit of land area can be written as:

$$n_{gi}(\varphi) = \frac{1-\alpha_g}{\alpha_g} \frac{1}{\omega_{gi}(\varphi)}. \tag{C.5}$$

C4.2 Specialization Across the Disaggregated Agricultural Goods

Agricultural land plots are allocated to either arable or pastoral farming based on which use offers the highest rental rate ($r_{gi}(\varphi)$). With a common wage across sectors and disaggregated agricultural goods within each location (w_i), we can equivalently characterize this allocation decision in terms of which use offers the lowest wage-rental ratio ($\omega_{gi}(\varphi)$):

$$\omega_i(\varphi) = \min_g \{\omega_{gi}(\varphi)\}. \tag{C.6}$$

The share of agricultural land allocated to each disaggregated agricultural good solves:

$$\ell_{gi}^A \equiv \frac{L_{gi}}{L_{Ai}} = \text{Prob}\left\{ g = \arg \min_{o \in \{G,F\}} \{\omega_{oi}\} \right\},$$

where the superscript A indicates that this land share is expressed as a share of agricultural land ($\ell_{gi}^A \equiv L_{gi}/L_{Ai}$). From the Fréchet distribution of agricultural productivity, we have:

$$F_{gi}(a) = \text{Prob}\left[a_{gi} \leq a\right] = e^{-E_{gi} a^{-\epsilon}}.$$

From zero-profits and maximization for each agricultural good, we have the following monotonic relationship between agricultural productivity and the wage-rental ratio for each land plot φ, agricultural good g and location i:

$$a_{gi}(\varphi) = \frac{1}{z_{Ai}(\varphi)\, \omega_{gi}(\varphi)} \left(\frac{w_i}{P_{gi}}\right)^{\frac{1}{\alpha_g}}.$$

Using this monotonic relationship, the wage-rental ratio has a Fréchet distribution:

$$F_{gi}(\omega) = \text{Prob}\left[\omega_{gi} \leq \omega\right] = 1 - e^{-\Phi_{gi}\omega^{\epsilon}}, \qquad \Phi_{gi} \equiv z_{Ai}^{\epsilon} E_{gi} \left(w_i/P_{gi}\right)^{-\epsilon/\alpha_g}.$$

The probability that each land plot φ is allocated to agricultural good g in location i is:

$$\begin{aligned}
\ell_{gi}^{A} &= \int_0^\infty \prod_{h\neq g} \left(1 - F_{hi}\left(\omega\right)\right) \mathrm{d}F_{gi}\left(\omega\right), \\
&= \int_0^\infty \left[\prod_{h\neq g} e^{-\Phi_{hi}\omega^{\epsilon}}\right] \Phi_{gi}\epsilon\omega^{\epsilon-1} e^{-\Phi_{gi}\omega^{\epsilon}} \mathrm{d}\omega, \\
&= \int_0^\infty \left[\prod_{h\in\{G,P\}} e^{-\Phi_{hi}\omega^{\epsilon}}\right] \Phi_{gi}\epsilon\omega^{\epsilon-1} \mathrm{d}\omega, \\
&= \int_0^\infty \left[e^{-\Phi_{Ai}\omega^{\epsilon}}\right] \Phi_{gi}\epsilon\omega^{\epsilon-1} \mathrm{d}\omega.
\end{aligned}$$

$$\Phi_{Ai} \equiv \sum_{g\in\{G,F\}} \Phi_{gi}.$$

Note that

$$\frac{\mathrm{d}}{\mathrm{d}\omega}\left[-\frac{1}{\Phi_{Ai}} e^{-\Phi_{Ai}\omega^{\epsilon}}\right] = \epsilon\omega^{\epsilon-1} e^{-\Phi_{Ai}\omega^{\epsilon}}.$$

Therefore

$$\ell_{gi}^{A} = \Phi_{gi}\left[-\frac{1}{\Phi_{Ai}} e^{-\Phi_{Ai}\omega^{\epsilon}}\right]_0^\infty,$$

$$\ell_{gi}^{A} = \frac{\Phi_{gi}}{\Phi_{Ai}} = \frac{E_{gi}\left(P_{gi}/w_i\right)^{\epsilon/\alpha_g}}{\sum_{o\in\{G,F\}} E_{oi}\left(P_{oi}/w_i\right)^{\epsilon/\alpha_o}}, \tag{C.7}$$

which corresponds to equation (14) in the paper.

C4.3 Expected Agricultural Bid Rent

The expected agricultural bid rent across the measure of land plots allocated to agriculture is:

$$r_{Ai} = \mathbb{E}_a\left[r_{gi}\right] = \mathbb{E}_a\left[\frac{w_i}{\omega_{gi}}\right] = w_i\mathbb{E}_a\left[\frac{1}{\omega_{gi}}\right].$$

Note that:

$$\text{Prob}\left[\omega_{gi} \leq \omega\right] = 1 - e^{-\Phi_{gi}\omega^{\epsilon}}, \qquad \Phi_{gi} \equiv z_{Ai}^{\epsilon} E_{gi}\left(w_i/P_{gi}\right)^{-\epsilon/\alpha_g}.$$

$$\text{Prob}\left[\omega_{Ai} \leq \omega\right] = 1 - e^{-\Phi_{Ai}\omega^{\epsilon}}, \qquad \Phi_{Ai} \equiv \sum_{g\in\{G,F\}} \Phi_{gi} = \sum_{g\in\{G,F\}} z_{Ai}^{\epsilon} E_{gi}\left(w_i/P_{gi}\right)^{-\epsilon/\alpha_g}.$$

Using these results, the expected wage-rental ratio can be written as:

$$\frac{1}{\omega_{Ai}} \equiv \mathbb{E}_a\left[\frac{1}{\omega_{gi}}\right] = \int_0^\infty \frac{1}{\omega}\left[\Phi_{Ai}\epsilon\omega_i^{\epsilon-1}\right] e^{-\Phi_{Ai}\omega^{\epsilon}} \mathrm{d}\omega. \tag{C.8}$$

Define the following change of variable:

$$y \equiv \Phi_{Ai}\omega^{\epsilon},$$

$$\omega = \left(\frac{y}{\Phi_{Ai}}\right)^{\frac{1}{\epsilon}},$$

$$\mathrm{d}y = \Phi_{Ai}\epsilon\omega^{\epsilon-1}\mathrm{d}\omega.$$

Using this change of variables, we can re-write the expectation (C.8) as:

$$\begin{aligned}
\frac{1}{\omega_{Ai}} \equiv \mathbb{E}_a\left[\frac{1}{\omega_{gi}}\right] &= \int_0^\infty \frac{1}{\omega}\left[\Phi_{Ai}\epsilon\omega^{\epsilon-1}\right] e^{-y}\mathrm{d}\omega, \\
&= \int_0^\infty \frac{1}{\omega} e^{-y}\mathrm{d}y, \\
&= \int_0^\infty y^{-\frac{1}{\epsilon}}\left[\sum_{g\in\{G,F\}} z_{Ai}^{\epsilon} E_{gi}\left(w_i/P_{gi}\right)^{-\epsilon/\alpha_g}\right]^{\frac{1}{\epsilon}} e^{-y}\mathrm{d}y.
\end{aligned}$$

Therefore the expected wage-rental ratio for each agricultural land plot within location i is the same across the two disaggregated agricultural goods and given by:

$$\frac{1}{\omega_{Ai}} \equiv \mathbb{E}_a\left[\frac{1}{\omega_{gi}}\right] = \gamma_\epsilon z_{Ai}\left[\sum_{g\in\{G,F\}} E_{gi}\left(w_i/P_{gi}\right)^{-\epsilon/\alpha_g}\right]^{\frac{1}{\epsilon}}, \qquad \gamma_\epsilon \equiv \Gamma\left(\frac{\epsilon-1}{\epsilon}\right), \tag{C.9}$$

where $\Gamma(\cdot)$ is the Gamma function and this expression corresponds to equation (11) in the paper.

Similarly, the expected rental rate for each agricultural land plot within location i is the same across the two disaggregated agricultural goods and given by:

$$r_i = w_i\mathbb{E}_a\left[\frac{1}{\omega_{gi}}\right] = \gamma_\epsilon w_i z_{Ai}\left[\sum_{g\in\{G,F\}} E_{gi}\left(w_i/P_{gi}\right)^{-\epsilon/\alpha_g}\right]^{\frac{1}{\epsilon}}. \tag{C.10}$$

Using profit maximization and zero profits, and our assumption of a Cobb-Douglas production technology, equilibrium employment per unit of agricultural land can be written as:

$$\mathbb{E}_a\left[n_{gi}\right] = \frac{1-\alpha_g}{\alpha_g}\frac{1}{\omega_{Ai}}. \tag{C.11}$$

Therefore, although the expected wage-rental ratio (as captured in $1/\omega_{Ai}$ in the second term on the right-hand side of (C.11)) is the same across the two disaggregated agricultural good within a given location, employment density per unit of agricultural land differs between arable and pastoral farming, because of the differences in factor intensity between these two disaggregated agricultural activities (as captured by the first term on the right-hand side of (C.11)).

C4.4 Multilateral Distribution of Bid Rents within Agriculture

The distribution of the wage-rental ratio in location i (ω_i) across the two disaggregated agricultural goods $g \in \{G, F\}$ is:

$$F_{Ai}(\omega) = 1 - \prod_{g \in \{G,F\}} e^{-\Phi_{gi}\omega^{\epsilon}} = 1 - e^{-\Phi_{Ai}\omega^{\epsilon}}.$$

We now show that this multilateral distribution of the wage-rental ratio across the two goods taken together is the same as the distribution of the wage-rental ratio for a given good conditional on agricultural land being allocated to that good:

$$\begin{aligned}
&= \frac{1}{\ell_{gi}^{A}} \int_0^{\omega} \prod_{k \neq g} \left(1 - F_{ki}(v)\right) f_{gi}(v)\, dv, \\
&= -\frac{1}{\ell_{gi}^{A}} \int_0^{\omega} \left[\prod_{k \neq g} e^{-\Phi_{ki} v^{\epsilon}}\right] \epsilon \Phi_{gi} v^{\epsilon-1} e^{-\Phi_{gi} v^{\epsilon}} dv, \\
&= -\frac{1}{\ell_{gi}^{A}} \int_0^{\omega} \left[\prod_{k \in \{G,F\}} e^{-\Phi_{ki} v^{\epsilon}}\right] \epsilon \Phi_{gi} v^{\epsilon-1} dv, \\
&= -\frac{\Phi_{Ai}}{\Phi_{gi}} \int_0^{\omega} \left[e^{-\Phi_{Ai} v^{\epsilon}}\right] \epsilon \Phi_{gi} v^{\epsilon-1} dv, \\
&= 1 - e^{-\Phi_{Ai}\omega^{\epsilon}}.
\end{aligned}$$

Intuitively, if a location has lower production costs and faces higher prices for one of the disaggregated agricultural goods, it allocates land plots with lower realizations for idiosyncratic productivity to that good. With a Fréchet distribution for agricultural productivity, this composition effect exactly offsets the lower production costs and higher prices, such that the distribution of the wage-rental ratio conditional on agricultural land being used for a good is the same across the two disaggregated agricultural goods.

C4.5 Average Productivity of Agricultural Land

We now use this common distribution of the wage-rental ratio across the two disaggregated agricultural goods to derive an expression for average land productivity conditional on agricultural land being allocated to a given disaggregated good. We have:

$$F_{Ai}(\omega) = 1 - e^{-\Phi_{Ai}\omega^{\epsilon}}. \tag{C.12}$$

Recall the monotonic relationship between the wage-rental ratio and agricultural productivity:

$$\omega_{gi} = \frac{1}{z_{Ai} a_{gi}} \left(\frac{w_i}{P_{gi}}\right)^{\frac{1}{\alpha_g}}.$$

Using this relationship, the distribution of productivity conditional on allocating agricultural land to a given disaggregated agricultural good in equation (C.12) can be written as:

$$\begin{aligned} F_{gi}(a) &= e^{-\Phi_{Ai} z_{Ai}^{-\epsilon} (w_i/P_{gi})^{\epsilon/\alpha_g} a^{-\epsilon}}, \\ &= e^{-E_{gi}(\Phi_{Ai}/\Phi_{gi}) a^{-\epsilon}}. \end{aligned}$$

Using this distribution, we can compute expected productivity conditional on agricultural land being allocated to a given disaggregated good as:

$$\begin{aligned} \mathbb{E}_i[a] &= \int_0^\infty a f_{gi}(a)\, da, \\ &= \int_0^\infty \epsilon E_{gi}(\Phi_{Ai}/\Phi_{gi})\, a^{-\epsilon} e^{-E_{gi}(\Phi_{Ai}/\Phi_{gi}) a^{-\epsilon}} da. \end{aligned}$$

Now define the following change of variables:

$$y = E_{gi}(\Phi_{Ai}/\Phi_{gi})\, a^{-\epsilon}, \qquad dy = -\epsilon E_{gi}(\Phi_{Ai}/\Phi_{gi})\, a^{-(\epsilon+1)} da,$$

$$a = \left(\frac{y}{E_{gi}(\Phi_{Ai}/\Phi_{gi})}\right)^{-\frac{1}{\epsilon}}, \qquad da = -\frac{dy}{\epsilon E_{gi}(\Phi_{Ai}/\Phi_{gi}) \left(\frac{y}{E_{gi}(\Phi_{Ai}/\Phi_{gi})}\right)^{\frac{\epsilon+1}{\epsilon}}}.$$

Using this change of variables, expected productivity conditional on agricultural land being allocated to a given disaggregated good can be written as:

$$\begin{aligned} \overline{a}_{gi} = \mathbb{E}_{gi}[a] &= \int_0^\infty \epsilon E_{gi}(\Phi_{Ai}/\Phi_{gi})\, a^{-\epsilon} e^{-E_{gi}(\Phi_{Ai}/\Phi_{gi}) a^{-\epsilon}} da, \\ &= \int_0^\infty \epsilon y e^{-y} \frac{dy}{\epsilon E_{gi}(\Phi_{Ai}/\Phi_{gi}) \left(\frac{y}{\Phi_{Ai}/\Phi_{gi}}\right)^{\frac{\epsilon+1}{\epsilon}}}, \\ &= \int_0^\infty y e^{-y} \frac{(y)^{-\frac{\epsilon+1}{\epsilon}}\, dy}{E_{gi}(\Phi_{Ai}/\Phi_{gi}) \left(E_{gi}(\Phi_{Ai}/\Phi_{gi})\right)^{-\frac{\epsilon+1}{\epsilon}}}, \\ &= \int_0^\infty y^{-1/\epsilon} e^{-y} \left(E_{gi}(\Phi_{Ai}/\Phi_{gi})\right)^{1/\epsilon} dy, \\ &= \gamma_\epsilon \left(E_{gi}(\Phi_{Ai}/\Phi_{gi})\right)^{1/\epsilon}, \\ &= \gamma_\epsilon \left(E_{gi}\right)^{1/\epsilon} \left(\ell_{gi}^A\right)^{-1/\epsilon}, \end{aligned}$$

which is decreasing in the share of agricultural land allocated to that disaggregated good (ℓ_{gi}^A).

C5 Production Across Sectors

In this subsection, we report the derivations for Subsection 5.5 of the paper.

C5.1 Production in Manufacturing and Services

The firm problem in the manufacturing and services sectors $k \in \{M, S\}$ is:

$$\max_{n_{ki}(\varphi)} \left\{\pi_{gi}(\varphi) = P_{ki}q_{ki}(\varphi) - w_i n_{ki}(\varphi) - r_{ki}(\varphi)\right\}. \tag{C.13}$$

From the first-order condition for profit maximization and zero profits, we obtain the following closed-form solutions for equilibrium employment per unit of land ($n_{ki}(\varphi)$), the wage-rental ratio ($\omega_{ki}(\varphi)$), and land rents ($r_{ki}(\varphi)$) in each sector $k \in \{M, S\}$:

$$n_{ki}(\varphi) = z_{ki}(\varphi)\left(\frac{(1-\alpha_k)\,P_{ki}\vartheta_k}{w_i}\right)^{\frac{1}{\alpha_k}}, \tag{C.14}$$

$$\omega_{ki}(\varphi) = \frac{w_i}{r_{ki}(\varphi)} = \frac{1}{z_{ki}(\varphi)}\left(\frac{w_i}{P_{ki}}\right)^{\frac{1}{\alpha_k}}, \tag{C.15}$$

$$r_{ki}(\varphi) = w_i\left(\frac{P_{ki}}{w_i}\right)^{\frac{1}{\alpha_k}} z_{ki}(\varphi). \tag{C.16}$$

C5.2 Specialization Across Sectors

Recall that a land plot is allocated to the use that offers the highest rental rate or equivalently the lowest wage-rental ratio. Therefore, the equilibrium wage-rental ratio across the agricultural, manufacturing and services sectors for an individual land plot φ in location i solves:

$$\omega_i(\varphi) = \min_{k\in\{A,M,S\}}\left\{\omega_{ki}(\varphi)\right\}, \tag{C.17}$$

where since the sectoral land use decision is made before the realizations for agricultural productivity (a_{gi}) are observed, the relevant wage-rental ratio for the agricultural sector is the expected wage-rental ratio equation (C.9).

From equations (C.9) and (C.15), the expected wage-rental ratio in agriculture and the wage-rental ratios in manufacturing and services can be written as:

$$\omega_{ki}(\varphi) = \frac{1}{z_{ki}\mathcal{P}_{ki}}, \qquad \mathcal{P}_{ki} \equiv \left(\frac{P_{ki}}{w_i}\right)^{\frac{1}{\alpha_k}}, \qquad k \in \{M, S\}, \tag{C.18}$$

$$\omega_{Ai}(\varphi) = \frac{1}{z_{Ai}\mathcal{P}_{Ai}}, \qquad \mathcal{P}_{Ai} \equiv \gamma_\epsilon\left[\sum_{g\in\{G,F\}} E_{gi}\left(\frac{P_{gi}}{w_i}\right)^{\frac{\epsilon}{\alpha_g}}\right]^{\frac{1}{\epsilon}}. \tag{C.19}$$

where $\mathcal{P}_{ki}$ denotes prices adjusted for the wage (w_i) and factor intensity (α_k, α_g), and recall that $\gamma_\epsilon \equiv \Gamma\left(\frac{\epsilon-1}{\epsilon}\right)$, where $\Gamma(\cdot)$ is the Gamma function.

From the monotonic relationship between the wage-rental ratio and sectoral productivity in equations (C.18) and (C.19), the wage-rental ratio in each sector $k \in \{A, M, S\}$ has the following Fréchet distribution:

$$F_{ki}(\omega) = 1 - e^{-\Psi_{ki}\omega^{\theta}}, \qquad \Psi_{ki} \equiv T_{ki}\mathcal{P}_{ki}^{\theta}.$$

The probability that a land plot if allocated to sector $k \in \{A, M, S\}$ is therefore:

$$\ell_{ki} = \int_0^{\infty} \prod_{s \neq k} (1 - F_s(\omega))\, \mathrm{d}F_k(\omega),$$

$$\ell_{ki} = \Psi_{ki} \int_0^{\infty} e^{-\Psi_i \omega^{\theta}} \theta \omega^{\theta - 1} \mathrm{d}\omega,$$

where

$$\Psi_i \equiv \sum_{s \in \{A,M,S\}} \Psi_{ki} = \sum_{s \in \{A,M,S\}} T_{si}\mathcal{P}_{si}^{\theta}.$$

Now note that:

$$\frac{\mathrm{d}}{\mathrm{d}\omega}\left[-\frac{1}{\Psi_i} e^{-\Psi_i \omega^{\theta}}\right] = \theta \omega^{\theta - 1} e^{-\Psi_i \omega^{\theta}}.$$

Therefore

$$\ell_{ki} = \Psi_{ki}\left[-\frac{1}{\Psi_i} e^{-\Psi_i \omega^{\theta}}\right]_0^{\infty},$$

and the share of land allocated to sector $k \in \{A, M, S\}$ is:

$$\ell_{ki} = \frac{\Psi_{ki}}{\Psi_i} = \frac{T_{ki}\mathcal{P}_{ki}^{\theta}}{\sum_{s \in \{A,M,S\}} T_{si}\mathcal{P}_{si}^{\theta}}, \tag{C.20}$$

which corresponds to equation (14) in the paper.

C5.3 Expected Bid Rent Across All Sectors

The expected bid rent across sectors for each of the continuum of land plots is:

$$r_i = \mathbb{E}_z[r_{ki}] = \mathbb{E}_z\left[\frac{w_i}{\omega_{ki}}\right] = w_i \mathbb{E}_z\left[\frac{1}{\omega_{ki}}\right].$$

Note that:

$$\begin{array}{rclcrcl} \text{Prob}[\omega_{ki} \leq \omega] & = & 1 - e^{-\Psi_{ki}\omega^{\theta}}, & & \Psi_{ki} & \equiv & T_{ki}\mathcal{P}_{ki}^{\theta}, \\ \text{Prob}[\omega_i \leq \omega] & = & 1 - e^{-\Psi_i\omega^{\theta}}, & & \Psi_i & \equiv & \sum_{s \in \{A,M,S\}} T_{si}\mathcal{P}_{si}^{\theta}. \end{array}$$

Using these results, the expected wage-rental ratio can be written as:

$$\mathbb{E}_z\left[\frac{1}{\omega_i}\right] = \int_0^{\infty} \frac{1}{\omega} \Psi_i \theta \omega^{\theta - 1} e^{-\Psi_i \omega^{\theta}} \mathrm{d}\omega. \tag{C.21}$$

Define the following change of variables:

$$y = \Psi_i \omega^\theta,$$

$$\omega = \left(\frac{y}{\Psi_i}\right)^{\frac{1}{\theta}},$$

$$\mathrm{d}y = \Psi_i \theta \omega^{\theta-1} \mathrm{d}\omega.$$

Using this change of variables, we can rewrite the expected wage-rental (C.21) ratio as:

$$\begin{aligned}\mathbb{E}_z\left[\frac{1}{\omega_i}\right] &= \int_0^\infty \frac{1}{\omega}\Psi_i \theta \omega^{\theta-1} e^{-y} \mathrm{d}\omega, \\ &= \int_0^\infty \frac{1}{\omega} e^{-y} \mathrm{d}y, \\ &= \int_0^\infty y^{-\frac{1}{\theta}} \Psi_i^{\frac{1}{\theta}} e^{-y} \mathrm{d}y,\end{aligned}$$

and hence:

$$\mathbb{E}_z\left[\frac{1}{\omega_i}\right] = \gamma_\theta \Psi_i^{\frac{1}{\theta}}, \qquad \gamma_\theta \equiv \Gamma\left(\frac{\theta-1}{\theta}\right),$$

where $\Gamma(\cdot)$ is the Gamma function. We thus obtain:

$$\mathbb{E}_z\left[\frac{1}{\omega_i}\right] = \gamma_\theta \left[\sum_{k\in\{A,M,S\}} T_{ki} \mathcal{P}_{ki}^\theta\right]^{\frac{1}{\theta}}, \tag{C.22}$$

which corresponds to equation (15) in the paper.

Using this expression for the expected wage-rental ratio (C.22), we can obtain the following closed-form solutions for land rents:

$$r_i = w_i \mathbb{E}_z\left[\frac{1}{\omega_i}\right] = w_i \gamma_\theta \left[\sum_{k\in\{A,M,S\}} T_{ki} \mathcal{P}_{ki}^\theta\right]^{\frac{1}{\theta}}. \tag{C.23}$$

and employment density by sector:

$$n_{Mi} = \frac{1-\alpha_M}{\alpha_M}\mathbb{E}_z\left[\frac{1}{\omega_i}\right], \tag{C.24}$$

$$n_{Si} = \frac{1-\alpha_S}{\alpha_S}\mathbb{E}_z\left[\frac{1}{\omega_i}\right], \tag{C.25}$$

$$n_{Hi} = \frac{1-\alpha_H}{\alpha_H}\mathbb{E}_z\left[\frac{1}{\omega_i}\right], \tag{C.26}$$

$$n_{Ai} = \sum_{g\in\{G,F\}} \frac{L_{gi}}{L_{Ai}}\frac{N_{gi}}{L_{gi}} = \sum_{g\in\{G,F\}} \ell_{gi}^A n_{gi} = \sum_{g\in\{G,F\}} \ell_{gi}^A \frac{1-\alpha_g}{\alpha_g}\mathbb{E}_z\left[\frac{1}{\omega_i}\right]. \tag{C.27}$$

C5.4 Multilateral Distribution of Bid Rents across Sectors

The distribution of the wage-rental ratio in location i across sectors $k \in \{A, M, S\}$ is:

$$F_i(\omega) = 1 - \prod_{k \in \{A,M,S\}} e^{-\Psi_{ki}\omega^\theta} = 1 - e^{-\Psi_i \omega^\theta}.$$

We now show that this multilateral distribution of the wage-rental ratio across sectors $k \in \{A, M, S\}$ is the same as the distribution of the wage-rental ratio for a given sector conditional on land being allocated to that sector:

$$\begin{aligned}
&= \frac{1}{\ell_{gi}} \int_0^\omega \prod_{k \neq g} (1 - F_{ki}(v)) f_{gi}(v) \, dv, \\
&= -\frac{1}{\ell_{gi}} \int_0^\omega \left[\prod_{k \neq g} e^{-\Psi_{ki} v^\theta} \right] \theta \Psi_{gi} v^{\theta-1} e^{-\Psi_{gi} v^\theta} dv, \\
&= -\frac{1}{\ell_{gi}} \int_0^\omega \left[\prod_{k \in \{A,M,S\}} e^{-\Psi_{ki} v^\theta} \right] \theta \Psi_{gi} v^{\theta-1} dv, \\
&= -\frac{\Psi_i}{\Psi_{gi}} \int_0^\omega \left[e^{-\Psi_i v^\theta} \right] \theta \Psi_{gi} v^{\theta-1} dv, \\
&= 1 - e^{-\Psi_i \omega^\theta}.
\end{aligned}$$

Intuitively, if a location has higher adjusted prices for a sector, it allocates land plots with lower realizations for sector productivity to that sector. With a Fréchet distribution for sector productivity, this composition effect exactly offsets the higher adjusted prices, such that the distribution of the wage-rental ratio conditional on land being allocated to a sector is the same across sectors.

C5.5 Average Land Productivity

Recall that the distribution of the wage-rental ratio across sectors $k \in \{A, M, S\}$ is equal to the distribution of the wage-rental ratio conditional on land being allocated to a given sector:

$$F_i(\omega) = 1 - e^{-\Psi_i \omega^\theta}.$$

Now recall the monotonic relationship between the wage-rental ratio and sectoral productivity:

$$\omega_{ki} = \frac{1}{z_{ki} \mathcal{P}_{ki}}.$$

Therefore the distribution of sectoral productivity conditional on land being allocated to a given sector is:

$$F_{ki}(z) = e^{-\Psi_i \mathcal{P}_{ki}^{-\theta} z^{-\theta}},$$

which can be written as:

$$F_{ki}\left(z\right) = e^{-T_{ki}(\Psi_i/\Psi_{ki})z^{-\theta}}, \qquad k \in \left\{A, M, S\right\}.$$

Expected sectoral productivity conditional on land being allocated to a sector is:

$$\begin{aligned}
\mathbb{E}_i\left[z\right] &= \int_0^\infty z f_n\left(z\right) dz, \\
&= \int_0^\infty \theta T_{ki}\left(\Psi_i/\Psi_{ki}\right) z^{-\theta} e^{-T_{ki}(\Psi_i/\Psi_{ki})z^{-\theta}} dz.
\end{aligned}$$

Now define the following change of variables:

$$y = T_{ki}\left(\Psi_i/\Psi_{ki}\right) z^{-\theta}, \qquad dy = -\theta T_{ki}\left(\Psi_i/\Psi_{ki}\right) z^{-(\theta+1)} dz,$$

$$z = \left(\frac{y}{T_{ki}(\Psi_i/\Psi_{ki})}\right)^{-\frac{1}{\theta}}, \qquad dz = -\frac{dy}{\theta T_{ki}\left(\Psi_i/\Psi_{ki}\right)\left(\frac{y}{T_{ki}(\Psi_i/\Psi_{ki})}\right)^{\frac{\theta+1}{\theta}}}.$$

Using this change of variables, expected productivity conditional on land being allocated to a sector can be written as:

$$\begin{aligned}
\overline{z}_{ki} = \mathbb{E}_{ki}\left[z\right] &= \int_0^\infty \theta T_{ki}\left(\Psi_i/\Psi_{ki}\right) z^{-\theta} e^{-T_{ki}(\Psi_i/\Psi_{ki})z^{-\theta}} dz, \\
&= \int_0^\infty \theta y e^{-y} \frac{dy}{\theta T_{ki}\left(\Psi_i/\Psi_{ki}\right)\left(\frac{y}{T_{ki}(\Psi_i/\Psi_{ki})}\right)^{\frac{\theta+1}{\theta}}}, \\
&= \int_0^\infty y e^{-y} \frac{(y)^{-\frac{\theta+1}{\theta}} dy}{T_{ki}\left(\Psi_i/\Psi_{ki}\right)\left(T_{ki}(\Psi_i/\Psi_{ki})\right)^{-\frac{\theta+1}{\theta}}}, \\
&= \int_0^\infty y^{-1/\theta} e^{-y} \left(T_{ki}(\Psi_i/\Psi_{ki})\right)^{1/\theta} dy, \\
&= \gamma_\theta \left(T_{ki}(\Psi_i/\Psi_{ki})\right)^{1/\theta}, \\
&= \gamma_\theta T_{ki}^{1/\theta} \ell_{ki}^{-1/\theta},
\end{aligned}$$

which is decreasing in the share of land allocated to that sector (ℓ_{ki}).

C6 Population Mobility

In this subsection, we report the derivations for Subsection 5.6 of the paper.

C6.1 Bilateral Distribution of Utility

The indirect utility of a worker ψ who chooses to live in location i can be written as follows:

$$u_i\left(\psi\right) = \frac{B_i b_i\left(\psi\right) w_i}{P_{Ci}}.$$

Amenities are drawn independently from the following distribution:

$$F\left(b\right)=e^{-b^{-\chi}},\qquad \chi>1.$$

Using the monotonic relationship between utility and amenities, we have:

$$b_i=\frac{u_iP_{Ci}}{B_iw_i}.$$

The distribution of utility for a worker who chooses to live in location i is therefore:

$$F_i\left(u\right)=e^{-\Theta_i u^{-\chi}},\qquad \chi>1.$$

$$\Theta_i\equiv\left(B_iw_i/P_{Ci}\right)^{\chi}.$$

C6.2 Location Choice Probabilities

The probability that a worker chooses to live in location i is:

$$\begin{aligned}
\lambda_i &= \text{Prob}\left[u_i\geq\max\left\{u_m\right\}\forall m\right],\\
&=\int_0^{\infty}\prod_{m\neq i}F_m\left(u\right)f_i\left(u\right)du,\\
&=\int_0^{\infty}\left[\prod_{m\neq i}e^{-\Theta_m u^{-\chi}}\right]\chi\Theta_i u^{-(\chi+1)}e^{-\Theta_i u^{-\chi}}du,\\
&=\int_0^{\infty}\left[\prod_{m\in\mathcal{J}}e^{-\Theta_m u^{-\chi}}\right]\chi\Theta_i u^{-(\chi+1)}du,\\
&=\int_0^{\infty}\left[e^{-\Theta u^{-\chi}}\right]\chi\Theta_i u^{-(\chi+1)}du,
\end{aligned}$$

where:

$$\Theta\equiv\sum_{m\in\mathcal{J}}\Theta_m.$$

Note that:

$$\frac{d}{du}\left[\frac{1}{\Theta}e^{-\Theta u^{-\chi}}\right]=\chi u^{-(\chi+1)}e^{-\Theta u^{-\chi}}.$$

Using this result, the probability that a worker chooses to live in location i is:

$$\lambda_i=\frac{\Theta_i}{\Theta}=\frac{\left(B_iw_i/P_{Ci}\right)^{\chi}}{\sum_{m\in\mathcal{J}}\left(B_mw_m/P_{Cm}\right)^{\chi}},$$

which corresponds to equation (19) in the paper.

C6.3 Multilateral Distribution of Utility

The distribution of utility across all locations $m \in \mathcal{J}$ is:

$$F_i\left(u\right) = \prod_{m \in \mathcal{J}} e^{-\Theta_m u^{-\chi}} = e^{-\Theta u^{-\chi}}.$$

We now show that this multilateral distribution of utility across all locations $m \in \mathcal{J}$ is the same as the distribution of utility conditional on choosing a given location i:

$$\begin{aligned}
&= \frac{1}{\lambda_i} \int_o^u \left[\prod_{m \neq n} F_m\left(v\right)\right] f_i\left(v\right) dv, \\
&= \frac{1}{\lambda_i} \int_o^u \left[\prod_{m \neq n} e^{-\Theta_m v^{-\chi}}\right] \chi \Theta_i v^{-(\chi+1)} e^{-\Theta_i u^{-\chi}} dv, \\
&= \frac{1}{\lambda_i} \int_o^u \left[\prod_{m \in \mathcal{J}} e^{-\Theta_m v^{-\chi}}\right] \chi \Theta_i v^{-(\chi+1)} dv, \\
&= \frac{\Theta}{\Theta_i} \int_o^u \left[e^{-\Theta v^{-\chi}}\right] \chi \Theta_i v^{-(\chi+1)} dv, \\
&= e^{-\Theta_i u^{-\chi}}.
\end{aligned}$$

Intuitively, if a destination has higher wages or a lower cost of living, it attracts workers with lower realizations for idiosyncratic preferences. With a Fréchet distribution for idiosyncratic preferences, this composition effect exactly offsets the higher wages or lower cost of living, such that the distribution of utility conditional on choosing a location is the same across all locations.

C6.4 Expected Utility

Expected utility is:

$$\begin{aligned}
\mathbb{E}\left[u\right] &= \int_0^\infty u f\left(u\right) du, \\
&= \int_0^\infty \chi \Theta u^{-\chi} e^{-\Phi u^{-\chi}} du.
\end{aligned}$$

Now define the following change of variables:

$$y = \Theta u^{-\chi}, \qquad dy = -\chi \Theta u^{-(\chi+1)} du,$$

$$u = \left(\frac{y}{\Theta}\right)^{-\frac{1}{\chi}}, \qquad du = -\frac{dy}{\chi \Theta u^{-(\chi+1)}} = -\frac{dy}{\chi \Theta \left(\frac{y}{\Theta}\right)^{\frac{\chi+1}{\chi}}}.$$

Using this change of variables, expected utility can be written as:

$$\begin{aligned}
\mathbb{E}\left[u\right] &= \int_0^\infty \chi \Theta u^{-\chi} e^{-\Theta u^{-\chi}} du, \\
&= \int_0^\infty \chi y e^{-y} \frac{dy}{\chi \Theta \left(\frac{y}{\Theta}\right)^{\frac{\chi+1}{\chi}}}, \\
&= \int_0^\infty y e^{-y} \frac{y^{-\frac{\chi+1}{\chi}} dy}{\Phi\left(\Theta\right)^{-\frac{\chi+1}{\chi}}}, \\
&= \int_0^\infty \Theta^{1/\chi} y^{-1/\chi} e^{-y} dy,
\end{aligned}$$

which in turn can be written as:

$$\mathbb{E}\left[u\right] = \gamma_\chi \Theta^{1/\chi} = \gamma_\chi \left[\sum_{m \in \mathcal{J}} \left(B_m w_m / P_{Cm}\right)^\chi\right]^{\frac{1}{\chi}},$$

$$\gamma_\chi \equiv \Gamma\left(\frac{\chi - 1}{\chi}\right),$$

where $\Gamma\left(\cdot\right)$ is the Gamma function and this expression corresponds to equation (20) in the paper.

C7 Market Clearing

No additional derivations required for Section 5.7 of the paper.

C8 General Equilibrium

In this subsection, we report additional results for Section 5.8 of the paper. The equilibrium spatial distribution of economic activity is determined by (i) model parameters (β_G, β_F, β_A, β_M, β_S, α_G, α_F, α_M, α_S, σ, ϵ, θ, χ); (ii) world prices and transport costs for the traded goods (P^*_G, P^*_F, P^*_M, τ_{Gi}, τ_{Fi}, τ_{Mi}; (iii) productivities (E_{Gi}, E_{Fi}, T_{Ai}, T_{Mi}, T_{Si}); (iv) amenities (B_i); (v) land supplies (L_i); (vi) labor supply ($\overline{N}$).

Given these exogenous primitives, the general equilibrium of the model can be referenced by the following endogenous variables: (i) the local prices of traded goods {P_{Gi}, P_{Fi}, P_{Mi}}, (ii) the local prices of non-traded services {P_{Si}}, (iii) the shares of arable and pastoral farming in agricultural land area (ℓ^A_{Gi}, ℓ^A_{Fi}); (iv) the shares of agriculture, manufacturing, and services in total land area (ℓ_{Ai}, ℓ_{Mi}, ℓ_{Si}); (v) the expected wage-rental ratio (ω_i); (vi) the wage (w_i); (vii) total employment shares (λ_i). Given these equilibrium objects, all other endogenous variables can be determined. The general equilibrium of the model solves the following system of equations:

$$\ell^A_{Gi} = \frac{E_{Gi}\left(P_{Gi}/w_i\right)^{\epsilon/\alpha_G}}{\sum_{g \in \{G,F\}} E_{gi}\left(P_{gi}/w_i\right)^{\epsilon/\alpha_g}}, \tag{C.28}$$

$$\ell_{Fi}^{A} = \frac{E_{Fi}\left(P_{Fi}/w_i\right)^{\epsilon/\alpha_F}}{\sum_{g\in\{G,F\}} E_{gi}\left(P_{gi}/w_i\right)^{\epsilon/\alpha_g}} = 1 - \ell_{Gi}^{A}, \tag{C.29}$$

$$\ell_{Ai} = \frac{T_{Ai}\mathcal{P}_{Ai}^{\theta}}{\sum_{s\in\{A,M,S\}} T_{si}\mathcal{P}_{si}^{\theta}}, \tag{C.30}$$

$$\ell_{Mi} = \frac{T_{Mi}P_{Mi}^{\theta}}{\sum_{s\in\{A,M,S\}} T_{si}\mathcal{P}_{si}^{\theta}}, \tag{C.31}$$

$$\ell_{Si} = \frac{T_{Si}\mathcal{P}_{Si}^{\theta}}{\sum_{s\in\{A,M,S\}} T_{si}\mathcal{P}_{si}^{\theta}} = 1 - \ell_{Ai} - \ell_{Mi}, \tag{C.32}$$

$$\mathcal{P}_{Ai} = \gamma_\epsilon \left[\sum_{g\in\{G,F\}} E_{gi}\left(\frac{P_{gi}}{w_i}\right)^{\frac{\epsilon}{\alpha_g}}\right]^{\frac{1}{\epsilon}}, \tag{C.33}$$

$$\mathcal{P}_{Mi} = \left(P_{Mi}/w_i\right)^{1/\alpha_M} \tag{C.34}$$

$$\mathcal{P}_{Si} = \left(P_{Si}/w_i\right)^{1/\alpha_S} \tag{C.35}$$

$$L_{Gi} = \ell_{Gi}^{A}\ell_{Ai}L_i, \tag{C.36}$$

$$L_{Fi} = \ell_{Fi}^{A}\ell_{Ai}L_i, \tag{C.37}$$

$$L_{Ai} = L_{Gi} + L_{Fi}, \tag{C.38}$$

$$L_{Mi} = \ell_{Mi}L_i, \tag{C.39}$$

$$L_{Si} = \ell_{Si}L_i, \tag{C.40}$$

$$w_{it}N_{Git} = \frac{1-\alpha_G}{\alpha_G} r_{it}L_{Git}, \tag{C.41}$$

$$w_{it}N_{Fit} = \frac{1-\alpha_F}{\alpha_F} r_{it}L_{Fit}, \tag{C.42}$$

$$w_{it}N_{Ait} = w_{it}N_{Git} + w_{it}N_{Fit}, \tag{C.43}$$

$$w_{it}N_{Mit} = \frac{1-\alpha_M}{\alpha_M} r_{it}L_{Mit}, \tag{C.44}$$

$$w_{it}N_{Sit} = \frac{1-\alpha_S}{\alpha_S} r_{it}L_{Sit}, \tag{C.45}$$

$$P_{Ai} = P_{Gi}^{\beta_G}P_{Fi}^{1-\beta_G}, \tag{C.46}$$

$$P_{Ci} = \left[\sum_{k\in\{A,M,S\}} \left(P_{ki}/\beta_k\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}, \tag{C.47}$$

$$\lambda_i = \frac{\left(B_i w_i/P_{Ci}\right)^{\chi}}{\sum_{m\in\mathcal{J}} \left(B_m w_m/P_{Cm}\right)^{\chi}}, \tag{C.48}$$

$$N_i = \lambda_i \overline{N}, \tag{C.49}$$

$$P_{Si} \left(N_{Si} \right)^{1-\alpha_S} \left(\overline{z}_{Si} \ell_{Si} L_i \right)^{\alpha_S} = x_{Si} \left[r_i L_i + w_i N_i \right], \tag{C.50}$$

$$x_{Si} = \frac{\left(P_{Si} / \beta_S \right)^{1-\sigma}}{\sum_{k \in \{A,M,S\}} \left(P_{ki} / \beta_k \right)^{1-\sigma}}, \tag{C.51}$$

$$\overline{z}_{Si} = \gamma_\theta T_{Si}^{1/\theta} \ell_{Si}^{-1/\theta}, \tag{C.52}$$

$$w_i N_i = \sum_{k \in \{A,M,S\}} w_i N_{ki}, \tag{C.53}$$

$$r_i = w_i \gamma_\theta \left[\sum_{k \in \{A,M,S\}} T_{ki} \mathcal{P}_{ki}^{\theta} \right]^{\frac{1}{\theta}}. \tag{C.54}$$

D Quantitative Analysis

In this section of the Online Appendix, we provide further details on our quantitative analysis of the model from Section 6 of the paper. Subsection D1 calibrates the model's parameters. Subsection D2 uses the observed endogenous variables and the structure of the model to recover other unobserved endogenous variables that are inputs into our counterfactuals. We show that the model rationalizes the observed data as an equilibrium outcome by allowing for unobserved productivities and amenities that vary by sector, location and year to control for other potential determinants of economic activity. Section D3 undertakes our counterfactuals that evaluate the impact of the Grain Invasion, which corresponds to a fall in the world relative price of arable products, holding constant all other exogenous variables. Throughout this section, we make explicit the dependence on time through the time subscript.

D1 Parameterization

In this subsection, we provide further details on our calibration of the model's parameters from Section 6.1 of the paper. We calibrate the model's parameters using historical data for our sample period and estimates from the related empirical literature.

D1.1 Demand Parameters

We set the elasticity of substitution across sectors equal to $\sigma = 0.5$, which is a central value in the existing macroeconomic literature on structural transformation, and is close to the value of $\sigma = 0.49$ estimated for late-19th-century Argentina in Fajgelbaum and Redding (2022). Using our assumption of CES preferences and data for a number of different countries and time periods, Bah (2007), Rogerson (2008), Duarte and Restuccia (2010), and Üngör (2017) obtain values for this elasticity of substitution of 0.44, 0.45, 0.40, and 0.47, respectively.

Our quantitative analysis does not require is to specify values for the CES weights for each sector (β_A, β_M, β_S). Instead, we use the observed data and the structure of our model to solve for implied expenditure shares for each sector and location in the initial equilibrium in the data, which capture these unobserved weights. We calibrate the consumer expenditure share parameters within the agricultural sector using the historical consumer expenditure survey data for Britain at the turn of the 20th century from Allen and Bowley (1935). We set the expenditure share on arable products as $\beta_G = 0.5$ and the expenditure share on pastoral products as $\beta_F = 1 - \beta_G = 0.5$, as summarized in Table D.1 below.

Table D.1: Shares of Agricultural Expenditure in the United Kingdom in 1904

Category	Sub-category	Share of Agric. Expenditure	Parameter
Arable	Cereals	0.185	
	Vegetables	0.08	
	Sugar	0.065	
	Tea, coffee	0.07	
	Miscellaneous	0.1	
	Sub-total	**0.50**	β_G
Pastoral	Meat	0.31	
	Dairy	0.19	
	Sub-total	**0.50**	β_F
	Total	1	

Note: Shares of agricultural expenditure in the United Kingdom in 1904 from Table B, page 34 in Allen and Bowley (1935); β_G is share of arable products in agricultural expenditure; $\beta_F = 1 - \beta_G$ is share of pastoral products in agricultural expenditure.

D1.2 Production Cost Shares

We calibrate the shares of land in production costs for arable farming (α_G), pastoral farming (α_F), manufacturing (α_M), and services (α_S) using data from Feinstein (1972) and Deane and Cole (1967).

Arable and Pastoral Farming We calibrate the land shares for arable farming (α_G) and pastoral farming (α_F) to match aggregate data on the land share for the agricultural sector as a whole

(α_A) and data on the relative labor intensity of arable and pastoral farming. Our model implies:

$$\begin{aligned}\alpha_{Ait} &= \tfrac{r_{it}L_{Ait}}{r_{it}L_{Ait}+w_{it}N_{Ait}},\\ &= \tfrac{r_{it}L_{Git}}{r_{it}L_{Git}+r_{it}L_{Fit}+w_{it}N_{Git}+w_{it}N_{Fit}} + \tfrac{r_{it}L_{Fit}}{r_{it}L_{Git}+r_{it}L_{Fit}+w_{it}N_{Git}+w_{it}N_{Fit}},\\ &= \tfrac{r_{it}L_{Git}}{r_{it}L_{Git}+r_{it}L_{Fit}+\frac{1-\alpha_G}{\alpha_G}r_{it}L_{Git}+\frac{1-\alpha_F}{\alpha_F}r_{it}L_{Fit}} + \tfrac{r_{it}L_{Fit}}{r_{it}L_{Git}+r_{it}L_{Fit}+\frac{1-\alpha_G}{\alpha_G}r_{it}L_{Git}+\frac{1-\alpha_F}{\alpha_F}r_{it}L_{Fit}},\\ &= \tfrac{r_{it}L_{Git}}{\frac{1}{\alpha_G}r_{it}L_{Git}+\frac{1}{\alpha_F}r_{it}L_{Fit}} + \tfrac{r_{it}L_{Fit}}{\frac{1}{\alpha_G}r_{it}L_{Git}+\frac{1}{\alpha_F}r_{it}L_{Fit}},\\ &= \alpha_G\left(\tfrac{\frac{1}{\alpha_G}\ell^A_{Git}}{\frac{1}{\alpha_G}\ell^A_{Git}+\frac{1}{\alpha_F}\ell^A_{Fit}}\right) + \alpha_F\left(\tfrac{\frac{1}{\alpha_F}\ell^A_{Fit}}{\frac{1}{\alpha_G}\ell^A_{Git}+\frac{1}{\alpha_F}\ell^A_{Fit}}\right).\end{aligned} \tag{D.1}$$

where we used the property that the wage (w_{it}) and expected rental rate (r_{it}) are equalized between arable and pastoral farming within each location. Our model also implies:

$$\frac{\frac{1-\alpha_G}{\alpha_G}}{\frac{1-\alpha_F}{\alpha_F}} = \frac{N_{Git}/L_{Git}}{N_{Fit}/L_{Fit}}. \tag{D.2}$$

From Table 23 of Feinstein (1972), the share of rent for farm land and buildings in total farm income in the United Kingdom in 1855 was $\alpha_A = 45/146 = 0.31$. From the data on farm area and employment in the individual-level 1851 population census, the relative labor intensity of arable to pastoral farming ($\frac{N_G/L_G}{N_F/L_F}$) is around 1.5. Therefore, we choose (α_G, α_F) to satisfy:

$$\alpha_G\left(\frac{\frac{1}{\alpha_G}\ell^A_G}{\frac{1}{\alpha_G}\ell^A_G+\frac{1}{\alpha_F}\ell^A_F}\right) + \alpha_F\left(\frac{\frac{1}{\alpha_F}\ell^A_F}{\frac{1}{\alpha_G}\ell^A_G+\frac{1}{\alpha_F}\ell^A_F}\right) - 0.31 = 0, \tag{D.3}$$

$$\frac{\frac{1-\alpha_G}{\alpha_G}}{\frac{1-\alpha_F}{\alpha_F}} - 1.5 = 0, \tag{D.4}$$

where ℓ^A_G and ℓ^A_F are the aggregate shares of agricultural land used for arable and pastoral farming, respectively. From these two empirical moments, we obtain a calibrated arable land share of $\alpha_G = 0.25$ and a calibrated pastoral land share of $\alpha_F = 0.34$.

Manufacturing and Services We assume a common share of land in production costs in manufacturing and services ($\alpha_M = \alpha_S = \alpha_N$), such that we focus on the differences in factor intensity between agriculture and non-agriculture.

We calibrate this common value for α_N such that the model's predictions are consistent with (i) our calibrated agricultural land share of $\alpha_A = 0.31$ from Feinstein (1972); (ii) the aggregate share of land in national income; (iii) the shares of agriculture, manufacturing and services (including housing) in national income. Note that the aggregate share of land in national income (α) can be written as follows:

$$\alpha_t \equiv \frac{r_t L}{Y_t} = \frac{r_t L_{At}}{Y_t} + \frac{r_t L_{Mt}}{Y_t} + \frac{r_t L_{St}}{Y_t},$$

$$\alpha_t = \frac{r_t L_{At}}{Y_{At}} \frac{Y_{At}}{Y_t} + \frac{r_t L_{Mt}}{Y_{Mt}} \frac{Y_{Mt}}{Y_t} + \frac{r_t L_{St}}{Y_{St}} \frac{Y_{St}}{Y_t},$$
$$\alpha_t = \alpha_A \frac{Y_{At}}{Y_t} + \alpha_N \left(\frac{Y_{Mt}}{Y_t} + \frac{Y_{St}}{Y_t} \right).$$

Re-arranging this relationship, we have:

$$\alpha_N = \frac{\alpha_t - \alpha_A \frac{Y_{At}}{Y_t}}{\frac{Y_{Mt}}{Y_t} + \frac{Y_{St}}{Y_t}}. \tag{D.5}$$

From Table 1 of Feinstein (1972), the share of rent in domestic income in 1855 was $\alpha_t = \frac{93}{629} = 0.15$. From Table 23 of Feinstein (1972), the share of rent for farm land and buildings in total farm income in the United Kingdom in 1855 was $\alpha_A = \frac{45}{146} = 0.31$, as noted above.

From Table 37 in Deane and Cole (1967), the share of agriculture in domestic income in 1851 was: $\frac{Y_{At}}{Y_t} = \frac{106.5}{453.8} = 0.23$. From the same table, the corresponding share of manufacturing and services (including housing) in domestic income in 1851 was:

$$\frac{Y_{Mt}}{Y_t} + \frac{Y_{St}}{Y_t} = \frac{453.8 - 106.5}{453.8} = \frac{347.3}{453.8} = 0.77.$$

Substituting these values into equation (D.5), we recover our calibrated value of α_N as:

$$\alpha_N = \frac{0.15 - (0.31 \times 0.23)}{0.77} = \frac{0.15 - 0.07}{0.77} = \frac{0.08}{0.77} = 0.10. \tag{D.6}$$

D1.3 Productivity Dispersion Parameters

The parameter ϵ controls the dispersion of agricultural productivity draws, which determines the elasticity of the arable and pastoral shares of agricultural land with respect to changes in the relative prices of these goods. We assume an elasticity of $\epsilon = 1.658$ based on the estimate using agricultural land use data in Sotelo (2020).

The parameter θ regulates the dispersion of sectoral productivity, which determines the elasticity of the agricultural, manufacturing and services land shares with respect to changes in the relative prices of these goods. Given the substantial heterogeneity in land suitability across these different sectors, we assume a lower elasticity of $\theta = 1.2$, which ensures greater substitutability of land between alternative uses within the agricultural sector than across sectors.

Finally, the parameter χ governs the dispersion of idiosyncratic preferences, which determines the elasticity of population with respect to real income in each location. We assume a value of $\chi = 2$ for this migration elasticity, which lies in the center of the range of estimates of this parameter in Bryan and Morten (2019), Fajgelbaum et al. (2019), and Galle et al. (2023).

D2 Model Inversion

In this subsection, we provide further details on our model inversion from Section 6.2 of the paper, including the proofs of Propositions 1 and 2.

In Proposition 1, we show that we observe enough information in the historical data to recover all of the model's endogenous variables required for our counterfactuals for the Grain Invasion below. We use the observed historical data and the equilibrium conditions of the model to back out the implied values of the unobserved endogenous variables that are consistent with the data being an equilibrium of the model.

In Proposition 2, we show that our model allows for unobserved changes over time in sectoral productivity for each location, agricultural productivity for each disaggregated good and location, and amenities for each origin and destination, such that the model exactly matches the observed data on the endogenous variables as an equilibrium outcome. Given the observed data on these endogenous variables, we use the model's equilibrium conditions to solve for unique values of prices adjusted for amenities and productivities for which the observed data are an equilibrium of the model.

D2.1 Proof of Proposition 1

Proposition. ***(Unobserved Endogenous Variables, Proposition 1 in the Paper)*** *Given the demand parameters (σ, β_G, β_F), production cost parameters (α_k), productivity and preference dispersion parameters (ϵ, θ, χ) and data by location i and year t on employment by sector (N_{Ait}, N_{Mit}, N_{Sit}), agricultural land shares for arable and pastoral farming (ℓ^A_{Git}, ℓ^A_{Fit}), total land area (L_i), and rateable values ($\mathbb{V}_{it} = r_{it}L_i$), there exist unique values for unobserved rental rates (r_{it}), wages (w_{it}), sectoral land shares (ℓ_{Ait}, ℓ_{Mit}, ℓ_{Sit}), employment in arable and pastoral farming (N_{Git}, N_{Fit}), and sectoral expenditure shares (x_{Ait}, x_{Mit}, x_{Sit}) that are consistent with the observed data being an equilibrium of the model.*

Proof. Given the demand parameters (σ, β_G, β_F), production cost parameters (α_k), productivity and preference dispersion parameters (ϵ, θ, χ) and data by location i and year t on employment by sector (N_{Ait}, N_{Mit}, N_{Sit}), agricultural land shares for arable and pastoral farming (ℓ^A_{Git}, ℓ^A_{Fit}), total land area (L_i), and rateable values ($\mathbb{V}_{it} = r_{it}L_i$), we use the equilibrium conditions of the model to recover the unobserved endogenous variables (r_{it}, α_{Ait}, w_{it}, ℓ_{Ait}, ℓ_{Mit}, ℓ_{Sit}, N_{Git}, N_{Fit}). Our quantitative analysis proceeds in the following steps.

Step 1: Rental Rates (r_{it})

From observed rateable values ($\mathbb{V}_{it} = r_{it}L_i$) and observed land area (L_i), we recover the implied rental rate:

$$r_{it} = \frac{\mathbb{V}_{it}}{L_i}. \tag{D.7}$$

Step 2: Agricultural Land Payments Share (α_{Ait})

We solve for the endogenous share of land in production costs for the agricultural sector as a whole in each location (α_{Ait}) using the observed agricultural land shares for arable and pastoral farming (ℓ^A_{Git}, ℓ^A_{Fit}) and our calibrated land cost shares for arable and pastoral farming (α_G, α_F). The share of land in production costs for the agricultural sector as a whole (α_{Ait}) is defined as:

$$\alpha_{Ait} = \frac{r_{it} L_{Ait}}{R_{Ait}},$$

where R_{Ait} denotes total revenue in the agricultural sector. Using cost minimization, zero profits and the Cobb-Douglas production technologies for arable and pastoral farming, we have:

$$r_{it} L_{Ait} = \alpha_G R_{Git} + \alpha_F R_{Fit},$$

where R_{Git} and R_{Fit} denote revenue in arable and pastoral farming, respectively. Using this expression in the definition of the share of land in production costs for the agricultural sector as a whole (α_{Ait}), we have:

$$\alpha_{Ait} = \alpha_G \frac{R_{Git}}{R_{Ait}} + \alpha_F \frac{R_{Fit}}{R_{Ait}}. \tag{D.8}$$

Now note that zero profits in arable and pastoral farming implies that the share of disaggregated agricultural good $g \in \{G, F\}$ in agricultural revenue can be written as:

$$\frac{R_{git}}{R_{Ait}} = \frac{w_{it} N_{git} + r_{it} L_{git}}{\sum_{h \in \{G,F\}} w_{it} N_{hit} + r_{it} L_{hit}}.$$

Additionally, cost minimization, zero profits and Cobb-Douglas production technologies for each disaggregated agricultural good $g \in \{G, F\}$ imply:

$$w_{it} N_{git} = \frac{1 - \alpha_g}{\alpha_g} r_{it} L_{git}.$$

Therefore the share of disaggregated agricultural good $g \in \{G, F\}$ in agricultural revenue can be written as:

$$\frac{R_{git}}{R_{Ait}} = \frac{\frac{1}{\alpha_g} r_{it} L_{git}}{\frac{1}{\alpha_G} r_{it} L_{Git} + \frac{1}{\alpha_F} r_{it} L_{Fit}}.$$

which simplifies to:

$$\frac{R_{git}}{R_{Ait}} = \frac{\frac{1}{\alpha_g} L_{git}}{\frac{1}{\alpha_G} L_{Git} + \frac{1}{\alpha_F} L_{Fit}}.$$

Using this result in equation (D.8), the share of land in production costs for the agricultural sector as a whole can be written as:

$$\alpha_{Ait} = \alpha_G \left(\frac{\frac{1}{\alpha_G} L_{Git}}{\frac{1}{\alpha_G} L_{Git} + \frac{1}{\alpha_F} L_{Fit}} \right) + \alpha_F \left(\frac{\frac{1}{\alpha_F} L_{Fit}}{\frac{1}{\alpha_G} L_{Git} + \frac{1}{\alpha_F} L_{Fit}} \right).$$

Dividing the numerator and denominator of the two fractions on the right-hand side by agricultural land area (L_{it}^{A}), we can equivalently write the share of land in production costs for the agricultural sector as a whole as:

$$\alpha_{Ait} = \alpha_G \left(\frac{\frac{1}{\alpha_G} \ell_{Git}^{A}}{\frac{1}{\alpha_G} \ell_{Git}^{A} + \frac{1}{\alpha_F} \ell_{Fit}^{A}} \right) + \alpha_F \left(\frac{\frac{1}{\alpha_F} \ell_{Fit}^{A}}{\frac{1}{\alpha_G} \ell_{Gi}^{A} + \frac{1}{\alpha_F} \ell_{Fit}^{A}} \right), \tag{D.9}$$

where the agricultural land shares for arable and pastoral farming (ℓ_{Git}^{A}, ℓ_{Fit}^{A}) are observed in the data. Therefore, given the observed data on agricultural land shares for arable and pastoral farming (ℓ_{Git}^{A}, ℓ_{Fit}^{A}) and the calibrated parameters (α_G, α_F), we recover the endogenous share of land in production costs for the agricultural sector as a whole in each location (α_{Ait}).

Step 3: Wages (w_{it})

We solve for wages (w_{it}) using the land market clearing condition that equates observed rateable values ($\mathbb{V}_{it} = r_{it} L_i$) with the sum of the payments for land use. Using cost minimization, zero profits and the Cobb-Douglas production technology, we can write this land market clearing condition as follows:

$$\mathbb{V}_{it} = \frac{\alpha_{Ait}}{1 - \alpha_{Ait}} w_{it} N_{Ait} + \frac{\alpha_N}{1 - \alpha_N} w_{it} \left[N_{it} - N_{Ait} \right]. \tag{D.10}$$

Re-arranging this land market clearing condition, we obtain the following closed-form solution for wages:

$$w_{it} = \frac{\mathbb{V}_{it}}{\frac{\alpha_{Ait}}{1 - \alpha_{Ait}} N_{Ait} + \frac{\alpha_N}{1 - \alpha_N} \left[N_{it} - N_{Ait} \right]}, \tag{D.11}$$

where the parameter $\alpha_N = \alpha_M = \alpha_S$ is known; we solved for α_{Ait} in the previous step; and we observe rateable values ($\mathbb{V}_{it}$), agricultural employment (N_{Ait}), and total employment (N_{it}).

Step 4: Sectoral Land Use (L_{Ait}, L_{Mit}, L_{Sit})

We solve for sectoral land use using cost minimization, zero profits, and the Cobb-Douglas production technologies:

$$L_{Ait} = \frac{\alpha_{Ait}}{1 - \alpha_{Ait}} \frac{w_{it}}{r_{it}} N_{Ait},$$

$$L_{Mit} = \frac{\alpha_N}{1 - \alpha_N} \frac{w_{it}}{r_{it}} N_{Mit},$$

$$L_{Sit} = \frac{\alpha_N}{1 - \alpha_N} \frac{w_{it}}{r_{it}} N_{Sit},$$

where the parameter $\alpha_N = \alpha_M = \alpha_S$ is known; we solved for (α_{Ait}, w_{it}, r_{it}) in previous steps; and we observe employment by sector (N_{Ait}, N_{Mit}, N_{Sit}). From these solutions for sectoral land use, we immediately recover sectoral land shares ($\ell_{Ait} = L_{Ait}/L_i$, $\ell_{Mit} = L_{Mit}/L_i$, $\ell_{Sit} = L_{Sit}/L_i$); and the shares of arable and pastoral farming in total land area ($\ell_{Git} = \ell_{Git}^{A} \ell_{Ait}$, $\ell_{Fit} = \ell_{Fit}^{A} \ell_{Ait}$).

Step 5: Employment in Arable and Pastoral Farming (N_{Git}, N_{Fit})

We solve for employment for each disaggregated agricultural good using cost minimization, zero profits, and the Cobb-Douglas production technologies, which imply:

$$N_{Git} = \frac{1-\alpha_G}{\alpha_G} \frac{r_{it}}{w_{it}} \ell_{Git}^{A} \ell_{Ait} L_i, \tag{D.12}$$

$$N_{Fit} = \frac{1-\alpha_F}{\alpha_F} \frac{r_{it}}{w_{it}} \ell_{Fit}^{A} \ell_{Ait} L_i, \tag{D.13}$$

where the parameters (α_G, α_F) are known; we solved for (r_{it}, w_{it}, ℓ_{Ait}) in previous steps; and we observe (ℓ_{Git}^{A}, ℓ_{Fit}^{A}, L_i).

Step 6: Adjusted Relative Prices ($\widetilde{P}_{Ait}$, 1, $\widetilde{P}_{Sit}$)

We express adjusted prices for each sector relative to the manufacturing sector:

$$\begin{aligned} \widetilde{P}_{Ait} &= \left(\beta_A P_{Ait}\right) / \left(\beta_M P_{Mit}\right), \\ \widetilde{P}_{Mit} &= 1 \\ \widetilde{P}_{Sit} &= \left(\beta_S P_{Sit}\right) / \left(\beta_M P_{Mit}\right). \end{aligned}$$

We solve for the adjusted relative agricultural price index ($\widetilde{P}_{Ait}$) and services price ($\widetilde{P}_{Sit}$) using expenditure minimization, profit maximization, zero profits, market clearing for non-traded services, and income equals expenditure. From profit maximization and zero profits, revenue in each sector is:

$$\begin{aligned} R_{kit} &= w_{it} N_{kit} + r_{it} L_{kit}, \\ &= \frac{w_{it} N_{kit}}{1-\alpha_k}. \end{aligned}$$

From expenditure minimization, expenditure in each sector is:

$$E_{kit} = \frac{\left(\beta_k P_{kit}\right)^{1-\sigma}}{\sum_{k \in \{A,M,S\}} \left(\beta_m P_{mit}\right)^{1-\sigma}} \left[w_{it} N_{it} + r_{it} L_i\right],$$

which can be written as:

$$E_{kit} = \frac{\left(\widetilde{P}_{kit}\right)^{1-\sigma}}{1 + \sum_{k \in \{A,S\}} \left(\widetilde{P}_{mit}\right)^{1-\sigma}} \left[w_{it} N_{it} + r_{it} L_i\right], \qquad k \in \{A, S\},$$

$$E_{Mit} = \frac{1}{1 + \sum_{k \in \{A,S\}} \left(\widetilde{P}_{mit}\right)^{1-\sigma}} \left[w_{it} N_{it} + r_{it} L_i\right].$$

Using market clearing for non-traded services, we obtain:

$$\frac{w_{it} N_{Sit}}{1-\alpha_S} = \frac{\left(\widetilde{P}_{Sit}\right)^{1-\sigma}}{1 + \sum_{m \in \{A,S\}} \left(\widetilde{P}_{mit}\right)^{1-\sigma}} \left[w_{it} N_{it} + r_{it} L_i\right]. \tag{D.14}$$

Using income equals expenditure, and using market clearing for non-traded services to eliminate the revenue and expenditure terms for that sector, we obtain:

$$\frac{w_{it}N_{Ait}}{1-\alpha_{Ait}} + \frac{w_{it}N_{Mit}}{1-\alpha_M} = \begin{bmatrix} \frac{\left(\widetilde{P}_{Ait}\right)^{1-\sigma}}{1+\sum_{m\in\{A,S\}}\left(\widetilde{P}_{kit}\right)^{1-\sigma}}\left[w_{it}N_{it}+r_{it}L_i\right] \\ \frac{1}{1+\sum_{m\in\{A,S\}}\left(\widetilde{P}_{kit}\right)^{1-\sigma}}\left[w_{it}N_{it}+r_{it}L_i\right] \end{bmatrix}. \tag{D.15}$$

Equations (D.14) and (D.15) provide a system of $2 \times N$ equations that determines unique values of the $2 \times N$ unknown relative adjusted prices ($\widetilde{P}_{Ait}$, $\widetilde{P}_{Sit}$), given the parameters ($\alpha_M = \alpha_S = \alpha_N$) and the observed data on employment by sector (N_{Ait}, N_{Mit}, N_{Sit}) and total land area (L_i), and our solutions for the agriculture cost share (α_{Ait}), wages (w_{it}), and rental rates (r_{it}) from previous steps. Given these solutions for relative adjusted prices ($\widetilde{P}_{Ait}$, $\widetilde{P}_{Sit}$), we recover the implied expenditure shares for each sector and location:

$$x_{Ait} = \frac{\left(\widetilde{P}_{Ait}\right)^{1-\sigma}}{1+\sum_{k\in\{A,S\}}\left(\widetilde{P}_{mit}\right)^{1-\sigma}},$$
$$x_{Mit} = \frac{1}{1+\sum_{k\in\{A,S\}}\left(\widetilde{P}_{mit}\right)^{1-\sigma}},$$
$$x_{Sit} = \frac{\left(\widetilde{P}_{Sit}\right)^{1-\sigma}}{1+\sum_{k\in\{A,S\}}\left(\widetilde{P}_{mit}\right)^{1-\sigma}}.$$

□

D2.2 Proof of Proposition 2

Proposition. ***(Model Inversion, Proposition 2 in the paper)*** *Given the demand parameters (σ, β_G, β_F), production cost parameters (α_k), productivity and preference dispersion parameters (ϵ, θ, χ) and data by location i and year t on employment by sector (N_{Ait}, N_{Mit}, N_{Sit}), agricultural land shares for arable and pastoral farming (ℓ^A_{Git}, ℓ^A_{Fit}), total land area (L_i), and rateable values ($\mathbb{V}_{it} = r_{it}L_i$), there exist unique values of productivity-adjusted prices for each disaggregated agricultural good ($\mathbb{E}_{git}$), productivity-adjusted prices for each sector ($\mathbb{T}_{kit}$), and amenity-adjusted aggregate prices ($\mathbb{B}_{it}$) (up to scale) that are consistent with the data being an equilibrium of the model.*

Proof. Given the demand parameters (σ, β_G, β_F), production cost parameters (α_k), productivity and preference dispersion parameters (ϵ, θ, χ) and data by location i and year t on employment by sector (N_{Ait}, N_{Mit}, N_{Sit}), agricultural land shares for arable and pastoral farming (ℓ^A_{Git}, ℓ^A_{Fit}), total land area (L_i), and rateable values ($\mathbb{V}_{it} = r_{it}L_i$), we use the equilibrium conditions of the model to recover unique values of productivity-adjusted prices for each disaggregated agricultural good ($\mathbb{E}_{git}$), productivity-adjusted prices for each sector ($\mathbb{T}_{kit}$), and amenity-adjusted aggregate prices ($\mathbb{B}_{it}$) (up to scale) that are consistent with the data being an equilibrium of the model. Our quantitative analysis proceeds in the following steps.

Step 1: Unobserved Endogenous Variables (r_{it}, α_{Ait}, w_{it}, ℓ_{Ait}, ℓ_{Mit}, ℓ_{Sit}, N_{Git}, N_{Fit})

We recover unique values for the unobserved endogenous variables (r_{it}, α_{Ait}, w_{it}, ℓ_{Ait}, ℓ_{Mit}, ℓ_{Sit}, N_{Git}, N_{Fit}) from the Proof of Proposition 1 in Online Appendix D2.1 above.

Step 2: Productivity-Adjusted Prices for Each Disaggregated Agricultural Good ($\mathbb{E}_{git}$)

We recover unique values for productivity-adjusted prices for each disaggregated agricultural good ($\mathbb{E}_{git}$) (up to scale) from the observed shares of agricultural land allocated to arable and pastoral farming (ℓ^A_{Git}, ℓ^A_{Fit}). We normalize productivity-adjusted prices for pastoral farming to one ($\mathbb{E}_{Fit} \equiv E_{Fit} P_{Fit}^{\epsilon/\alpha_F} = 1$). Given this normalization, we recover unique values for productivity-adjusted prices for arable farming ($\mathbb{E}_{Git} \equiv E_{Git} P_{Git}^{\epsilon/\alpha_F}$) from equation (10) for agricultural land shares in the paper:

$$\ell^A_{Git} = \frac{\mathbb{E}_{Git} w_{it}^{-\epsilon/\alpha_G}}{\mathbb{E}_{Git} w_{it}^{-\epsilon/\alpha_G} + w_{it}^{-\epsilon/\alpha_F}}, \tag{D.16}$$

where the parameters (α_G, α_F, ϵ) are known; ℓ^A_{Git} is observed; and we solved for w_{it} in Step 1.

Step 3: Productivity-Adjusted Prices for Each Sector ($\mathbb{T}_{ki}$)

We recover unique values for productivity-adjusted prices for each sector ($\mathbb{T}_{git}$) (up to scale) from our solutions for sectoral land shares (ℓ_{kit}) from Proposition 1 above. We normalize sectoral productivity for agriculture to one ($T_{Ait} = 1$). Given this normalization, we recover unique values for productivity-adjusted prices ($\mathbb{T}_{kit} = T_{kit} P_{kit}^{\theta/\alpha_k}$) for each of the other sectors $k \in \{M, S\}$ from equation (14) for land shares in the paper:

$$\ell_{kit} = \frac{\mathbb{T}_{kit} w_{it}^{-\theta/\alpha_k}}{\sum_{k\in\{M,S\}} \mathbb{T}_{kit} w_{it}^{-\theta/\alpha_k} + \gamma_\epsilon^\theta \left[\mathbb{E}_{Git} w_{it}^{-\epsilon/\alpha_G} + w_{it}^{-\epsilon/\alpha_F}\right]^{\frac{\theta}{\epsilon}}}, \qquad k \in \{M, S\}, \tag{D.17}$$

where the parameters (α_G, α_F, $\alpha_M = \alpha_S = \alpha_N$, ϵ, θ) are known; we used $\mathbb{E}_{Fit} \equiv E_{Fit} P_{Fit}^{\epsilon/\alpha_F} = 1$ from Step 2 above; and we solved for (ℓ_{Mit}, ℓ_{Sit}, w_{it}, $\mathbb{E}_{Git}$) in previous steps.

Step 4: Amenity-adjusted Aggregate Prices

We recover unique values for amenity-adjusted aggregate prices ($\mathbb{B}_{it}$) (up to scale) from our observed data on total employment shares (λ_{it}). We normalize the geometric mean of amenity-adjusted aggregate prices to one ($\left(\prod_{m\in\mathcal{J}} \mathbb{B}_{mt}\right)^{\frac{1}{J}}$, where $J = |\mathcal{J}|$). Given this normalization, we recover unique values for amenity-adjusted aggregate prices ($\mathbb{B}_{it} \equiv (B_{it}/P_{Cit})^{\chi}$) from equation (19) for total employment shares in the paper:

$$\lambda_{it} = \frac{\mathbb{B}_{it} w_{it}^{\chi}}{\sum\limits_{m\in\mathcal{J}} \mathbb{B}_{mt} w_{mt}^{\chi}},$$

where the parameter χ is known; total employment shares (λ_{it}) are observed; and we solved for wages (w_{it}) in Step 1. □

D3 Counterfactuals

In this subsection, we provide further details on our counterfactuals for the impact of the Grain Invasion in Subsection 6.3 of the paper, including the Proof of Proposition 3. We evaluate the impact of the Grain Invasion as a fall in the world price of arable products, holding constant other exogenous determinants of economic activity.

We undertake these counterfactuals starting from the observed equilibrium in the data in 1851, immediately following the Repeal of the Corn Laws, and the first year for which we observe employment by sector and location. We denote the value of variable in the counterfactual equilibrium by a prime (e.g., x'_j), the value of a variable in the observed equilibrium in the data without a prime (e.g., x_j), and the relative changes in variables between the counterfactual and observed equilibria by a hat (e.g., $\widehat{x}_j \equiv x'_j/x_j$). We show that we can re-write the counterfactual equilibrium conditions in the model in terms of the observed values of the endogenous variables in the initial equilibrium in the data and the counterfactual relative changes in variables.

D3.1 Proof of Proposition 3

Proposition. ***(Exact-Hat Algebra, Proposition 3 in the paper)*** *Given the demand parameters (σ, β_G, β_F), production cost parameters (α_k), productivity and preference dispersion parameters (ϵ, θ, χ), and data by location i and year t on employment by sector (N_{Ait}, N_{Mit}, N_{Sit}), agricultural land shares for arable and pastoral farming (ℓ^A_{Git}, ℓ^A_{Fit}), total land area (L_i), rateable values ($\mathbb{V}_{it} = r_{it}L_i$), and counterfactual changes in world relative prices for traded goods ($\widehat{P}^*_{Gt}$, $\widehat{P}^*_{Ft}$, $\widehat{P}^*_{Mt}$), the solution for counterfactual changes in the model's endogenous variables does not require information on the level of the location characteristics (E_{git}, T_{kit}, τ_{kit}, B_{it}).*

Proof. We consider an exogenous counterfactual change in world prices for traded goods ($\widehat{P}_{Git} = \widehat{P}^*_{Gt}$, $\widehat{P}_{Fit} = \widehat{P}^*_{Ft}$, $\widehat{P}_{Mit} = \widehat{P}^*_{Mt}$). We hold constant productivities: $\widehat{E}_{git} = 1$ for $g \in \{G, F\}$ and $\widehat{T}_{kit} = 1$ for $k \in \{A, M, S\}$. We also hold constant trade costs: $\widehat{\tau}_{kit} = 1$ for $k \in \{G, F, M\}$. Given our assumed counterfactual change in exogenous traded goods prices ($\widehat{P}^*_{Gt}$, $\widehat{P}^*_{Ft}$, $\widehat{P}^*_{Mt}$), and initial guesses for changes in wages ($\widehat{w}^0_{it}$) and changes in the price of non-traded services ($\widehat{P}^0_{Sit}$), we solve the following system of general equilibrium conditions:

$$\widehat{\ell}^A_{Git}\ell^A_{Git} = \frac{\ell^A_{Git}\left(\widehat{P}^*_{Gt}\right)^{\epsilon/\alpha_G}\left(\widehat{w}^0_{it}\right)^{-\epsilon/\alpha_G}}{\ell^A_{Git}\left(\widehat{P}^*_{Gt}\right)^{\epsilon/\alpha_G}\left(\widehat{w}^0_{it}\right)^{-\epsilon/\alpha_G} + \ell^A_{Fit}\left(\widehat{P}^*_{Ft}\right)^{\epsilon/\alpha_F}\left(\widehat{w}^0_{it}\right)^{-\epsilon/\alpha_F}}, \tag{D.18}$$

$$\widehat{\ell}^A_{Fit}\ell^A_{Fit} = 1 - \widehat{\ell}^A_{Git}\ell^A_{Git}, \tag{D.19}$$

$$\widehat{\mathcal{P}}_{Ait} = \left[\ell^A_{Git}\left(\widehat{P}^*_{Gt}\right)^{\epsilon/\alpha_G}\left(\widehat{w}^0_{it}\right)^{-\epsilon/\alpha_G} + \ell^A_{Fit}\left(\widehat{P}^*_{Ft}\right)^{\epsilon/\alpha_F}\left(\widehat{w}^0_{it}\right)^{-\epsilon/\alpha_F}\right]^{\frac{1}{\epsilon}}, \tag{D.20}$$

$$\widehat{\mathcal{P}}_{Mit} = \left(\widehat{P}^*_{Mt}/\widehat{w}^0_{it}\right)^{1/\alpha_M}, \tag{D.21}$$

$$\widehat{\mathcal{P}}_{Sit} = \left(\widehat{P}^0_{Sit}/\widehat{w}^0_{it}\right)^{1/\alpha_S}, \tag{D.22}$$

$$\widehat{\ell}_{Ait}\ell_{Ait} = \frac{\ell_{Ait}\widehat{\mathcal{P}}^\theta_{Ait}}{\sum\limits_{k\in\{A,M,S\}}\ell_{kit}\widehat{\mathcal{P}}^\theta_{kit}}, \tag{D.23}$$

$$\widehat{\ell}_{Mit}\ell_{Mit} = \frac{\ell_{Mit}\widehat{\mathcal{P}}^\theta_{Mit}}{\sum\limits_{k\in\{A,M,S\}}\ell_{kit}\widehat{\mathcal{P}}^\theta_{kit}}, \tag{D.24}$$

$$\widehat{\ell}_{Sit}\ell_{Sit} = \frac{\ell_{Sit}\widehat{\mathcal{P}}^\theta_{Sit}}{\sum\limits_{k\in\{A,M,S\}}\ell_{kit}\widehat{\mathcal{P}}^\theta_{kit}}, \tag{D.25}$$

$$\widehat{L}_{Git} = \widehat{\ell}^A_{Git}\widehat{\ell}_{Ait}, \tag{D.26}$$

$$\widehat{L}_{Fit} = \widehat{\ell}^A_{Fit}\widehat{\ell}_{Ait}, \tag{D.27}$$

$$\widehat{L}_{Ait} = \widehat{\ell}_{Ait}, \tag{D.28}$$

$$\widehat{L}_{Mit} = \widehat{\ell}_{Mit}, \tag{D.29}$$

$$\widehat{L}_{Sit} = \widehat{\ell}_{Sit}, \tag{D.30}$$

$$\widehat{r}_{it} = \widehat{w}_{it}\left[\sum_{k\in\{A,M,S\}}\ell_{ki}\widehat{\mathcal{P}}^\theta_{kit}\right]^{\frac{1}{\theta}}, \tag{D.31}$$

$$\widehat{w}^0_{it}\widehat{N}_{Ait} = \frac{w_{it}N_{Git}}{w_{it}N_{Ait}}\widehat{r}_{it}\widehat{L}_{Git} + \frac{w_{it}N_{Fit}}{w_{it}N_{Ait}}\widehat{r}_{it}\widehat{L}_{Fit}, \tag{D.32}$$

$$\widehat{w}^0_{it}\widehat{N}_{Mit} = \widehat{r}_{it}\widehat{L}_{Mit}, \tag{D.33}$$

$$\widehat{w}^0_{it}\widehat{N}_{Sit} = \widehat{r}_{it}\widehat{L}_{Sit}, \tag{D.34}$$

$$\widehat{P}_{At} = \left(\widehat{P}^*_{Gt}\right)^{\beta_G}\left(\widehat{P}^*_{Ft}\right)^{1-\beta_G} \tag{D.35}$$

$$\widehat{P}_{Cit} = \left[x_{Ait}\widehat{P}^{1-\sigma}_{At} + x_{Mit}\left(\widehat{P}^*_{Mt}\right)^{1-\sigma} + x_{Sit}\left(\widehat{P}^0_{Sit}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}, \tag{D.36}$$

$$\widehat{x}_{Sit}x_{Sit} = \frac{x_{Sit}\left(\widehat{P}^0_{Sit}\right)^{1-\sigma}}{x_{Ait}\widehat{P}^{1-\sigma}_{At} + x_{Mit}\left(\widehat{P}^*_{Mt}\right)^{1-\sigma} + x_{Sit}\left(\widehat{P}^0_{Sit}\right)^{1-\sigma}}, \tag{D.37}$$

$$\widehat{\lambda}_{it}\lambda_{it} = \frac{\lambda_{it}\left(\widehat{w}_{it}^{0}/\widehat{P}_{Cit}\right)^{\chi}}{\sum_{m\in\mathcal{J}}\lambda_{mt}\left(\widehat{w}_{mt}^{0}/\widehat{P}_{Cmt}\right)^{\chi}}, \tag{D.38}$$

$$\widehat{N}_{it} = \widehat{\lambda}_{it}, \tag{D.39}$$

where we have re-written the system of general equilibrium conditions from Online Appendix C8 above in terms of the initial values of the endogenous variables and counterfactual changes in the endogenous variables. We also have used the following result for revenue in the non-traded services sector:

$$R_{Sit} \equiv P_{Sit}\left(N_{Sit}\right)^{1-\alpha_S}\left(\overline{z}_{Sit}\ell_{Sit}L_i\right)^{\alpha_S} = P_{Sit}\left(N_{Sit}\right)^{1-\alpha_S}\left(\gamma_\theta T_{Si}^{1/\theta}\ell_{Sit}^{\frac{\theta-1}{\theta}}L_i\right)^{\alpha_S}.$$

From solving this system of equations we obtain the following implied changes in wages ($\widehat{w}_{it}$) and the price of non-traded services ($\widehat{P}_{it}^{S}$) in each location:

$$\widehat{w}_{it} = \frac{1}{\widehat{N}_{it}}\left[\sum_{k\in\{G,F,M,S\}}\widehat{w}_{kt}^{0}\widehat{N}_{kit}\frac{w_{it}N_{kit}}{w_{it}N_{it}}\right], \tag{D.40}$$

$$\widehat{P}_{Sit} = \widehat{x}_{Sit}x_{Sit}\frac{\left[\widehat{r}_{it}\frac{r_{it}L_i}{r_{it}L_i+w_{it}N_{it}} + \widehat{w}_{it}^{0}\widehat{N}_{it}\frac{w_{it}N_{it}}{r_{it}L_i+w_{it}N_{it}}\right]}{\left(\widehat{N}_{Sit}\right)^{1-\alpha^S}\left(\widehat{\ell}_{Sit}\right)^{\frac{\alpha^S(\theta-1)}{\theta}}}. \tag{D.41}$$

We update our initial guesses for changes in wages ($\widehat{w}_{it}^{0}$) and changes in the price of non-traded services ($\widehat{P}_{Sit}^{0}$) until the implied changes in these variables in equations (D.40) and (D.41) are equal to our guesses, and all the other equilibrium conditions are satisfied.

Note that the system of general equilibrium conditions (D.18)-(D.41) is written solely in terms of the observed variables of the endogenous variables in an initial equilibrium and the counterfactual changes in the endogenous variables. Therefore, the solution to this system of equations for a counterfactual equilibrium does not require information on the unobserved level of the location characteristics (E_{git}, T_{kit}, τ_{kit}, B_{it}). □

E Theoretical Extensions

In this section of the Online Appendix, we report theoretical extensions of our baseline model from Section 5 of the paper. In Subsection E1, we introduce migration dynamics. In Subsection E2, we examine non-homothetic preferences. In Subsection E3, we consider a constant elasticity supply function for land and buildings.

E1 Migration Dynamics

To examine the long-run impact of the Grain Invasion, our baseline specification in Section 5 of the paper considers a static specification of worker location decisions. In this section of the Online Appendix, we consider a generalization of our theoretical model to incorporate migration dynamics following Artuç et al. (2010) and Caliendo et al. (2019).

At the beginning of each period t, the economy inherits a mass of workers in each location i (N_{it}). After supplying labor and consuming in each period t, workers observe idiosyncratic mobility shocks for each location. The value function for a worker living in location n in period t is equal to the current flow of utility from living in that location, plus the continuation value from the optimal choice of location:

$$V_{it} = \ln u_{it} + \max_{\{m\}_{m=1}^{m=J}} \left\{ \beta V_{mt+1} - \mu_{im} + \rho \epsilon_{mt} \right\}, \tag{E.1}$$

where β is the discount rate; u_{it} takes the same form as in equation (3) in the paper; μ_{im} is the cost of moving from location i to location m; ϵ_{mt} is the idiosyncratic mobility shock; and ρ regulates the relative importance of these idiosyncratic shocks.

We make the conventional assumption that idiosyncratic mobility shocks (ϵ_{mt}) are drawn from an extreme value distribution: $F(\epsilon) = e^{-e^{(-\epsilon - \overline{\gamma})}}$, where $\overline{\gamma}$ is the Euler-Mascheroni constant. Using standard properties of this extreme value distribution, the expected value of living in location i is:

$$v_{it} = \ln u_{it} + \rho \ln \left(\sum_{m=1}^{J} \exp\left(\beta v_{mt} - \mu_{im}\right)^{1/\rho} \right), \tag{E.2}$$

and the probability that a worker moves from location i to location m is :

$$\xi_{imt} = \frac{\exp\left(\beta v_{mt+1} - \mu_{im}\right)^{1/\rho}}{\sum_{j=1}^{J} \exp\left(\beta v_{jt+1} - \mu_{ij}\right)^{1/\rho}}. \tag{E.3}$$

Given these migration probabilities, the mass of workers in each location i (N_{it}) evolves according to the following dynamic equation:

$$N_{mt+1} = \sum_{i=1}^{J} \xi_{imt} N_{it}. \tag{E.4}$$

Therefore, a shock such as the Grain Invasion that affects the distribution of continuation values (v_{it}) across locations will lead to a gradual change in the distribution of total employment. Intuitively, workers wait to incur the mobility costs (μ_{im}) until they receive a favorable idiosyncratic shock (ϵ_{mt}), which gives rise to gradual adjustment to shocks.

E2 Non-homothetic Preferences

In our baseline specification in Section 5 of the paper, we assume homothetic constant elasticity of substitution (CES) preferences across sectors, as in a long line of research in macroeconomics on structural transformation following Baumol (1967). This assumption of homotheticity is primarily motivated by our quantitative application, for which only limited historical expenditure survey data are available. In this section of the Online Appendix, we generalize our theoretical model to allow for non-homothetic preferences following Comin et al. (2021). Although non-homotheticity introduces a new expenditure channel for distributional effects, geography remains an important dimension along which the income distributional effects of trade occur.

E2.1 Preferences

In particular, we generalize our specification of the upper tier of utility to the non-separable class of CES functions in Sato (1975), which satisfy implicit additivity in Hanoch (1975), as recently used in the macroeconomics literature in Comin et al. (2015) and Matsuyama (2019). To streamline notation, we drop the subscript for location throughout this section of the Online Appendix. The non-homothetic CES consumption index (C_t) is defined by the following implicit function:

$$\sum_{k\in\{A,M,S\}} \left(\frac{\beta_k C_{kt}}{C_t^{(\epsilon_k-\sigma)/(1-\sigma)}} \right)^{\frac{\sigma-1}{\sigma}} = 1, \tag{E.5}$$

where C_{kt} denotes consumption of sector k at time t; β_k is the CES weighting parameter for sector k; σ is the constant elasticity of substitution between sectors; ϵ_k is the constant elasticity of consumption of sector k with respect to the consumption index (C_t) that allows preferences to be non-homothetic. Our baseline homothetic CES specification corresponds to the special case of equation (E.5) in which $\epsilon_k = 1$ for all $k \in \{A, M, S\}$.

E2.2 Expenditure Minimization

The Lagrangian for the utility maximization problem is:

$$\mathcal{L} = C_t + \rho \left(1 - \sum_{k\in\{A,M,S\}} \left(\frac{\beta_k C_{kt}}{C_t^{(\epsilon_k-\sigma)/(1-\sigma)}} \right)^{\frac{\sigma-1}{\sigma}} \right) + \lambda \left(X_t - \sum_{k\in\{A,M,S\}} P_{kt} C_{kt} \right). \tag{E.6}$$

The first-order condition with respect to consumption of each sector (C_{kt}) can be written as:

$$P_{kt} C_{kt} = \frac{\rho}{\lambda} \left(\frac{1-\sigma}{\sigma} \right) \zeta_{kt}, \tag{E.7}$$

where we define ζ_{kt} as:

$$\zeta_{kt} \equiv \left(\frac{\beta_k C_{kt}}{C_t^{(\epsilon_k-\sigma)/(1-\sigma)}} \right)^{\frac{\sigma-1}{\sigma}}. \tag{E.8}$$

From the first-order condition (E.7) and utility function (E.5), total expenditure is given by:

$$X_t = \sum_{k \in \{A,M,S\}} P_{kt} C_{kt} = \frac{1-\sigma}{\sigma} \frac{\rho}{\lambda}. \tag{E.9}$$

Using this result in the first-order condition (E.7), we find that ζ_{kt} equals the share of sector k in expenditure at time t:

$$x_{kt} = \frac{P_{kt} C_{kt}}{X_t} = \zeta_{kt} = \left(\frac{\beta_k C_{kt}}{C_t^{(\epsilon_k - \sigma)/(1-\sigma)}} \right)^{\frac{\sigma-1}{\sigma}}. \tag{E.10}$$

Re-arranging this relationship, we obtain the demand function for sector k:

$$C_{kt} = (\beta_k)^{\sigma-1} \left(\frac{P_{kt}}{X_t} \right)^{-\sigma} (C_t)^{\epsilon_k - \sigma} = (\beta_k)^{\sigma-1} \left(\frac{P_{kt}}{P_t} \right)^{-\sigma} C_t^{\epsilon_k}, \tag{E.11}$$

which highlights that ϵ_k controls the elasticity of demand for sector k with respect to the real consumption index (C_t). Using this demand function (E.11), the expenditure share (E.10) can be re-written as:

$$x_{kt} = (\beta_k)^{\sigma-1} \left(\frac{P_{kt}}{P_t} \right)^{1-\sigma} C_t^{\epsilon_k - 1}. \tag{E.12}$$

Additionally, using the CES demand function (E.11) in utility in equation (E.5), we can solve for the expenditure function:

$$X_t = P_t C_t = \left[\sum_{k \in \{A,M,S\}} \left(\frac{P_{kt}}{\beta_k} \right)^{1-\sigma} C_t^{\epsilon_k - \sigma} \right]^{\frac{1}{1-\sigma}}. \tag{E.13}$$

Therefore the consumption price index is given by:

$$P_t = \frac{1}{C_t} \left[\sum_{k \in \{A,M,S\}} \left(\frac{P_{kt}}{\beta_k} \right)^{1-\sigma} C_t^{\epsilon_k - \sigma} \right]^{\frac{1}{1-\sigma}}, \tag{E.14}$$

or equivalently:

$$P_t = \left[\sum_{k \in \{A,M,S\}} \left(\frac{P_{kt}}{\beta_k} \right)^{1-\sigma} (X_t / P_t)^{\epsilon_k - 1} \right]^{\frac{1}{1-\sigma}}. \tag{E.15}$$

Combining equations (E.12) and (E.15), the share of sector k in expenditure at time t can be written as:

$$x_{kt} = \frac{(P_{kt}/\beta_k)^{1-\sigma} (X_t/P_t)^{\epsilon_k - 1}}{\sum_{\ell \in \{A,M,S\}} (P_{\ell t}/\beta_\ell)^{1-\sigma} (X_t/P_t)^{\epsilon_\ell - 1}} = \frac{(P_{kt}/\beta_k)^{1-\sigma} (X_t/P_t)^{\epsilon_k - 1}}{P_t^{1-\sigma}}. \tag{E.16}$$

Equations (E.15) and (E.16) correspond to the non-homothetic generalizations of equations (5) and (8) in Section 5 of the paper, respectively. Given estimates of the parameters (σ, $\{\epsilon_k\}$), this non-homothetic specification can be implemented empirically. Although non-homotheticity introduces an additional expenditure channel for distributional effects, geography remains an important dimension along which the income distributional consequences of trade occur.

E3 Endogenous Supply of Floor Space

In our baseline specification in the paper, we assume that land and buildings are in perfectly inelastic supply. In this section of the Online Appendix, we consider an extension with a constant elasticity supply function for land and buildings.

We assume that floor space ($\mathbb{L}_j$) is produced using land (L_j) and capital (M_j) by a competitive construction sector using a Cobb-Douglas production technology:

$$\mathbb{L}_j = L_j^{\eta} M_j^{1-\eta}, \qquad 0 < \eta < 1, \tag{E.17}$$

where capital is assumed to be in perfectly elastic supply at a constant price Ξ. From the first-order condition for profit maximization, capital and land use are related as:

$$M_j = \left[\frac{(1-\eta)\, r_j}{\Xi} \right]^{\frac{1}{\eta}} L_j. \tag{E.18}$$

Substituting equilibrium capital use (E.18) into the floor space production technology (E.17), we obtain a constant elasticity supply function for floor space as in Saiz (2010):

$$\mathbb{L}_j = \varrho_j L_j, \qquad \varrho_j = r_j^{\mu} \left[\frac{1-\mu}{\Xi} \right]^{\frac{1-\eta}{\eta}}, \qquad \mu = \frac{1-\eta}{\eta}, \tag{E.19}$$

where this supply of floor space takes the same value across all land plots within a given location.

F Data Appendix

This section of the Online Appendix reports additional information on the data from Section 3 of the paper. In Subsection F1, we provide further details on the data sources and definitions. In Subsection F3, we discuss the harmonization of parishes and other administrative units over time. In Subsection F4, we summarize the classification of parishes into rural and urban parishes.

F1 Data Sources

In Subsection F1.1, we introduce our exogenous measure of exposure to the Grain Invasion. In subsection F1.2, we discuss our consistent population data. In Subsection F1.3, we present further information on our rateable values data. In Subsection F1.4, we provide further details on our data on poor law relief. In Subsection F2.2, we summarize our other geographic data.

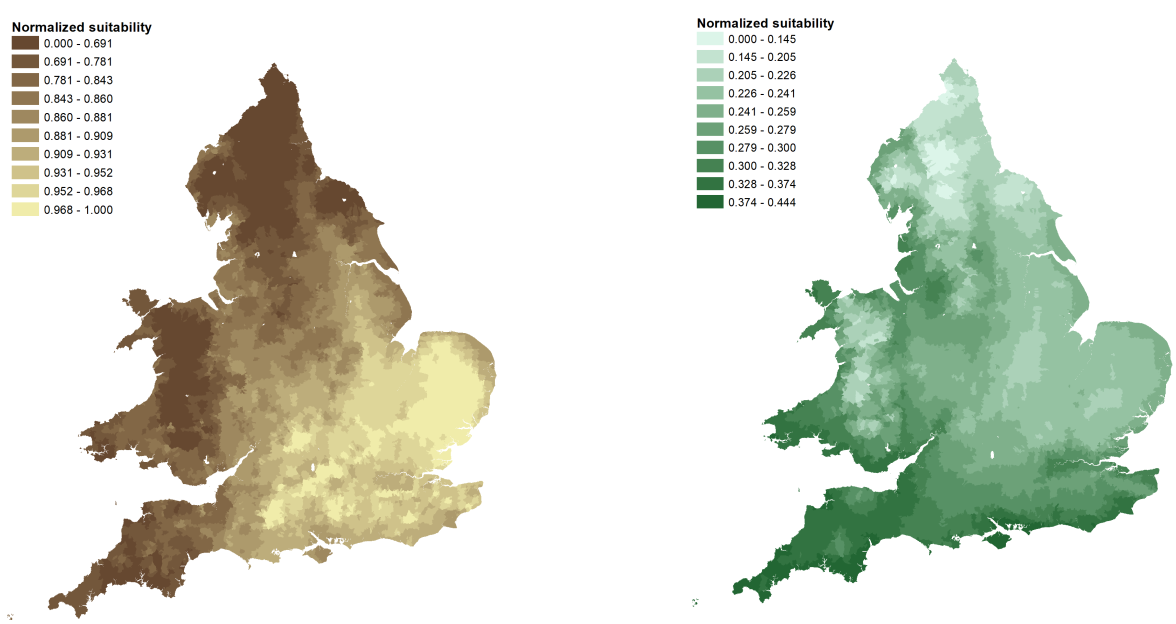
F1.1 Agro-climatic Conditions

Our measure of exogenous exposure to the Grain Invasion relies on variation in climate and soil conditions across England and Wales. The low-lying areas of West Britain are exposed to winds from the Atlantic and experience frequent rain all along the year. This climate is more conducive for pastoral farming. By contrast, Eastern Britain is drier and better suited for arable farming.

We capture these differences in suitability with data on the suitability of climate and soil conditions for the cultivation of wheat from the United Nations Food and Agricultural Organization Global Agro-Ecological Zones (GAEZ) dataset, as used in Costinot, Donaldson and Smith (2016). Although this GAEZ measure is computed for the period 1961-1990, these differences in relative agro-climatic conditions between the Western and Eastern parts of England and Wales have been stable for centuries.[3] Consistent with 19th-century farming practices in England and Wales, we use the measure of wheat suitability for low-input and rain-fed cultivation. The data come at a 1km resolution, which we map into our consistent parishes by using the average suitability across 1km cells and within parishes.

We illustrate the spatial variation in wheat suitability in Panel (a) of Figure F.1. In our empirical analysis, we also use grass suitability, as a measure of suitability for pastoral farming. Panel (b) of Figure F.1 shows the distribution of such suitability across England and Wales.

Figure F.1: Soil Suitability for Wheat and Grass Production Under Low Inputs



(a) Wheat

(b) Grass

Notes: Panel (a) displays the suitability for wheat production under low inputs across consistent parishes. Panel (b) displays the suitability for grass production under low inputs across consistent parishes.

[3]Despite this stability in relative agro-climatic conditions, there was an increase in overall temperature levels during the medieval warm period (900–1300 CE) and a decrease in overall temperature levels during the little ice age (1300–1850 CE), as discussed for example in Fagan (2019).

F1.2 Population Data

Population data from the population censuses of England and Wales from 1801–1891 was provided by the *Cambridge Group for the History of Population and Social Structure* (Cambridge Group), as documented in Wrigley (2011). The original sources for the population data are as follows:

- 1801 Census Report, Abstract of answers and returns, PP 1801, VI;
- 1811 Census Report, Abstract of answers and returns, PP 1812, XI;
- 1821 Census Report, Abstract of answers and returns, PP 1822, XV;
- 1831 Census Report, Abstract of the Population Returns of Great Britain, PP 1833, XXXVI to XXXVII;
- 1841 Census Report, Enumeration Abstract, PP 1843, XXII;
- 1851 Census Report: Population Tables, part II, vols. I to II, PP 1852-3, LXXXVIII, parts I to II;
- 1861 Census Report: Population tables, vol. II, PP 1863, LIII, parts I to II;
- 1871 Census Report: vol. III, Population abstracts: ages, civil condition, occupations and birthplaces of people, PP 1873, LXXI, part I;
- 1891 Census Report: vol. II, Area, Houses and Population: registration areas and sanitary districts, PP 1893-4, CV—which also includes the 1881 data, as used in our analysis.

Population data for 1901 stem from the Integrated Census Microdata Project (I-CeM). The majority of parishes did not experience any change in boundaries from 1891–1901. Therefore, we can simply extend the parish panel for 1801–1891 discussed above to 1901. However, from 1911 onwards, there are a number of major changes in parish boundaries. Most notably, the City of London (COL) consisted of more than 100 parishes in the censuses for 1801–1901, which were amalgamated into a single parish in 1907. Therefore, we end our sample period in 1901, to avoid the problems introduced by these changes in parish boundaries.

F1.3 Rateable Value Data

We measure the value of land and buildings using rateable values, which correspond to the annual flow of rent for the use of land and buildings, and equal the price times the quantity of floor space in the model. In particular, these rateable values correspond to "The annual rent which a tenant might reasonably be expected, taking one year with one another, to pay for a hereditament, if the tenant undertook to pay all usual tenant's rates and taxes [...] after deducting the probable annual average cost of the repairs, insurance and other expenses" (see Stamp 1920).

These rateable values have a long history in England and Wales, dating back to the 1601 Poor Relief Act, and were originally used to raise revenue for local public goods. These rateable values covered all categories of property, including public services (such as tramways, electricity works, etc.), government property (such as courts, parliaments, etc.), private property (including factories, warehouses, wharves, offices, shops, theaters, music halls, clubs, and all residential dwellings), and other property (including colleges and halls in universities, hospitals and other charity properties, public schools, and almshouses). All of these categories of properties were included, regardless of whether or not their owners were liable for taxation.

There were three categories of exemptions: (1) Crown property occupied by the Crown (Crown properties leased to other tenants are included); (2) places for divine worship (church properties leased to other tenants are included); (3) concerns listed under No. III Schedule A, namely (i) mines of coal, tin, lead, copper, mundic, iron, and other mines, (ii) quarries of stone, slate, limestone, or chalk, (iii) ironworks, gasworks, salt springs or works, alum mines or works, waterworks, streams of water, canals, inland navigations, docks, drains and levels, fishings, rights of markets and fairs, tolls, railways and other ways, bridges, ferries, and cemeteries.

Different types of rateable values can be distinguished, depending on the use of the revenue raised: Schedule A Income Taxation, Local Authority Rates, and Poor Law Rates. Where available, we use the Schedule A rateable values, since Schedule A is the section of the national income tax concerned with income from property and land, and these rateable values are widely regarded as corresponding most closely to market valuations. For example, Stamp (1920) argues that "It is generally acknowledged that the income tax, Schedule A, assessments are the best approach to the true values" (p.25). Where these Schedule A rateable values are not available, we use the Local Authority rateable values, Poor Law rateable values, or property valuations for income tax. For years in which more than one of these measures is available, we find that they are highly correlated with one another across parishes.

The original sources for rateable values are as follows:

- **1815**: Property valuations for income tax. Return to an address of the Honourable the House of Commons, dated 21 February 1854; House of Commons Papers, vol. LVI.1, paper no: 509.

- **1843**: Property valuations for income tax. Return to an address of the Honourable the House of Commons, dated 21 February 1854; House of Commons Papers, vol. LVI.1, paper no: 509.

- **1852**: Poor law rateable values. Return to an address of the Honourable the House of Commons, dated 21 February 1854; House of Commons Papers, vol. LVI.1, paper no: 509.
- **1865**: Poor law rateable values. Return Pursuant to an Address of the House of Lords, dated 24th March 1868 of the Parishes of England and Wales, Stating the Counties in which they are Situated; their Area; their Population and the Gross Estimated Rental, Ordered to be printed 24th March 1868. House of Lords Papers, vol. XVIII, paper no. 54.
- **1870**: Rateable values. Return to an Address of the Honourable The House of Commons, dated 28 April 1871; House of Commons Papers, vol. LV.329, paper no. 201.
- **1881**: Poor law rateable values. A Statement of the Names of the Several Unions And Poor Law Parishes In England And Wales; And of the Population, Area, And Rateable Value Thereof in 1881. London: Her Majesty's Stationery Office, 1887.
- **1896**: Schedule A rateable values. Agricultural Rates Act, 1896. Reports separate data on the rateable value of agricultural land and the rateable value of other land and buildings. Return to an order of the Honourable the House of Commons, dated 27 July 1897; House of Commons Papers, paper no: 368; 1897.

Rateable values are reported at the level of a parish in any given year. However, parish geographies change over time; the raw data cannot be immediately linked to our consistent mappable units between 1801 and 1901. To map rateable values into consistent spatial units over time, we proceed in two steps. We first geolocate each entry in the raw data and create a set of points characterized by a measure of local land value (the rateable value per acre within the local parish). We then use spatial interpolation methods to estimate a smooth rateable value raster at a resolution of 250 meters for all of England and Wales. From this estimated raster, we can calculate the average rateable value within consistent parish boundaries by multiplying the average rateable value per acre with the consistent parish's acreage.

The geolocation procedure works as follows:

1. We manually match the rateable value information reported at the parish level with the *CGKO* shapefile (Cambridge Group Kain Oliver) using both, parish and place names, as well as the corresponding registration district (poor law union) and county in which the reported parish or place is nested. The CGKO shapefile consists of roughly 23,000 spatial units and is derived from Kain and Oliver's digital maps of parish and township boundaries. The file maps data at the level of parishes, townships, or places from censuses collected between 1801–1891.

2. In a few cases, the parish units used to report rateable values do not match census parishes, for instance, because evaluators chose to aggregate information. In these cases, we look up the location manually and geolocate it. This leaves us with point coordinates for the specific rateable value payment.

3. Since parish boundaries change over time, we cannot be sure that the rateable value reported for a given parish also applies to the matched census parish. Therefore, we refrain from using the rateable value per parish and focus instead on the *rateable value per acre* assigned to either the parish centroid or the coordinates of a manually located entry. Conveniently, in all years except 1896, rateable values were reported along with the corresponding acreage of the parish they apply to. In 1896, we have to rely on the matched CGKO parishes unit's acreage. This procedure leaves us with 10,238 observation points for the years 1815, 1843 and 1852, with 13,698 points in 1865; 18,575 points in 1881; and 13,357 points in 1896 across England and Wales. Note that CGKO units are very accurate and in some cases, a rateable value reported for one parish may be linked to one census parish which is subdivided into multiple places and thus spatial units. In this case, we would assign the same rateable value per acre to each the centroid of each subdivision.[4]

4. In the last step, we log-transform the data to ensure that they are approximately normally distributed.

Our preferred spatial interpolation procedure is kriging (for a recent application to land prices, see Davis et al. (2021) and Larson and Shui (2022).[5] Before describing the kriging procedure, we introduce the general idea underlying spatial interpolation models.[6] Consider a series of N location centroids p_i, where we observe a rateable value $R(p_i)$. To create a smooth price surface, we want to use the observed rateable values to predict rateable values at additional observations $\hat{R}(p_j)$ with $j \notin N$. To do so, we estimate:

$$\hat{R}(p_j) = \sum_{i=1}^{n} \lambda_i R(p_{i(j)}) \tag{F.1}$$

where λ_i are weights such that $\hat{R}(p_j)$ is a weighted average of the rateable values in observed neighboring locations $R(p_{i(j)})$. Note that we use the notation $i(j)$ to indicate that i is located in the vicinity to j. The main difference across interpolation methods is how each of those computes the weights $\{\lambda_i\}$. Inverse Distance Weighting (IDW) assumes weights which are proportional to

[4]When we interpolate between these points, we account for clustering, as discussed further below.

[5]In robustness checks, we compare kriging with two other methods of spatial interpolation—inverse distance weighting and triangular irregular networks—using 20% holdout sample and computing the root mean square error (RMSE). We find that kriging performs significantly better than these other interpolation methods.

[6]For a more detailed discussion, we refer the interested reader to Sherman (2011).

the inverse distance, possibly raised to some power depending on how quickly one expects the spatial correlation to decay with distance. The second alternative method, Triangular Irregular Network (TIN), joins all observations into a mesh of triangles which contain no sampled points. The mesh covers the entire observed area and it is used to infer a triangulated surface of rateable values per acre.

By contrast, kriging uses optimization methods to calibrate the weights in Equation (F.1) to the spatial covariance structure of the observed points. Following the discussion in Sherman (2011), suppose that the data follow the process:

$$R(p) = \mu + \delta(p),$$

where μ is some constant and $\delta(p)$ is assumed to be Intrinsically Stationary (IS), i.e., the following two properties must hold for both $\delta(p)$ and $R(p)$:

1. $E[R(p) - \mu] = 0$ (the estimates must be centered around the unconditional mean),
2. $Var[R(p+h) - R(p)] := 2\gamma(h)$ for all h (the variance between any two points only depends on the distance h between these two points and not on direction or on a general region).

The function $\gamma(h)$ is called the semi-variogram function. Under the previous assumptions, the best linear estimator of an unobserved point $R(p_j)$ is the kriging estimator. The kriging estimator generates a set of optimal weights λ_i by minimizing:

$$\begin{gathered} E\left\{\left[\sum_{i=1}^{n} \lambda_i R(p_i) - R(p_0)\right]^2\right\} \\ \text{s.t.} \sum_{i=1}^{n} \lambda_i = 1. \end{gathered} \tag{F.2}$$

where the semivariogram $\gamma(h)$ is estimated in a first step from the relationship between the pairwise variances between locations i and j and the distance h_{ij}. This estimation depends on the number of neighbors and the assumed functional form for the function γ. As discussed in Sherman (2011), co-kriging is a natural extension of this procedure that supplements the rateable value information with more ubiquitous information about other characteristics correlated across space, e.g., elevation. The underlying idea is that elevation is informative about variation in rateable values. For example, locations with higher elevations arguably have lower rateable values because of their lower agricultural or construction potential. Incorporating such additional information helps improve the interpolation precision. Also, we sometimes observe clusters of points with rateable value information in urban areas where the average parish size is smaller. To account for uneven sampling density, we implement cell declustering (Li and Heap 2011). Cell

Table F.1: Interpolation RMSE (20% holdout sample)

Year	TIN	IDW	Kriging	Kriging Decluster	CoKriging	CoKriging Decluster
1815	0.569	0.543	0.532	0.538	0.531	0.526
1843	0.569	0.529	0.520	0.517	0.532	0.519
1852	0.553	0.528	0.525	0.524	0.517	0.515
1865	0.607	0.581	0.564	0.565	0.557	0.557
1881	0.614	0.602	0.579	0.572	0.574	0.568
1896	0.854	0.820	0.792	0.783	0.785	0.780
Mean	0.624	0.597	0.582	0.580	0.581	0.578

declustering works by overlaying a rectangular grid of cells over the data, then weighting each point inversely with the number of points in that cell. We set the declustering grid size to a rectangular grid that divides England and Wales into a 60-by-60 grid.

To shed some light on the quality of the resulting interpolation, we remove a random 20% validation sample from the data and perform the interpolation on the remaining 80% of data. Following Davis et al. (2021), we compute the root mean square error (RMSE) between each point in the validation sample and the predicted surface in a given year. We report the RMSE for each method and for each year of rateable value data in Table F.1. We find that kriging produces smaller errors than TIN or IDW, and consistently so across years (in line with the findings in Davis et al. 2021). There is a small, but detectable, increase in precision when we use co-kriging or declustering. Since declustered co-kriging provides the best results, we use it in our baseline interpolation procedure.

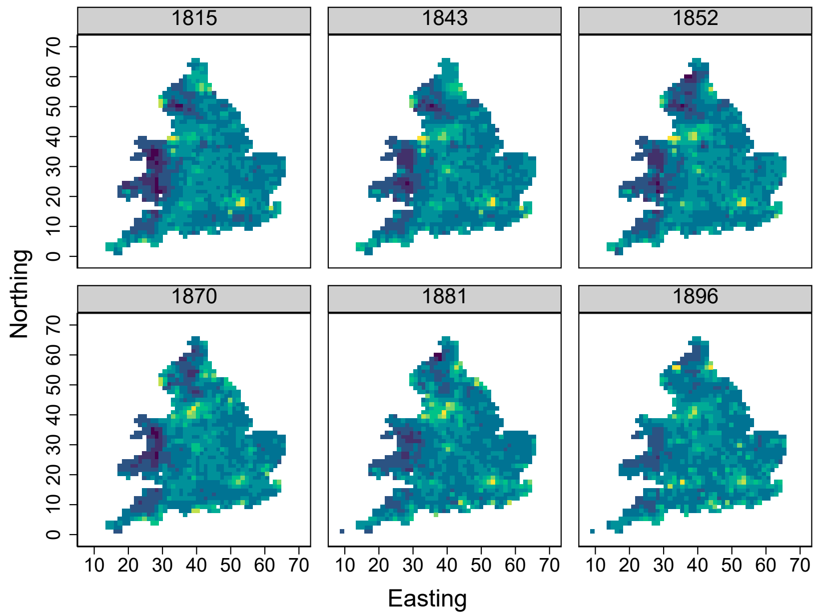
More specifically, we implement a declustered co-kriging interpolation for all years with 12 neighbors and an exponential functional form for γ on the full set of rateable values per acre in the sample. We transform the resulting grid into a raster of rateable value per acre by inverting the initial log transformation. Finally, we overlay this grid with our consistent mappable units to construct rateable values at the level of our main geography. This procedure produces a consistent panel of rateable value per acre across 11,445 parishes.

Figure F.2 shows the spatial variation in rateable value per acre across England and Wales, where blue to yellow represents lower to higher rateable value per acre. Land value appears to be high in larger metropolitan areas, e.g., London, Liverpool, or Manchester, and low in rural areas with limited agricultural use, e.g., the highlands in Wales or inner Cornwall.

F1.4 Poor Law Payments Data

The poor law corresponded to a system of welfare transfers for the poor, dating back to the 1601 Poor Relief Act, as discussed in Boyer (1990) and Renwick (2017). Originally, these welfare transfers were administered at the parish level, and provided money and resources for those in

Figure F.2: Rateable Value Per Acre—Heatplot Over England and Wales



Notes: The figure shows heatmaps of the rateable value per acre for 1851, 1843, 1852, 1865, 1881, and 1896. Blue colors indicate lower rateable values per acre while yellow colors indicate higher values.

need, typically referred to as "outdoor relief." After the 1834 New Poor Law Act, parishes were grouped into Poor Law Unions (registration districts), and there was a move towards housing poor law recipients in designated workhouses, typically referred to as "indoor relief." Nevertheless, substantial amounts of outdoor relief continued to be given, even after 1834. We construct data on the number of paupers relieved for each Poor Law Union for 1841, 1851, 1860, 1870, 1881 and 1891 by digitizing the data reported in the publications of the Houses of Parliament.

The original sources for the poor law data for each year are as follows:

- **1844**: Return of the Average Annual Expenditure of the Parishes Comprised in each of the Unions in England and Wales and the Number of Indoor and Outdoor Paupers relieved. House of Lords Papers; Returns Paper Number 70; Volume 15; 19th Century House of Lords Sessional Papers.
  - Number of indoor plus outdoor paupers relieved per capita in 1841.
- **1851**: Return of all the Indoor and Outdoor Paupers Relieved by the Different Unions in England and Wales. House of Lords Papers; Returns Paper Number 111; Volume 16; 19th Century House of Lords Sessional Papers.

  - Number of indoor plus outdoor paupers relieved per capita in 1851.

- **1860 and 1870**: Return for the Years 1850, 1860, 1870, and 1874 of the Number of Paupers in Receipt of Relief in each Union in England. House of Commons Papers; Paper Number 214; Volume 63; 19th Century House of Commons Sessional Papers.

  - Number of indoor plus outdoor paupers relieved per capita in 1860 and 1870.

- **1883**: Return of Paupers in Receipt of Relief on the 1st January 1883. House of Commons Papers; Paper Number 217; Volume 67; 19th Century House of Commons Sessional Papers.

  - Number of indoor plus outdoor paupers relieved per capita in 1883.

- **1891**: Return of Sum expended for In-maintenance and Out-door Relief in England 1890-91 (and Wales); Return of Number of Paupers in receipt of Relief, January 1891. House of Commons Papers 266; Volume 68; 19th Century House of Commons Sessional Papers.

  - Number of indoor plus outdoor paupers relieved per capita in 1891.

F2 Agricultural Land Use Data

We construct our agriculture land use by combining two data sources: the 1836 Tithe Surveys discussed in Section B3 above and the Agricultural Censuses from 1871-1901.

Tithe Survey Data: The Tithe Survey conducted from 1837–1855 (90 percent was completed by 1845) collected data on agricultural land use by parish for all land that had not been enclosed. There were 14,829 Tithe districts in England and Wales and, for 6,726 of these districts, surveyors' reports are still extant in the National Archives, as digitized by Kain (1986). Although the Tithe districts are not randomly allocated, they are evenly distributed across space, which allows us to compute interpolated land use for all parishes in England and Wales, as shown in Figure B.3 in Section B3 above. We define arable, pastoral and agricultural land use as follows.

- **Arable Land Use:** (i) Corn crops (including wheat, barley, oats, rye, beans and peas); (ii) Green crops (including potatoes, turnips and swedes, mangold, carrots, cabbage, kohlrabi, and oil seed rape, vetches and other green crops except clover or grass); (iii) clover, sanfoin and grasses under rotation; (iv) flax; (v) hops; (vi) bare fallow or uncropped arable land.

- **Pastoral Land Use:** Permanent pasture or grass not broken up in rotation (exclusive of heath or mountain land).

- **Agricultural Land Use:** The sum of arable and pastoral land use.

We thus obtain data on the land area used for arable and pastoral farming for each parish in England and Wales in 1836.

Agricultural Census Data: The agricultural censuses for 1871, 1881, 1891 and 1901 report agricultural land use for each county in England and Wales. We define arable, pastoral and agricultural land use based on the same categories as for the Tithe Survey data above.

- **Arable Land Use:** (i) Corn crops (including wheat, barley, oats, rye, beans and peas); (ii) Green crops (including potatoes, turnips and swedes, mangold, carrots, cabbage, kohlrabi, and oil seed rape, vetches and other green crops except clover or grass); (iii) clover, sanfoin and grasses under rotation; (iv) flax; (v) hops; (vi) bare fallow or uncropped arable land.

- **Pastoral Land Use:** Permanent pasture or grass not broken up in rotation (exclusive of heath or mountain land).

- **Agricultural Land Use:** The sum of arable and pastoral land use above.

For each census decade from 1871-1901, we allocate the county totals for arable and pastoral land area across parishes within counties using the shares of each parish in the county totals for arable and pastoral land area in the 1836 Tithe Surveys. At the county level, our data on arable and pastoral land area are exactly equal to the totals reported in the Agricultural Censuses. At the parish level, we apportion the county totals across parishes using the information from the 1836 Tithe Survey. This apportioning process allows for arbitrary changes in patterns of land use across counties (as captured in the data reported in the Agricultural Censuses), but assumes common trends for a given type of land use across parishes within counties (where these trends can differ both across counties and between arable and pastoral land use).

Agricultural Land use 1841-1901: From combining the Agricultural Censuses from 1871-1901 with the 1836 Tithe Survey, we obtain data on arable and pastoral land area and agricultural land area (arable plus pastoral) for each parish for each census decade from 1871-1901.

For each census decade from 1841-1861, we construct agricultural land area for each parish by linear interpolation across years within parishes, in between the endpoints provided by our data on parish land use in 1836 (from the 1836 Tithe Survey) and parish land use in 1871 (from the combination of the 1871 Agricultural Census and the 1836 Tithe Survey).

We thus obtain our data on the arable share of agricultural land ($\ell_{Git}^{A} = L_{Git} / \left(L_{Git} + L_{Fit}\right)$) and the pastoral share of agricultural land ($\ell_{Fit}^{A} = L_{Fit} / \left(L_{Git} + L_{Fit}\right)$) used in our quantitative analysis of the model.

County Agricultural Land Use 1878-1901: From 1878 onwards, the agricultural censuses report wheat land use as a separate sub-category for each county within the more aggregated category of corn crops, as used in Figure 6 in the paper.

F2.1 Agricultural Prices

We use the agricultural prices data from Clark (2004), which constructs annual price indexes for English net agricultural output from 1209-1914 based on the prices of 26 agricultural goods. The paper reports an overall agricultural price index, separate price indexes for arable, pastoral and wood products, and price series for each of the 26 agricultural goods.

Relative Price of Wheat: We display the wheat price relative to the pastoral price index from 1701-1901 in Figure 1 in the paper.

Arable Price Index: The arable price index is based on the prices of the following eleven crops: Wheat, Rye, Barley, Oats, Peas, Beans, Potatoes, Hops, Straw, Mustard Seed, and Saffron. Products are weighted by estimated output shares, with the cereals of wheat, barley and oats dominating the arable price index.

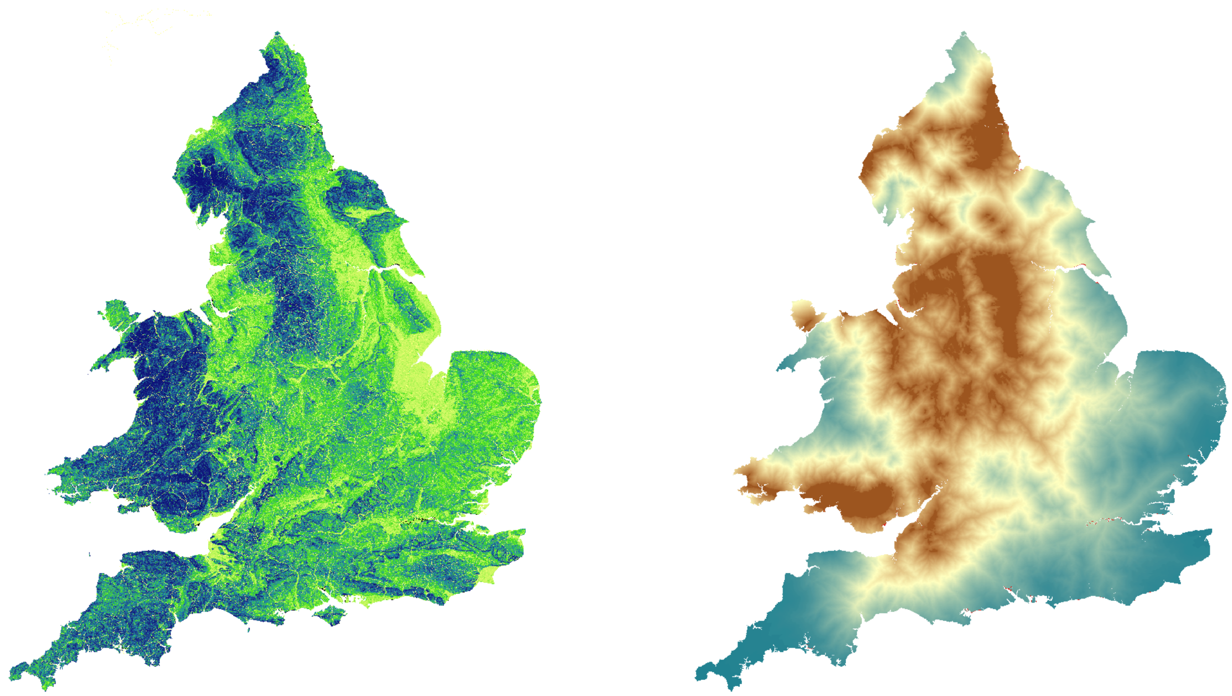
Pastoral Price Index: The pastoral price index is based on the prices of the following eleven products: hay, cheese, butter, milk, beef, mutton, pork, bacon, tallow, wool and eggs. Products are again weighted by estimates of their relative importance.

Relative Arable Price: We use the ratio of the arable price index to the pastoral price index to construct our measures of changes in the relative price of arable products for different time periods for our counterfactuals in Section 6.3 of the paper.

F2.2 Other Geographic Data

Our baseline event-study specification in equation (1) in the paper uses various geographic controls interacted with year fixed effects. Among those controls, two measures rely on transportation costs: (i) a measure of access to resources (travel time to coal); (ii) and a measure of market access (travel time to the nearest market town). We construct these two measures as follows. In a first step, we collect waterways, railways and roads digitized by The Cambridge Group for the History of Population and Social Structure for the project “Transport, urbanization and economic development in England and Wales c.1670–1911”. We consider a given transportation network based on waterways in 1817, roads in 1830 and railway lines in 1846, and we penalize the travel cost by roads versus railways as in Glaeser and Kohlhase (2004). We further assume that the travel cost by waterways is as low as by railways, and we assume that traveling along smaller roads is twice as costly as traveling along the turnpike roads.[7] Panel (a) of Figure F.3 illustrates the spatial distribution of transportation costs across England and Wales.

Figure F.3: Transportation Costs Across England and Wales and Access to Coal



(a) Travel Cost

(b) Distance to Coal

Note: Panel (a) displays the transportation cost as computed from waterways in 1817, roads in 1830 and railway lines in 1846; panel (b) shows the spatial distribution of travel time to the closest coal field through the previous transportation network.

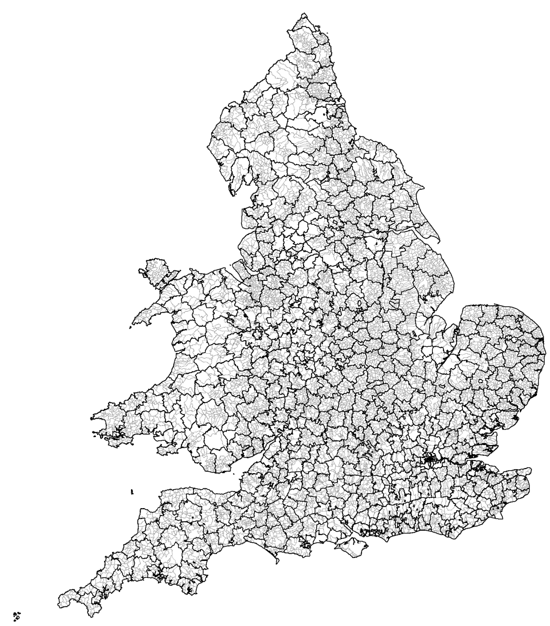
In a second step, we collect data on coal fields in England and Wales—and on the network of market towns—and we construct the travel time to the nearest coal field—and to the nearest town—for any location. Panel (b) of Figure F.3 displays the former—a measure of access to coal resources. There is some correlation between this measure, which could affect the pace of structural transformation, and our exogenous measure of exposure to the Grain Invasion. This is the reason why we control for differential trends along travel time to the nearest coal field in our baseline event-study specification (1).

[7] We also penalize travel using a linear penalty with the local slope: traveling along turnpike roads is twice as costly when the slope is 5 degrees or more.

F3 A Consistent Panel of Parishes

The smallest unit of observation and the lowest tier of local government are civil parishes, which we refer to simply as *parishes*. The boundaries of these parishes can change across the population censuses. To create a consistent panel of mappable spatial units over time, the Cambridge Group has developed a two-stage procedure. First, they spatially match parish level polygons and geographical units from each census to derive all spatial units that existed in any period between 1801–1891. They refer to this as *CGKO* (Cambridge Group Kain Oliver) map. Next, they employ a *Transitive Closure Algorithm* from graph theory (see for example Cormen, Leiserson, Rivest and Stein, 2009) to determine the lowest common unit between parish polygons in different years, which defines the *mappable units*. This procedure implies that these mappable units do not necessarily represent real parishes, but for simplicity we continue to refer to them as *parishes*.

Figure F.4: Consistent Parishes (Gray) and Registration Districts (Black)



Notes: This Figure displays the output of the transitive closure algorithm implemented by the Cambridge Group for History of Population and Social Structure. Consistent mappable units based on parishes are displayed in gray; registration districts are displayed with black borders.

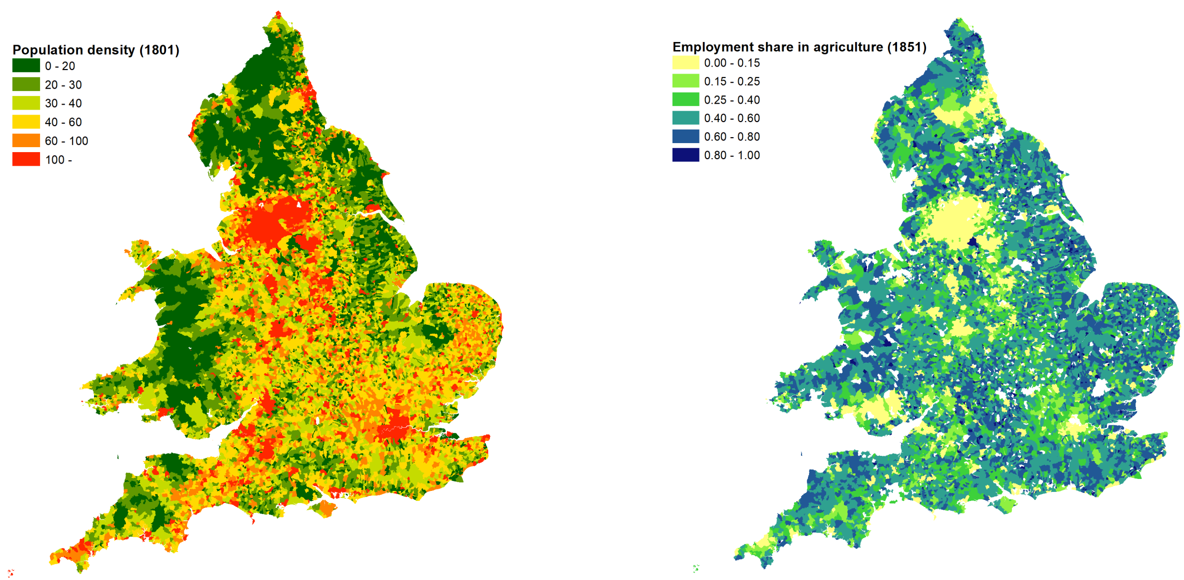
F4 Rural/Urban Classification of Parishes

The empirical analysis in Section 4 of the paper sometimes uses a classification of our baseline administrative units into rural and urban parishes. In this section of the Online Appendix, we describe the construction of this classification and its robustness to alternative approaches.

F4.1 The Structural Transformation of England and Wales

In the first half of the nineteenth century, England and Wales was already on the path to structural transformation. In 1801, about 57% of the population already lived in parishes with a population density of over 100 inhabitants per square kilometer, and about 63% lived within 5 kilometers of one of the 70 largest towns or within 1 kilometer of one of the 500 market towns. In 1851, the aggregate employment share in agriculture was about 22%, compared to 34% in manufacturing, 27% in services, and the remaining 17% in mining, construction and transportation. Figure F.5 illustrates the structural transformation of England and Wales, with a large concentration of non-agricultural workers in London and in Lancashire. Nevertheless, many parishes in rural areas continued to have agricultural employment shares of 60 percent or more.

Figure F.5: Spatial Distribution of Population Density and Agricultural Employment



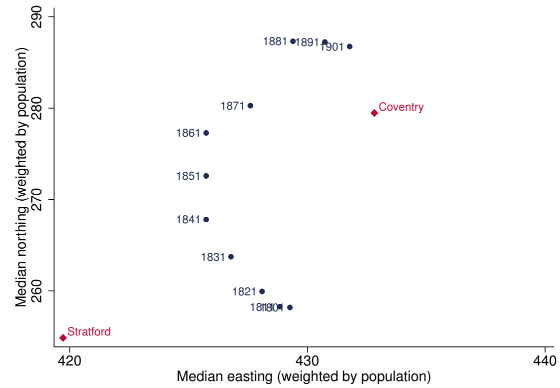
(a) Population density (1801) (b) Agricultural employment (1851)

Notes: This figure displays the distribution of population density (1801) and agricultural employment (1851) across consistent parishes. Agriculture includes both farming and animal husbandry, but it excludes estate work, forestry, and fishing.

The structural transformation of the economy accelerates during the second half of the century—a period which coincides with the Grain Invasion; the employment share in agriculture decreases to 13% in 1881 and 8% in 1911. The movement of agricultural workers away from parishes with high wheat suitability however tends to be quite local, as documented in the next section, such that the epicenter of England and Wales does not move greatly (see Figure F.6).

Figure F.6: Spatial distribution of population density and agricultural employment.

(a) Average location of workers (1801–1911).



(b) Median location of workers (1801–1911).

Notes: This Figure displays the average and median Easting and Northing, as weighted by parish populations from 1801 to 1901. We also report the locations of Coventry, and Stratford or Rugby (separated by about 20 and 10 kilometers, respectively).

F4.2 K-Means Clustering

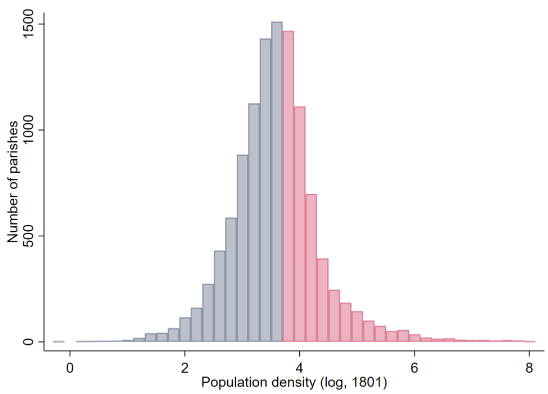
One issue with using consistent parishes as our baseline units is that there is no official classification of such units as rural or urban. In order to design a classification that is exogenous to the subsequent evolution of parishes, we proceed as follows. We first create a measure of population density at baseline, in 1801. We then apply a k-means clustering partitioning of parishes into 2 groups along this baseline characteristic. We label the group with high population density as being "urban" and the remaining parishes as being "rural."

Figure F.7 displays the distribution of population density and agricultural employment across the two groups at baseline (panels a and b) and later on, in 1881 (panels c and d). As apparent in the figure, the k-means clustering generates a very clear partitioning of the data along population density, and a noisier partitioning along initial agricultural employment. The partitioning in 1881 becomes slightly blurrier along population density.

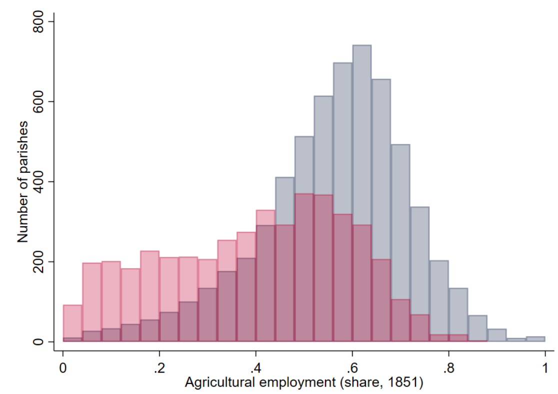
F4.3 Robustness to Alternative Rural/Urban Classifications

The previous approach has the advantage of defining urban locations based on population density, as in seminal models of the monocentric city. We now combine this characterization of urban locations with an industry-based characterization using industrial employment shares. In the top panels of Figure F.8, we report the partitioning of parishes through k-means clustering based on two variables: (log) population density (1801) and agricultural employment share (1851). In the bottom panels of Figure F.8, we report the partitioning of parishes based on only agricultural

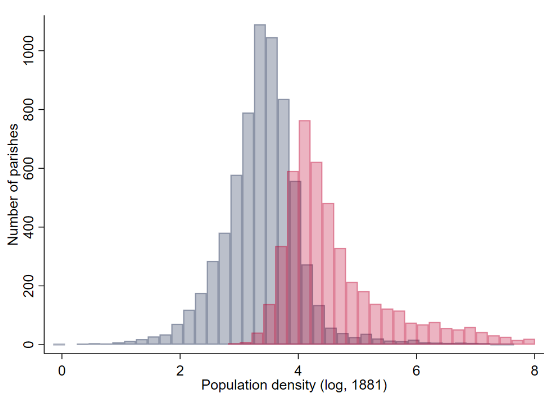
Figure F.7: Rural/Urban Classification of Parishes



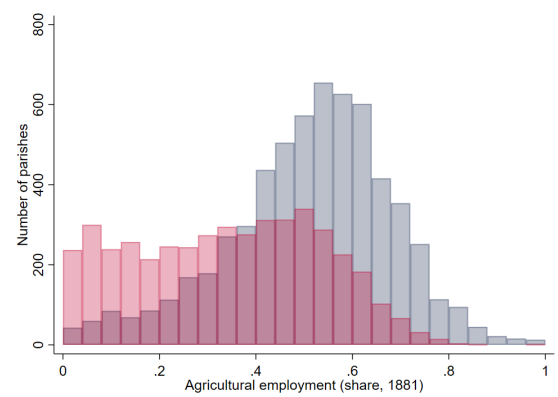
(a) Population Density (1801)



(b) Agricultural Employment (1851)



(c) Population Density (1881)

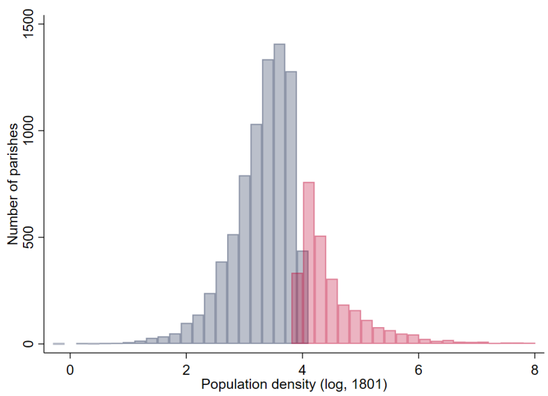


(d) Agricultural Employment (1881)

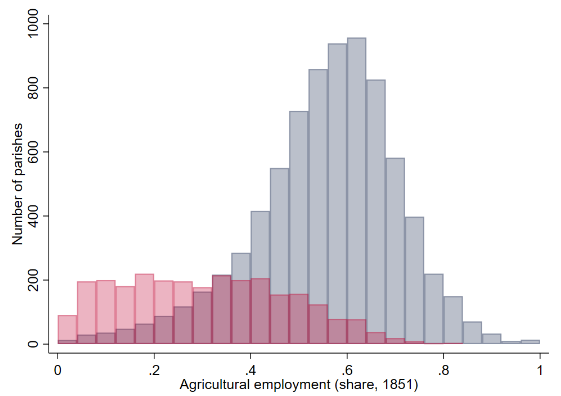
Notes: This Figure displays the distribution of population density and agricultural employment at the parish level, in 1801/1851 and in 1881, for rural (in blue) and urban parishes (in red). Parishes are classified into urban/rural categories according to a k-means clustering partitioning the 11,500 parishes into 2 groups along (log) population density (1801).

employment share (1851).

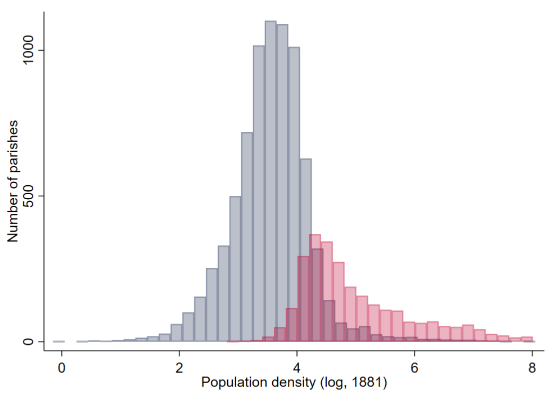
Figure F.8: Rural/Urban Classification of Parishes—Sensitivity Analysis



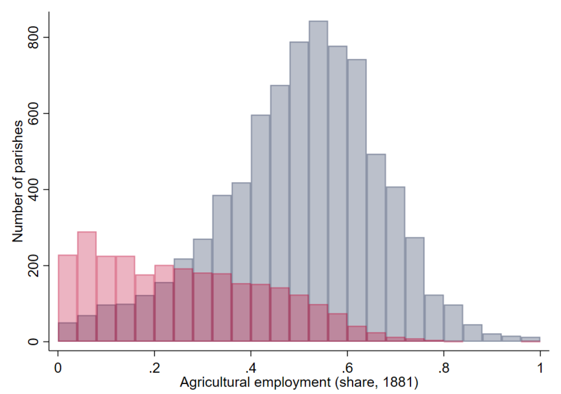
(a) Population Density (1801), Clustering Based on Population and Industry



(b) Agricultural Employment (1851), Clustering Based on Population and Industry



(c) Population Density (1801), Clustering Based on Industry



(d) Agricultural Employment (1851), Clustering Based on Industry

Notes: This Figure displays the distribution of population density and agricultural employment at the parish level, in 1801/1851, for rural (in blue) and urban parishes (in red). Parishes are classified into urban/rural categories according to k-means clustering partitioning the 11,500 parishes into 2 groups along: population density (1801) and agricultural employment (1851)—panels a and b; agricultural employment (1851)—panels c and d.

G Matching Census Records

Analyzing the reallocation of labor across space, occupations and industries involves tracking individuals over time. This section of the Online Appendix describes the matching procedure that we use to link the records of the Integrated Census Microdata Project (I-CeM) across census waves, discusses the quality of matches and the selection of individuals into matching, and provides descriptive statistics about spatial mobility and occupational mobility.

G1 Description of the Matching Procedure

The matching procedure combines the variables made available by the Integrated Census Microdata Project (I-CeM), most notably the age of each respondent, their gender, their county of birth, and the household structure (as in Price et al. 2021), and the (usually unavailable) names of individuals.[8]

The matching procedure proceeds in three steps, quite similar in nature to the ABE-JW algorithm described in Abramitzky et al. (2021). In a first step, we perform some basic, systematic cleaning including: the substitution of obvious abbreviations in names (e.g., "GEO" for GEORGES); the standardization of names as one first name and one surname (without middle names—as they are infrequently documented, and reported differently across census years), the harmonization of county codes.

In a second step, we select possible matching pairs using a mix of block matching (on the phonetic encoding of the first name, the phonetic encoding of the surname, the county of birth, the gender) and fuzzy matching (on the first name, the surname and age of the household member). The phonetic encoding of names relies on the metaphone algorithm developed for English spelling and pronunciation, and trained on the first names and family names commonly found in the United States. "Redding" and "Reading" would both be transformed into "RTNK", and thus verify the phonetic-based matching condition. The matching condition on names is based on the Jaro-Winkler distance, a measure of distance between sequences of characters. The Jaro-Winkler distance is not very intuitive; we use it only to exclude obvious mismatches. In order to properly characterize the distance between names, we compute instead the Levenshtein distance—much slower to process—on the selected set of possible pairs and further restrict the set of matched pairs to those with at most 2 substitutions needed to match names and surnames across waves. The pair ("Robt Little", "Robert Little") would be kept but not ("Robt Little", "Robert Litte"). The matching condition on age imposes that an individual is no more than 1 year younger or older

[8] More specifically, we extract the following variables from the Integrated Census Microdata Project (I-CeM): a unique individual identifier, a unique household identifier, the location of the record in the census (piece, folio, page), a unique parish identifier, information on the county, registration district and sub-district, the gender, age, occupation, birth information, marital status of each household member, the address of the household.

than what would have been implied by the approximate time difference between two census waves (e.g., pairs of entries are kept if individuals in 1861 are 9, 10 or 11 years older than its possible counterpart in 1851).

In a third step, we select the most suitable matches among a large set of matched pairs.[9] Consider a match (i_1, j_1), and an associated score $f(i_1, j_1)$ based on the distance between the respective characteristics of i_1 (in census t) and j_1 (in census $t + 1$). Tie-breaking is difficult because (i) individual i_1 also may be matched with a list of individuals $\{j_2, \ldots, j_J\}$ (in census $t+1$) and (ii) individual j_1 also may be matched with a list of individuals $\{i_2, \ldots, i_I\}$ (in census t). We adopt a conservative approach and only select a matched pair as viable if there does not exist *any* other pair involving one of the nodes and with a score at least as good. This approach implies that ties are dropped, and thus differs from probabilistic matching—usually involving machine learning—allowing different matches of the same individual to be kept with a probabilistic weight.

The score function is constructed on a priority principle. We partition the score function by (a) other common household matches, (b) matching proximity in names (i.e., needed substitutions between sequences of characters), (c) matching proximity in age, in that order of priority. In other words, a matched pair that has, at least, one other match involving another household member would be given a higher score than any competing matches without "common household matches"—irrespective of the other characteristics of the match or the quality of the match for the household member. For instance, a "Thomas Hardy" in 1881 matched to a "Thomas Hardy" in 1891, and identified within the same household together with another possible match ("Emma Gifford") in both periods would be given a higher score than any other single matches involving "Thomas Hardy" in 1881 or in 1891. Within each partition, we give priority to the best match along the specific dimension considered: In the counterfactual situation in which common matches would exist across different households, we give priority to the highest number of such matches.

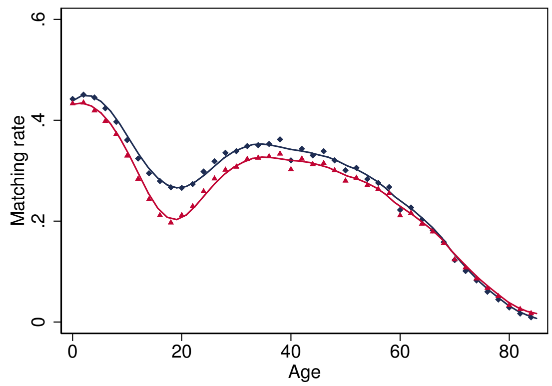
G2 Matching Quality and Selection

In this section, we describe the outcome of the matching procedure. More specifically, we discuss: the matching rate; indicators of matching quality; and a description of selection into matching.

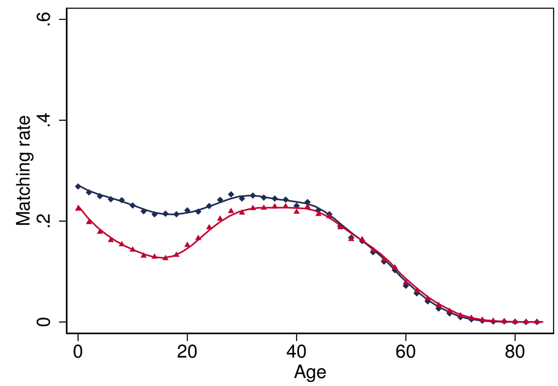
Table G.1 reports the matching rate in 1851–1861, 1861–1881, 1881–1891, and 1891–1901, and its variation across gender and age is displayed in Figure G.1. Our conservative procedure leads to a matching rate of about 32-34% in 1851–1861 and 1881–1891, 22% in 1861–1881 (with a gap of two decades) and 45% in 1891–1901. The matching rate between consecutive waves shows strong empirical regularities: matching decreases by 10-15% for teenagers/young adults due to

[9]The previous procedure generates, for instance, about 80,000,000 matched pairs between the 1891 and 1901 Census records (each consisting of about 25-30 million unique entries). A non-negligible number of entries in one wave are matched with more than 10 possible records in the other wave—an issue that would not arise for a "Yanos Zylberberg" for instance.

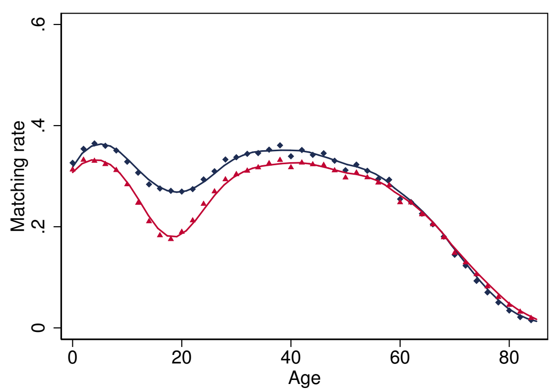
Figure G.1: Matching Rates Across Consecutive Census Waves (1851–1901).



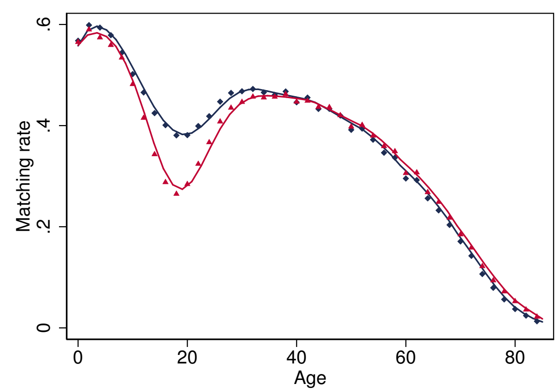
(a) 1851–1861



(b) 1861–1881



(c) 1881–1891



(d) 1891–1901

Notes: This Figure displays the matching rate between two consecutive Census waves as a function of age (in the first Census wave) and by gender (red: female, blue: male). The matching rate consistently decreases for teenagers/young adults (especially women, given the change in names associated with marriage) and for older individuals (given the high mortality rate after 50). The individual Census records in 1871 are not included by the Integrated Census Microdata Project.

migration and marriage (especially women, given the subsequent change in surname); matching decreases steadily for older individuals, and similarly so across gender. We show in Table G.1 that there remains a large unexplained variation in matching probability, even after controlling for age, marital status, occupation/industry and parish of residence: matching depends crucially on the number of individuals sharing the same name and county of birth—age being quite uniformly distributed. The matching probability thus mostly relates to the triple (and not so rare) coincidence of having a common first name (e.g., Sarah), a common name (e.g., Smith), and a common county of birth (e.g., Middlesex).

Our conservative matching procedure implies that the average match quality is likely to be

Table G.1: Matching Probability in 1851–61, 1861–81, 1881–91 and 1891–01

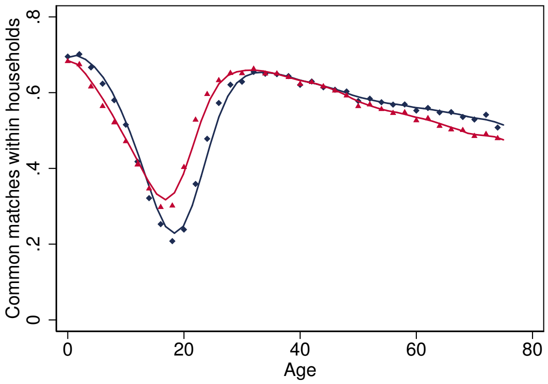
Census wave	1851–61	1861–81	1881–91	1891–01	*all*
Matching rate (individual)	.335	.215	.314	.457	*.341*
Matching rate (household)	.623	.487	.602	.727	*.620*
R-square (age)	.029	.022	.012	.041	
R-square (+ marital)	.036	.026	.019	.052	
R-square (+ occupation)	.038	.033	.029	.056	
R-square (+ industry)	.037	.032	.029	.055	
R-square (+ parish)	.059	.052	.049	.072	
Observations (males)	8,447,553	9,299,034	12,504,744	13,773,762	*44,025,093*

This table reports statistics about matching: (i) the matching rate at the individual level, (ii) the matching rate at the household level, i.e., the probability that a male within a household in t has at least one household member matched in $t+1$, (iii) the R-square of regressions of a matching dummy on age fixed-effects, marital fixed-effects, marital fixed-effects, industry fixed-effects, and parish fixed-effects.

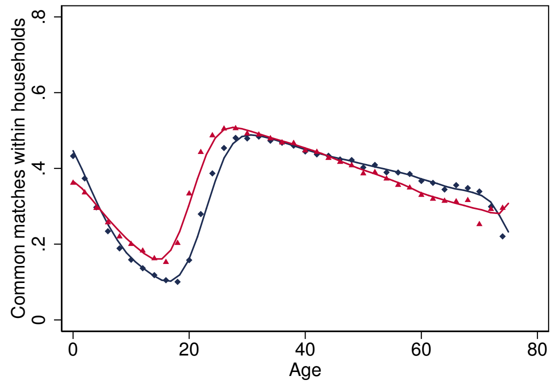
high. While we do not implement any external validation exercise, we now provide reassuring descriptive statistics. First, the match is (i) perfect along all core matching variables (names, age, county of birth, gender) for more than 85% of matches, (ii) similarly so for all consecutive pairs of Censuses, and (iii) consistently so across gender and age. This would be consistent with a reporting error that is orthogonal to age and gender. Second, the vast majority of imperfect matches for names involve only one substitution between two letters. Third, the proportion of co-household members that are matched is high—about 60% on average—and there is a dip reflecting expected changes of households for young adults (a few years earlier for women). We report these statistics in Figure G.2.

There is differential selection into matching, as shown by the importance of age and gender in predicting successful matches (see Figure G.1). This selection is likely to be correlated with socio-economic characteristics of household members, e.g., their education (through spelling, uncommon names etc.). We now provide additional insight about how such selection could correlate with important dimensions of the empirical analysis conducted in Section 4, notably the occupational structure and the exposure to the trade shock. In Panel A of Table G.2, we report the correlation between matching and individual characteristics at baseline, whether the individual is engaged in agriculture or proxies for socio-economic vulnerability, in separate regressions for the different (initial) census waves (i.e., 1851, 1861, 1881, 1891). As apparent, there is an occupational/urban gradient in the probability to be matched. Farmers are much more likely to be matched, possibly because the county of birth is used as a block variable and is less specific to a random urban dweller than to a random rural dweller. By contrast, paupers, beggars or servants are less likely to be matched. In Panel B of Table G.2, we report the correlation between

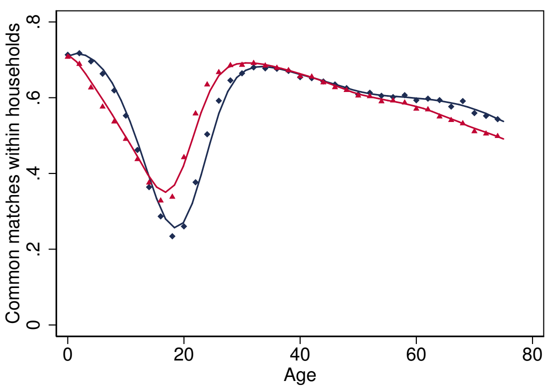
Figure G.2: Common Matches Within Households (1851–1901)



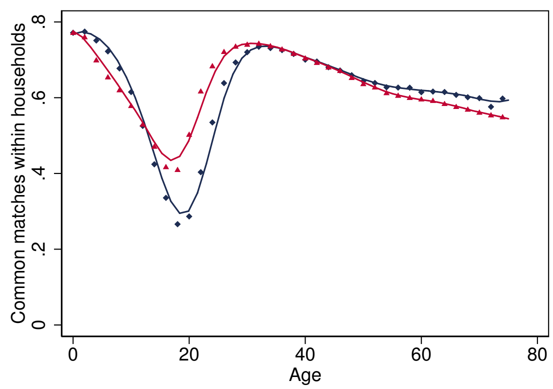
(a) 1851–1861



(b) 1861–1881



(c) 1881–1891



(d) 1891–1901

Notes: This Figure displays the share of (other) household members of a Census wave matched in the next Census wave as a function of age (in the first Census wave) and by gender (red: female, blue: male). The matching rate markedly decreases for teenagers/young adults migrating out of their parent household into a new household. The individual Census records in 1871 are not included by the Integrated Census Microdata Project.

matching and geographic characteristics of the origin parish. There are little signs of differential selection into matching along our trade exposure variable, or evidence of differential selection along geographic characteristics and over time.

G3 Spatial Mobility and Occupational Mobility

The matching of individuals across consecutive Census waves allows us to document the relocation of labor across space and across occupations/industries. To this purpose, we pool all consecutive observations between 1851–1861, 1861–1881, 1881–1891, 1891–1901, and provide average

Table G.2: Selection into Matching in 1851–61, 1861–81, 1881–91 and 1891–01

Census wave	1851	1861	1881	1891
Panel A: selection along agricultural occupation and social status				
Farmer	.0921	.0690	.1110	.0896
	(.0010)	(.0009)	(.0010)	(.0011)
Farm helper	.1080	.0810	.1040	.0882
	(.0015)	(.0013)	(.0013)	(.0015)
Agricultural laborer	.0188	.0399	.0731	.0154
	(.0005)	(.0005)	(.0006)	(.0006)
Farmer, other	.0191	.0361	.0547	.0292
	(.0015)	(.0012)	(.0006)	(.0012)
Pauper	-.0428	-.0168	.0053	-.0426
	(.0017)	(.0025)	(.0036)	(.0035)
Beggar	-.0721	-.0426	-.0731	-.0086
	(.0041)	(.0033)	(.0024)	(.0022)
Servant	-.0584	-.0280	-.0550	-.0605
	(.0011)	(.0004)	(.0010)	(.0010)
Middle class	-.0002	-.0152	-.0439	.0050
	(.0004)	(.0004)	(.0004)	(.0004)
Upper class	-.0427	-.0564	-.1080	-.0722
	(.0009)	(.0007)	(.0007)	(.0008)
Observations	8,447,553	9,299,033	12,494,825	13,773,758
Panel B: selection along geographic characteristics				
Wheat suitability	.0087	.0065	-.0009	.0047
	(.0065)	(.0051)	(.0062)	(.0057)
Grass suitability	-.0369	-.0265	-.0307	-.0363
	(.0048)	(.0034)	(.0043)	(.0045)
Distance to coal (100 kms)	-.0001	-.0130	-.0181	-.0048
	(.0036)	(.0029)	(.0045)	(.0052)
Distance to London (100 kms)	-.0095	-.0018	.0011	-.0019
	(.0036)	(.0025)	(.0034)	(.0039)
Distance to nearest town (100 kms)	.0743	.0805	.0759	.0628
	(.0114)	(.0079)	(.0118)	(.0136)
Observations	6,541,950	7,340,195	10,398,756	11,073,523

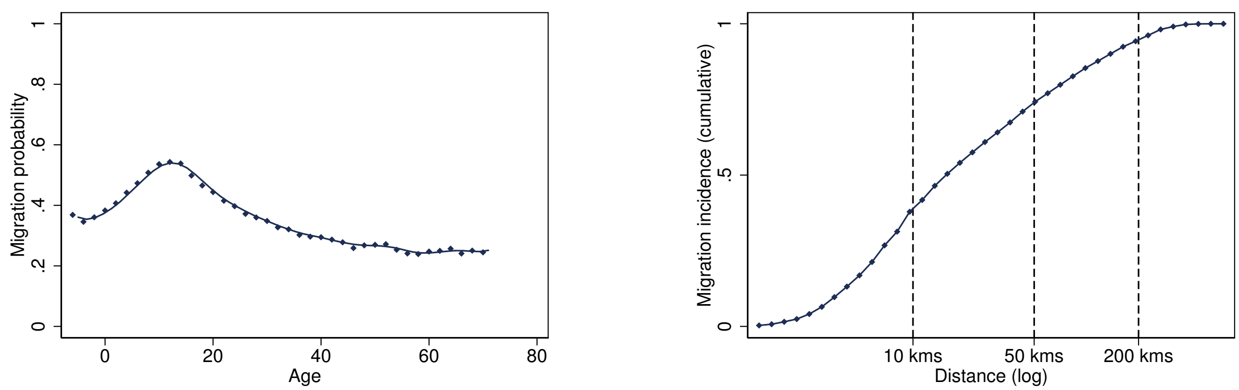
Standard errors are reported between parentheses and are clustered at the parish-level. The unit of observation is an individual at a given time t. In Panel A, all specifications condition the analysis on age fixed effects and marital status. In Panel B, all specifications condition the analysis on age fixed effects, marital status, occupation (2-digit) and industry (1-digit) fixed effects. See Equation (2) for a more formal description of the empirical strategy.

statistics in the pooled sample of males.[10]

[10]The wave-specific statistics are similar one to another; there are limited trends in the relationship between migration or job changes and baseline characteristics.

Panel (a) of Figure G.3 shows that migration spells are very likely on average (around 40% of males report a different registration district among the pooled sample of matched individuals), with a significantly higher migration incidence for young adults looking for a suitable job or a suitable partner. These migration spells tend to be local, with more than half of these spells occurring between parishes that are less than 20 kms apart (see Panel (b) of Figure G.3).

Figure G.3: Migration as a Function of Age and Cumulative Distribution of Distance Between Sending and Receiving Parishes



(a) Migration as a Function of Age

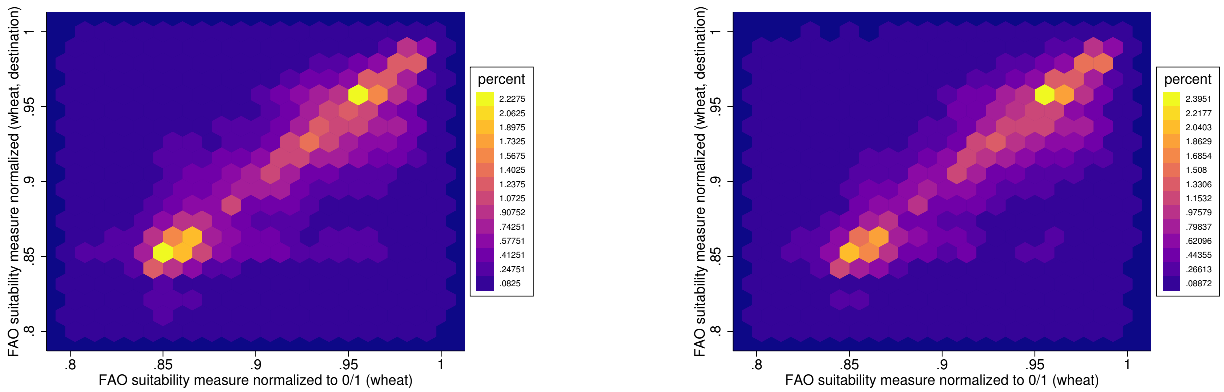
(b) Cumulative Distribution of Distance

Notes: Panel (a) displays the probability to report a different registration district among the pooled sample of matched individuals between 1851–1861, 1861–1881, 1881–1891, 1891–1901, as a function of the average age between the two consecutive Census waves. Panel (b) displays the cumulative distribution of distance between sending and receiving parishes.

Even if a large number of migration spells are across urban centers or from rural hinterlands to the nearby urban area, there is a general relocation of workers from areas with higher wheat suitability to areas with lower wheat suitability following the Grain Invasion in the second half of the 19th century. Figure G.4 provides further evidence on this relocation of labor from high wheat suitability areas towards other locations. Panel (a) displays the joint distribution of wheat suitability at destination and at origin for all migration spells between 1851–61, 1861–81, 1881–91, 1891–1901; panel (b) limits the sample to farmers or agricultural workers at the time of migration.

Figure G.5 shows that the probability to report a different 2-digit occupation (Panel a), or a different 1-digit industry (Panel b), is very high among the pooled sample of matched individuals. About 50% of young active adults change their industry between two Census waves, and this job mobility remains high and stable even for older workers (at about 40%).

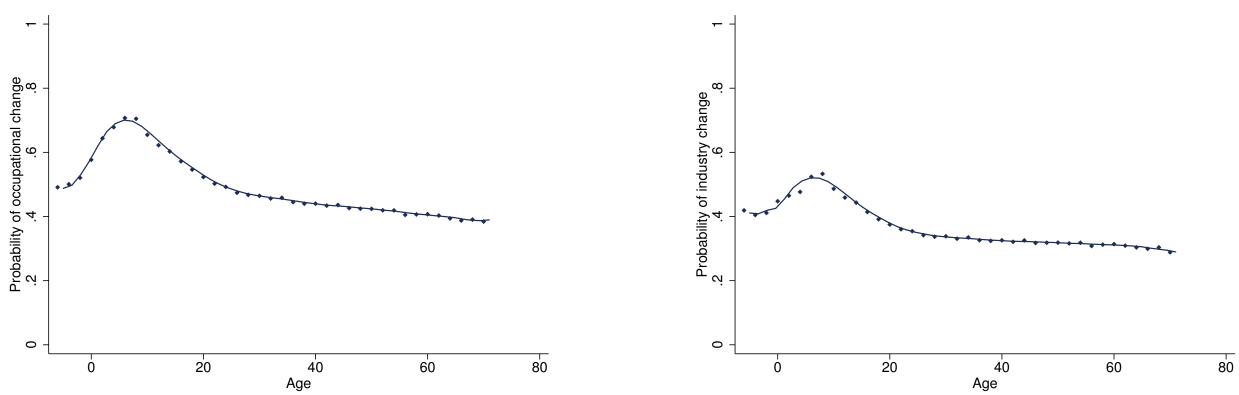
Figure G.4: Wheat Suitability at Destination and at Origin



(a) All Migration Spells (b) Migration Spells for Farmers

Notes: This Figure represents the joint distribution of wheat suitability at destination and at origin for all migration spells between 1851–61, 1861–81, 1881–91, 1891–1901, and for migration spells of farmers.

Figure G.5: Change in Occupation or Industry as a Function of Age



(a) Changes in Occupation (b) Changes in Industry

Notes: Panel (a) (resp. b) displays the probability to report a different 2-digit occupation (resp. 1-digit industry) among the pooled sample of matched individuals between 1851–1861, 1861–1881, 1881–1891, 1891–1901, as a function of the average age between the two consecutive Census waves.

References

Abramitzky, R., K. Eriksson, J. Feigenbaum, L. Platt Boustan and S. Perez (2021) “Automated Linking of Historical Data,” *Journal of Economic Literature*, 59(3), 865-918.

Adão, R., E. Morales, and M. Kolesár (2019) “Shift-Share Designs: Theory and Inference,” *Quarterly Journal of Economics*, 134(4), 1949-2010.

Allen, R. G. D and A. L. Bowley (1935) *Family Expenditure: A Study of its Variation*, London: P. S. King and Son, Ltd.

Artuç, Erhan and Chaudhuri, Shubham and McLaren, John (2010) “Trade Shocks and Labor Adjustment: A Structural Empirical Approach,” *American Economic Review*, 100, 1008–1045.

Bah, El-hadj M (2007) “A Three-Sector Model of Structural Transformation and Economic Development,” *Working Paper no. 10654*, Munich Personal RePEc Archive.

Bank of England (2024) *A Millennium of Macroeconomic Data*, London: Bank of England, https://www.bankofengland.co.uk/statistics/research-datasets.

Baumol, William J. (1967) “Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis,” *American Economic Review*, 57(3), 415-426.

Bellerby, J. R. (1953) “Distribution of Farm Income in the United Kingdom 1867–1938,” *Journal of Proceedings of the Agricultural Economics Society*, 10(2), 127-138.

Board of Agriculture and Fisheries (1906) *Report on the Decline in the Agricultural Population of Great Britain, 1881–1906*, London: Darling and Son.

Bohr, Clement, Martí Mestieri and Frederic Robert-Nicoud (2023) “Heterothetic Cobb-Douglas: Theory and Applications,” UPF mimeo.

Borusyak, Kirill, Xavier Jaravel and Jann Spiess (2024) “Revisiting Event-Study Designs: Robust and Efficient Estimation,” *Review of Economic Studies*, forthcoming.

Boyer, G. R. (1990) *An Economic History of the English Poor Law 1750-1850*, Volume 1, Cambridge: Cambridge University Press.

Bryan, Gharad and Melanie Morten (2019) “The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia,” *Journal of Political Economy*, 127(5), 1993-2568.

Caird, J. (1852) *English Agriculture in 1850-51*, London: Longman, Brown, Green and Longmans.

Caliendo, Lorenzo, Maximiliano Dvorkin and Fernando Parro (2019) “Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock,” *Econometrica*, 87(3), 741-835.

de Chaisemartin, Clément, Xavier D’Haultfœuille (2020) “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects,” *American Economic Review*, 110(9), 2964-96.

Callaway, Brantly and Pedro H.C. Sant'Anna (2021) "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 225(2), 200-230.

Clark, Gregory (2004) "The Price History of English Agriculture, 1209-1914," *Research in Economic History*, 22, 41-123.

Cobbett, W. (1830) *Rural Rides*, London: T. Nelson and Sons.

Comin, Diego, Danial Lashkari and Martí Mestieri (2021) "Structural Change With Long-Run Income and Price Effects," *Econometrica*, 89(1), 311-374.

Costinot, A., D. Donaldson, and C. Smith (2016) "Evolving Comparative Advantage and the Impact of Climate Change in Agricultural Markets: Evidence from 1.7 Million Fields around the World," *Journal of Political Economy*, 124(1), 205-248.

Cormen, Thomas H., Charles E. Leiserson, Ronald L. Rivest and Clifford Stein (2009) *Introduction to Algorithms*, Third Edition, Cambridge MA: MIT Press.

Davis, M. A., W. D. Larson, S. D. Oliner, and J. Shui (2021) "The Price of Residential Land for Counties, ZIP Codes, and Census Tracts in the United States," *Journal of Monetary Economics*, 118, 413-431.

Deane, Phyllis and W. A. Cole (1967) British Economic Growth 1688-1959: Trends and Structure, Cambridge: Cambridge University Press.

Duarte, Margarida and Diego Restuccia (2010) "The Role of the Structural Transformation in Aggregate Productivity," *Quarterly Journal of Economics*, 125(1), 129-73.]

Lord Ernle, Rowland E. Prothero (1912) *English Farming, Past and Present*, London: Longmans, Green and Company.

Fajgelbaum, Pablo D., Eduardo Morales, Juan Carlos Suárez-Serrato and Owen Zidar (2019) "State Taxes and Spatial Misallocation," *Review of Economic Studies*, 86, 333-376.

Fajgelbaum, Pablo and Stephen Redding (2022) "Trade, Structural Transformation and Development: Evidence from Argentina 1869-1914," *Journal of Political Economy*, 130(5), 1249-1318.

Fagan, B. (2019) *Little Ice Age*, London: Basic Books

Feinstein, C. H. (1972) *National Income, Expenditure and Output of the United Kingdom, 1873-1896*, Cambridge: Cambridge University Press.

Galle, S., A. Rodríguez-Clare and M. Yi (2022) "Slicing the Pie: Quantifying the Aggregate and Distributional Consequences of Trade," *Review of Economic Studies*, 90(1), 331-375.

Glaeser, Edward L. and Janet E. Kohlhase (2004) "Cities, Regions and the Decline of Transport Costs," *Papers in Regional Science*, 83(1), 197-228.

Hanoch, Giora (1975) "Production and Demand Models with Direct or Indirect Implicit Additivity," *Econometrica*, 43(3), 395-419.

Imlah, Albert Henry (1958) *Economic Elements in the Pax Britannica: Studies in British Foreign Trade in the Nineteenth Century*, Cambridge MA: Harvard University Press.

Kain, R. J. (1986) *An Atlas and Index of the Tithe Files of Mid-nineteenth Century England and Wales*, Cambridge: Cambridge University Press .

Larson, W. and J. Shui (2022) "Land Valuation using Public Records and Kriging: Implications for Land versus Property Taxation in Cities," *Journal of Housing Economics*, 58, 101871.

Li, J. and A. D. Heap (2011) "A Review of Spatial Interpolation Methods for Environmental Scientists," *Ecological Informatics*, 3-4, 228-241.

Matsuyama, Kiminori (2019) "Engel's Law in the Global Economy: Demand-Induced Patterns of Structural Change, Innovation, and Trade," *Econometrica*, 87(2), 497-528.

Mitchell, B. R. (1988) *British Historical Statistics*, Cambridge: Cambridge University Press.

O'Rourke, K. H. (1997) "The European Grain Invasion, 1870-1913," *Journal of Economic History*, 57(4), 775-801.

Perry, P. J. (1972) "Where Was the Great Agricultural Depression? A Geography of Agricultural Bankruptcy in Late Victorian England and Wales," *Agricultural History Review*, 20(1), 30-45.

Perry, P. J. (1973) *British agriculture 1875-1914*, London: Methuen and Co Ltd.

Price, J., K. Buckles, J. Van Leeuwen and I. Riley (2021) "Combining Family History and Machine Learning to Link Historical Records," *Explorations in Economic History*, 80, 101391.

Renwick, C. (2017) *Bread of All: The Origins of the Welfare System*, London: Penguin Random House.

Rogerson, Richard (2008) "Structural Transformation and the Deterioration of European Labor Market Outcomes," *Journal of Political Economy*, 116(2), 235-59.

Saiz, Albert (2010) "Geographic Determinants of Housing Supply," *Quarterly Journal of Economics*, 125(3), 1253-1296.

Sato, R (1975) "The Most General Class of CES Functions," *Econometrica*, 43, 999-1003.

Sherman, M. (2011) *Spatial Statistics and Spatio-temporal Data: Covariance Functions and Directional Properties*, London: John Wiley and Sons.

Sotelo, Sebastian (2020) "Domestic Trade Frictions and Agriculture," *Journal of Political Economy*, 128(7), 2690-2738.

Sun, Liyang and Sarah Abraham (2021) "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 225(2), 175-199.

Üngör, Murat (2017) "Productivity Growth and Labor Reallocation: Latin America versus East Asia," *Review of Economic Dynamics*, 24, 25-42.

Wrigley, Edward Anthony (2011) *The Early English Censuses*, Oxford: Oxford University Press.