Online Appendix of:

Rage Against the Machines: Labor-Saving Technology and Unrest in Industrializing England

Bruno Caprettini and Hans-Joachim Voth

February 18, 2020

A Data appendix

In this appendix, we describe all the steps that are necessary to create the dataset used in our analysis.

A.1 Dataset construction

To construct our database, we start from the map of ancient parishes of England and Wales prepared by Southall and Burton (2004). This map derives from earlier electronic maps by Kain and Oliver (2001), and contains a GIS database of all parishes of England and Wales in 1851. It consists of 22,729 separate polygons, each identifying a separate location. These are smaller than a parish, so that a given parish is often composed of several polygons. Because we observe all our variables at the parish level, we start by aggregating the 22,729 polygons into 11,285 parishes.[1]

Next, we aggregate a subset of these parishes into larger units of observation. We do this in two cases. First, large urban areas such as London, Liverpool or Manchester consists of several distinct parishes. Treating these areas as separate observations is incorrect, because we always observe riots and threshing machines for a whole city, and we generally not able to assign them to any specific area within a city. Thus, all parishes belonging to a city form a single observation. We also aggregate different parishes into larger units when the information from at least one of our sources does not allow us to compute one of our variables more precisely. This happens when one of our sources records a riot, a threshing machine, or Census population for a large area comprising several parishes. In these cases, we also aggregate all variables at the level of the larger unit of observation. Table 1 reports the full list of towns constructed aggregating more than one parish.[2]

At the end of this process, we are left with 10,700 separate observations. Of these, we are able to match 9,737 to the 1801 Population Census based on county and parish name. We drop 59 observations that report 0 workers and 1 that reports 0 men.[3] Finally, the area of two parishes was so small that we could not evaluate the suitability of the soil from the geographical raster: we drop these parishes as well. The final sample has a maximum of 9,674 observations.

[1]We do this based on the fields GAZ_CNTY and PAR, which identify county and parish.

[2]There is a second reason for aggregating parishes within cities. Because most of riots and almost all machines appear in rural areas, keeping separate observations for each urban parish effectively duplicates observations with no riots and no machines. This would introduce the "Moulton problem" (Moulton, 1990) and, by biasing standard errors downwards, it would artificially increase the precision of our estimates.

[3]These 0s create missings in the share of agricultural workers and in the log sex ratio.

Table 1: List of cities and towns created by aggregating more than one parish.

County	City or village	Parishes aggregated
London	London	80
Yorkshire, West Riding	York	55
Norfolk	Norwich	36
Devon	Exeter	25
Kent	Canterbury	24
Lincolnshire	Lincoln	21
Gloucestershire	Bristol	20
Oxfordshire	Oxford	13
Cheshire	Chester	13
Suffolk	Ipswich	13
Hampshire	Winchester	12
Gloucestershire	Gloucester	12
Essex	Colchester	12
Cambridgeshire	Cambridge	12
Leicestershire	Leicester	11
Worcestershire	Worcester	11
Sussex	Chichester	11
Sussex	Hastings	7
Shropshire	Shrewsbury	7
Hampshire	Southampton	7
Sussex	Lewes	6
Herefordshire	Hereford	6
Lincolnshire	Stamford	5
Surrey	Guildford	5
Bedfordshire	Bedford	5
Northamptonshire	Northampton	5
Berkshire	Wallingford	5
Yorkshire, East Riding	Beverley	4
Brecknockshire	Brecon	4
Derbyshire	Derby	4
Cambridgeshire	Ely	4
Huntingdonshire	Huntingdon	4
Norfolk	Lynn	4
Wiltshire	Salisbury	4
Kent	Sandwich	4
Suffolk	Sudbury	4
Yorkshire, North Riding	Thornton Dale and Ellerburn	4
Middlesex	Westminster	4
Norfolk	Wiggenhall St German	4
Somerset	Bath	3
Norfolk	Bircham	3
Dorset	Blandford	3
Buckinghamshire	Brickhill	3
Glamorganshire	Cardiff	3
Kent	Dover	3
Worcestershire	Droitwich	3
Suffolk	Fornham	3
Hertfordshire	Hertford	3
Essex	Maldon	3
Nottinghamshire	Nottingham	3
Berkshire	Reading	3
Kent	Rochester	3
Lincolnshire	Saltfleetby	3
Huntingdonshire	Sawtry	3
Dorset	Shaftesbury	3
Lincolnshire	Wainfleet	3
Dorset	Wareham	3
Berkshire	Windsor	3
Berkshire	Abingdon	2
Cambridgeshire	Abington	2
Norfolk	Alpington and Yelverton	2
Hampshire	Alresford	2
Devon	Axminster and Uplyme	2
Kent	Barming	2
Oxfordshire	Barton	2
Norfolk	Bawburgh and Bowthorpe	2
Norfolk	Beckham	2
Norfolk	Beechamwell	2
Norfolk	Beeston and Bittering	2
Sussex	Bersted and Pagham	2
Northamptonshire	Boddington	2
Somerset	Brewham	2
Berkshire	Bucklebury Stanford	2
Suffolk	Bungay	2
Suffolk	Bury St Edmunds	2
Cumberland	Carlisle	2
Carmarthenshire	Carmarthen	2
Wiltshire	Cheverell	2
Wiltshire	Chitterne	2
Wiltshire	Codford	2
Wiltshire	Collingbourne	2
Warwickshire	Coventry	2
Northamptonshire	Cranford	2
Wiltshire	Cricklade	2
Devon	Dartmouth	2
Kent	Deptford	2
Dorset	Dorchester	2
Worcestershire	Evesham	2
Yorkshire, West Riding	Ferry Fryston	2
Gloucestershire	Forest Of Dean	2
Norfolk	Forncett	2
Norfolk	Glandford and Bayfield	2
Lincolnshire	Great Limber and Brocklesby	2
Worcestershire	Great Witley and Martley	2
Suffolk	Hargrave and Southwell Park	2
Yorkshire, East Riding	Hull	2
Suffolk	Icklingham	2
Norfolk	Lamas and Little Hautbois	2
Cornwall	Landrake and St Erney	2
Cornwall	Launceston	2
Wiltshire	Lavington	2
Leicestershire	Leicester Forest	2
Norfolk	Long Stratton	2
Lincolnshire	Ludford	2
Dorset	Lulworth	2
Dorset	Lytchett	2
Wiltshire	Manningford	2
Wiltshire	Marlborough	2
Lincolnshire	Mumby	2
Suffolk	Newmarket	2
Wiltshire	Orcheston	2
Norfolk	Oxwick and Pattesley	2
Pembrokeshire	Pembroke	2
Cornwall	Perranuthnoe and St Hilary	2
Worcestershire	Pershore	2
Northamptonshire	Peterborough	2
Somerset	Pilton and North Wootton	2
Devon	Plymouth	2
Devon	Plympton	2
Norfolk	Poringland	2
Norfolk	Ranworth With Panxworth	2
Nottinghamshire	Retford	2
Kent	Romney	2
Norfolk	Rudham	2
Wiltshire	Savernake	2
Yorkshire, West Riding	Sawley and Tosside	2
Wiltshire	Sherston	2
Lincolnshire	Sleaford	2
Kent	Snodland and Paddlesworth	2
Lincolnshire	Somercotes	2
Norfolk	Somerton	2
Norfolk	South Walsham	2
Norfolk	Sporle and Palgrave	2
Middlesex	St Andrew Holborn and St George The Martyr	2
Cornwall	St Columb	2
Middlesex	St Giles in the Fields and St George Bloomsbury	2
Lincolnshire	Stoke	2
Buckinghamshire	Stony Stratford	2
Herefordshire	Sutton	2
Nottinghamshire	Sutton Bonington	2
Glamorganshire	Swansea	2
Somerset	Taunton	2
Herefordshire	Tedstone	2
Norfolk	Terrington	2
Norfolk	Thetford	2
Wiltshire	Tisbury	2
Norfolk	Upton and Fishley	2
Norfolk	Walpole	2
Norfolk	Walton	2
Norfolk	Warham	2
Warwickshire	Warwick	2
Norfolk	Weasenham	2
Suffolk	Whelnetham	2
Dorset	Whitchurch and Catherson	2
Cambridgeshire	Wisbech	2
Norfolk	Witchingham	2
Norfolk	Wretham	2

A.2 Variable construction

Riots before Swing (1758-1829). We collect new data on pre-1830 arsons and machine attacks from the British Library and Findmypast (2016).[4] We search for the words 'arson' and 'machine attack' within the universe of articles published in one of the 60 regional newspaper printed between 1750 and 1832. The search yielded a total of 6,392 articles for 'arson' and 15,986 articles for 'machine attack.' We read in full each of the 'arson' articles and a 35% random sample of the 'machine attack' articles. We first determine whether an article describes a recent episode of civil unrest. If it does, we manually geo-locate the event on the map of England (Southall and Burton, 2004). The final database contains 610 episodes of arson and 69 attacks on machines between 1758 and 1829. We validate this data by looking for similar articles during the Swing riots of 1830-32, and by comparing these episodes with Swing riots coded as 'arson' or 'attacks on machines' in the database compiled by Holland (2005). Both arsons and attacks on machines are strongly correlated in the two data sources: the t-stat of a regression of arsons is 4.53; the t-stat of a regression of machine attacks is 8.09.

Swing riots (1830-32). Data on Swing riots comes from a database compiled by the Family and Community Historical Research Society (Holland, 2005). It contains a comprehensive list of Captain Swing incidents between January 1830 and December 1832. The information comes from judicial records and historical newspapers and contains date, parish, and type of crime perpetrated by rioters. We consider only episodes that occurred between August 1830 and December 1832. For each of these episodes, we manually match the parish of the riot to the historical map of English and Welsh parishes (Southall and Burton, 2004). On this map, we identify the location of these riots by county (variable GAZ_CNTY) and either the name of the parish (variable PAR) or the name of the place (variable PLA). In our baseline results, we use a variable that contains every episode listed in the database, irrespective of the nature of the protest.

Attacks on threshing machines during Swing (1830-32). This is a subset of the Swing riots from Holland (2005). We classify as attack on a threshing machine every event that was recorded as "MACHINE BREAKING (Threshing machines).

Threshing machine adoption (1800-29). We assemble a list of threshing machines in use before the riots from two data sources. The first is built from threshing machine advertisements found in English and Welsh newspapers. The second are the reports of threshing machines in the *General Views of Agriculture*. We collect newspaper advertisements from the *British Newspaper Archive* compiled by the British Library and Findmypast (2016).[5] Within the universe of all articles published by the 60 regional newspaper in the database between 1800 and 1830, we search for the exact string 'threshing machine.' We restrict our search to articles classified as either 'advertisement' or 'classifieds.' Next, we read in full each article retrieved. We use all information from any article that advertises the sale or the lease of a threshing machine or of a farm that lists a threshing machine among its assets. In one case,

[4]See: http://www.britishnewspaperarchive.co.uk/. We collected these articles during the spring of 2019.
[5]We collected these articles during the spring of 2016.

we also exploit the information provided by a threshing machine manufacturer, who lists names and locations of his clients. These clients are farmers located in parishes all over the country (see Figure 8). We drop all advertisements of threshing machine producers that only provide information about the location of the factory, usually an industrial town. We also only consider ads for a single threshing machine whenever we find the same advertisement printed more than once. We manually geo-locate the farm mentioned in each advertisement, based on the map prepared by Southall and Burton (2004).

We complement this source with a list of threshing machines in the *General Views of Agriculture*, covering all English counties. In the second edition, the volume for each county contains an entire chapter on threshing machines, relating information on every machine including the name of the owner and its place of operation. We locate each of these machines on the map of Southall and Burton (2004) and ensure that we do not double count any machine from the newspapers, comparing the names of the owners in the two sources. Whenever we link a parish to either an advertisement or a machine from the *General Views*, we add 1 to the count of threshing machines in a parish.

Density (1801-31). Parish population comes from the decennial Censuses of England of 1801-31 (Southall et al., 2004). The original variables are POP_1801 in 1801 and TOT_POP in the other years. We merge census information, geolocating parishes on the historical map of English and Welsh parishes by the Census variables county (ANC_CNTY) and parish (ANC_PAR). The total area of the parish (in square km) is calculated with ArcGIS based on the map of historical parishes of England and Wales described in Appendix A.1. Density is population per square km. We use the natural logarithm of this variable in all regressions.

Sectoral shares (1801-31). We construct sectoral shares from data in the decennial Censuses of England, using the years 1801-31 (Southall et al., 2004). We calculate three shares: for agriculture, trade and other activities. In 1801 these shares reflect the number of workers employed in these three sectors (we use the variables OC_AGRIC, OC_TRADE and OC_OTHER). For the other years the shares represent the share of families chiefly employed in the three sectors (we use the variables $FAMAGRI$, $FAMTRADE$ and $FAMOTHER$). The data in Southall et al. (2004) do not allow to calculate other shares. Census data come at the parish level and we merge it to the historical map of English and Welsh parishes as we did with the population.

Sex ratio (1801-31). We compute the sex ratio using data from the the decennial Censuses of England of 1801-31. The variable is equal to the total number of men (variable MA_1801 in 1801, TOT_MALE in the other years) divided by the total number of women (variable FE_1801 in 1801, TOT_FEM in the other years). Census data are available at the parish level, and we geo-locate parishes on the historical map of English and Welsh parishes as we did with the population. We use the natural logarithm of this variable in all regressions.

Share of land cultivated with cereals (1801). The 1801 Corn Returns record land use information for almost 4000 parishes (Turner, 2005). We merge the Crop Returns to the historical map of English and Welsh parishes using the Census variables county (ANC_CNTY) and parish (ANC_PAR). We construct the share of land cultivated with cereals as the sum of the area devoted to wheat, oat, barley and rye (variables $WHEAT$, $OATS$, $BARLEY$ and RYE) divided by the total area cultivated.

Ratio of sales of wheat to oat. Brunt and Cannon (2013) digitized information from the crop returns. Their database records weekly information on quantity and value sold for different crops across 174 English market towns in 1820-30. We assign each English parish to the closest market town based on the distance to the centroid of the parish. We construct two ratios. The first is the ratio of the average value of wheat sold to the average value of oat sold. The second is the ratio of the average quantity of wheat sold to the average quantity of oat sold.

Distance to Elham (first riot). We construct this variable as the straight-line distance of the centroid of every parish in our map to Elham, the parish that saw the first episode of the Swing riots according to Griffin (2012). We use the natural logarithm of this variable in all regressions.

Distance to closest town with a newspaper. To construct this variable, we first determine which of the newspapers in *British Newspaper Archive* was in print before 1830. Next, we manually geo-locate the cities in which these newspapers were printed. We then calculate the straight-line distance of the centroid of every parish in our map to each of these cities. Finally, we keep only the distance to the closest city. We use the natural logarithm of this variable in all regressions.

Distance to closest manufacturing city. We consider 15 manufacturing centers in 1801: Stockport in Cheshire, Blackburn, Bolton-le-Moors, Liverpool, Manchester, Oldham, Preston and Whalley in Lancashire, London, Norwich in Norfolk, Wolverhampton and Birmingham in Warwickshire and three cities in Yorkshire, West Riding: Halifax, Leeds and Sheffield. We identify these cities by selecting the top 15 parishes in terms of 1801 share of employment in "trade among those that had at least 18,000 inhabitants in 1801. In the 1801 census, these centers had on average 46 percent of workers employed in trade and less than 2.7 percent employed in agriculture. In the rest of English parishes, 11.6 percent of workers were chiefly employed in trade and 38.6 percent in agriculture. We use the coordinates of the centroid of these cities and of every parish in England to construct the straight-line distance of every parish to the closest of manufacturing center. We then divide the sample into two groups: above and below the median distance to these cities. The median parish in terms of distance to manufacturing cities is Waterstock in Oxfordshire, which lies 74 km from Blackburn.

Share of heavy soil. Heavy soils are soils rich in clay and to a lesser extent loam. For every parish we take the share under heavy soil of all the cells that fall inside the parish. To calculate it, we collect information on soil composition from the *British Geological Survey Soil Parent Material Model* (British Geological Survey, 2009). The dataset focuses upon the material from which top soils and subsoils develop (A and B horizons). The original data is a raster that covers the land mass of Britain on a grid of 1×1 km. We superimpose the raster on our historical map of English and Welsh parishes by intersecting every cell of the raster with the parish it falls in. We use the soil group variable to classify cells into light and heavy soils. Light soils are soils rich in sand and silt.

Cereal suitability index. We construct our own cereal suitability index based on detailed weather data and an agronomic model from FAOs ECOCROP.[6] Weather data is from Hijmans et al. (2005*a*,*b*): they provide average monthly precipitation and three average monthly temperatures (minimum, maximum and mean) over a grid of 30 × 30 arc-seconds. Averages are computed over the years 1960-90. We use these data to estimate cereal suitability following Wigton-Jones (2019): Appendix A.3 describes the procedure in more detail. It yields an index for every grid cell covering England and Wales: We resample this raster on a grid of 2.88 arc-seconds with the "nearest method. Next, we superimpose this raster on our historical map of English and Welsh parishes. For every cell of the raster we take the centroid and assign it to the parish where the centroid falls. Finally, for each parish we take the average index of all the cells that fall inside the parish.

Abnormal precipitation (spring and summer 1830) and temperature. We take historical precipitation from Pauling et al. (2006). They used documentary evidence and natural proxies to estimate seasonal precipitation for the period 1500-1900 over a 0.5 × 0.5 degrees grid covering Europe (approximately 55.5 × 55.5 km). To construct abnormal precipitation in the spring (summer) of 1830 across England and Wales, we take average spring (summer) precipitation in 1830 and subtract the average spring (summer) precipitation in the years 1800-1828. We do this for every cell that covers the British Isle, obtaining a new raster with the abnormal precipitation in the spring (summer) of 1830. Next, we resample this raster on a finer grid of 88.8 × 88.8 m with the "nearest" method, and superimpose it onto our historical map of English and Welsh parishes. For every cell of the raster, we take its centroid and assign it to the parish witin which the centroid falls. Finally, for every parish we calculate the average abnormal precipitation in the spring (summer) of 1830 of every cell that falls inside the parish.

For abnormal temperature, we follow the same procedure using historical temperature data from Luterbacher et al. (2004).

Share of land enclosed (1820). Data on enclosures are from Gonner (1912, p.270-78), who reports information on the percentage of common land that was enclosed before 1870. Gonner collected information across 345 'registration districts' covering 6,715 parishes. In order to estimate the percentage of land enclosed in 1820, we combine the information on this table with information from the table on page 279-281 of the same book. In this second table, Gonner reports the share of land in commons enclosed in each decade between 1760 and 1870 for every county in England and Wales. We estimate the share of land enclosed in 1820 by multiplying district-level enclosures in 1870 with the proportion of enclosures that happened before 1820 in the county of every district. We use the registration district reported in the 1801 Census to match each parish to its registration district. The parishes in the registration districts of Biggleswade (Bedford), Billericay, Colchester, Ongar, Romford (Essex) and Market Harborough (Leicester) have the median level of enclosure. We define parishes with 'high' enclosures those parishes with more than this level of enclosures.

Poor Rates per capita (1801). We calculate poor relief based on data from the "Poor

[6] See http://ecocrop.fao.org/ecocrop/srv/en/home.

Law Commissioners Report compiled by the 1832 Royal Commission on the Operation of the Poor Laws, published in 1834.[7] The report is based on assistant commissioners sent all across the country to collect information, combined with questionnaires directly sent to parishes. Returns are available for 1,391 parishes. We have valid information on poor rates and population for 1,251 of these parishes. From the report, we digitized the population in 1801 (first entry of question A on the questionnaire) and *Poor Rates* collected in 1803 (first entry of question B on the questionnaire). The variable is calculated as the total value of poor rates in 1803 divided by the 1801 population in the parish.

Unemployment (winter and summer 1834). We collect data on winter and summer unemployment from the same "Poor Law Report" of 1834. To reconstruct parish-level unemployment, we digitize the answers to question 5 and 6.[8] Question 5 reads: 'number of agricultural labourers in your parish?'; question 6 reads: 'number of labourers generally out of employment, and how maintained in summer and in winter?' We construct unemployment as the number of labourers out of employment divided by the total number of labourers. We calculate this ratio separately for winter and for summer, and we set to missing 6 parishes where unemployment is above 100 percent. We construct relative unemployment as the difference between winter and summer unemployment.

[7]Full title: *Report from his Majestys commissioners for inquiring into the administration and practical operation of the Poor Laws.*

[8]Officials were sent to survey parishes in 3 different waves between 1833 and 1834, and the questionnaire they used varied slightly between these waves. Question 5 and 6 in the first two issues became question 6 and 7 in the 3rd issue.

A.3 Cereal suitability index

This section describes the construction of our cereal suitability index from the FAOs agronomic model ECOCROP.[9] It follows closely the work of Wigton-Jones (2019).

1. The index requires the following 8 parameters:
   - minimum temperature ($\underline{\theta}$): temperature below which cereals die;
   - optimal temperature range ($\underline{\theta}^*$—$\overline{\theta}^*$): optimal temperature range for growing cereals;
   - maximum temperature ($\overline{\theta}$): temperature above which cereals die;
   - minimum rainfall ($\underline{\rho}$): cumulative rainfall during growing season below which cereals die;
   - optimal rainfall range ($\underline{\rho}^*$—$\overline{\rho}^*$): optimal cumulative rainfall range during growing season;
   - maximum rainfall ($\overline{\rho}$): cumulative rainfall during growing season above which cereals die.

2. We use these parameters together with average monthly temperature (T_m^{avg}) and rainfall (R_m^{avg}) to construct two sets of monthly indexes: temperature suitability (I_m^T) and rainfall suitability (I_m^R). The indexes take the following values:

$$I_m^T = \begin{cases} 0 & \text{if} \quad & T_m^{\text{avg}} & < \underline{\theta} \\ f_1(T_m^{\text{avg}}) & \text{if} \quad \underline{\theta} \leq & T_m^{\text{avg}} & < \underline{\theta}^* \\ 1 & \text{if} \quad \underline{\theta}^* \leq & T_m^{\text{avg}} & < \overline{\theta}^* \\ f_2(T_m^{\text{avg}}) & \text{if} \quad \overline{\theta}^* \leq & T_m^{\text{avg}} & < \overline{\theta} \\ 0 & \text{if} \quad \overline{\theta} \leq & T_m^{\text{avg}} & \end{cases}$$

$$I_m^R = \begin{cases} 0 & \text{if} \quad & R_m^{\text{avg}} & < \underline{\rho} \\ g_1(R_m^{\text{avg}}) & \text{if} \quad \underline{\rho} \leq & R_m^{\text{avg}} & < \underline{\rho}^* \\ 1 & \text{if} \quad \underline{\rho}^* \leq & R_m^{\text{avg}} & < \overline{\rho}^* \\ g_2(R_m^{\text{avg}}) & \text{if} \quad \overline{\rho}^* \leq & R_m^{\text{avg}} & < \overline{\rho} \\ 0 & \text{if} \quad \overline{\rho} \leq & R_m^{\text{avg}} & \end{cases}$$

3. We choose the functions $f_1(T^{\text{avg}})$, $f_2(T^{\text{avg}})$, $g_1(R^{\text{avg}})$ and $g_2(R^{\text{avg}})$ so that the index function is linear and continuous (see Figure 1).

4. We also set $I_m^T = 0$ whenever the mean maximum (minimum) temperature rises above the maximum (falls below the minimum) temperature that kills cereals.

5. We obtain monthly indexes by multiplying temperature and rainfall indexes: $I_m = I_m^T \times I_m^R$.

[9] See http://ecocrop.fao.org/ecocrop/srv/en/home.

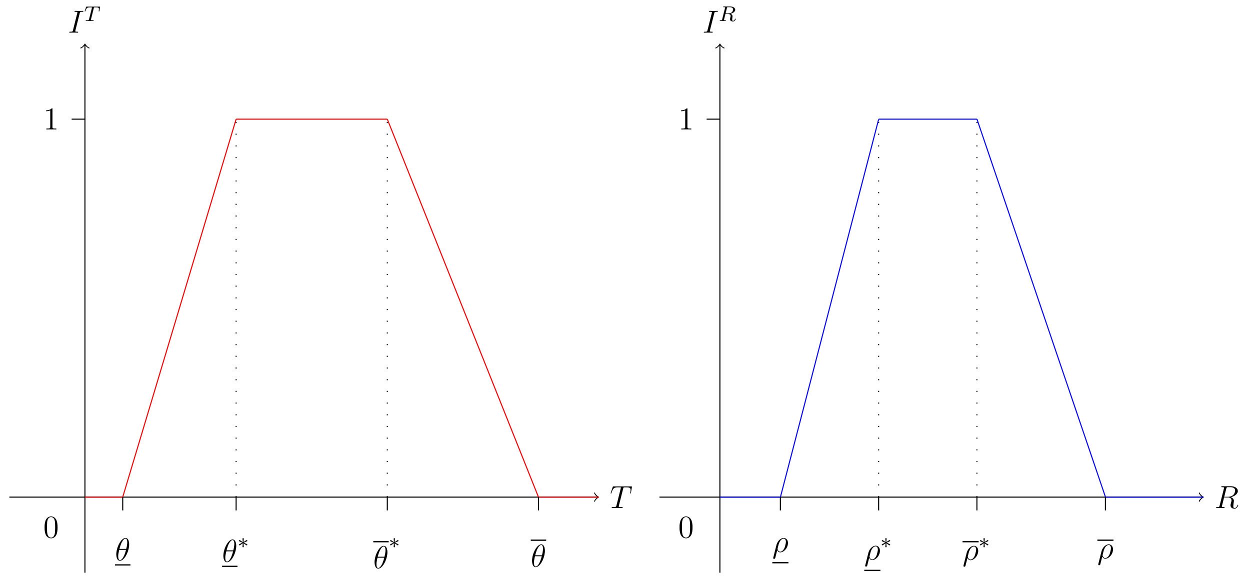


Figure 1: Examples of temperature and rainfall suitability indexes

6. Cereals need 100-120 days to grow. As Wigton-Jones (2019), we do not take a stance on which month the growing season should start. Instead, we calculate separate indexes for each of the 12 months. We consider that during any spell of 4 consecutive months, the worse conditions will determine productivity (Liebig's law). Thus, for every month we take the minimum suitability index among the 4 months starting then: this is the index of that growing season. We assume that farmers will select the best growing season among the 12 possible, and take the highest of the 12 indexes to be the suitability index.

The FAO provides parameters for 4 cereals: wheat (*triticum aestivum*), oat (*avena sativa*), barley (*hordeum vulgare*) and rye (*secale cereale*). However, it provides no parameter for cereals as a whole. Because we want to capture weather conditions that make cultivation of *any* cereal possible, for every parameter we select the most constraining among the values provided for the 4 cereals. Table 2 provides the parameters of the four crops and the combined parameter for the cereal family.

Figure 2 plots variations in cereal suitability across England.

Table 2: FAO's ECOCROP parameters.

		Wheat	Oat	Barley	Rye	Cereals
Minimum temperature (°C)	$\underline{\theta}$	5	5	2	3	5
Minimum optimal temperature (°C)	$\underline{\theta}^*$	15	16	15	15	16
Maximum optimal temperature (°C)	$\overline{\theta}^*$	23	20	20	20	20
Maximum temperature (°C)	$\overline{\theta}$	27	30	40	31	27
Minimum rainfall (mm)	$\underline{\rho}$	99	82	66	132	132
Minimum optimal rainfall (mm)	$\underline{\rho}^*$	247	197	164	197	247
Maximum optimal rainfall (mm)	$\overline{\rho}^*$	296	329	329	329	296
Maximum rainfall (mm)	$\overline{\rho}$	526	493	658	658	493

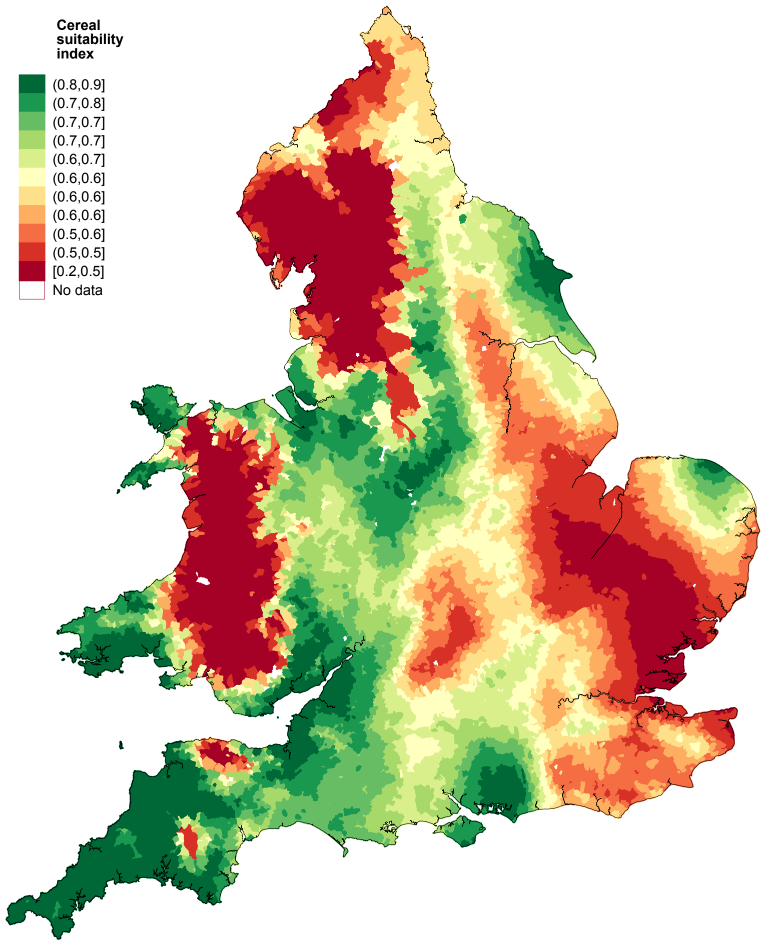


Figure 2: Cereal suitability index

Notes. Cereal suitability index. Source: own calculation based on weather data from Hijmans et al. (2005*b*) and parameters from the FAO-ECOCROP model.

A.4 Historical weather in England and Wales

We compute a cereal suitability index with weather records from Hijmans et al. (2005*b*). One possible concern with this procedure is that it uses average weather conditions for the period 1961-1990, which may be different from weather conditions that affected cereal suitability in 1800-30. To determine how much weather changed over the last 200 years, we perform two separate tests.

In the first one, we use historical records of temperature and precipitation on a 0.5°× 0.5°grid that covers Europe[10] to compare average temperature and precipitation in the period 1801-1830 and 1961-1990. The four panels of Figure 3 plot average temperature in the years 1801-1830 (on the x-axes) against the average temperature in the period 1961-1990 (on the y-axes) for the four seasons of the year across the 135 cells that cover England and Wales. The four panels of Figure 4 repeat the exercise for precipitation, and Table 3 reports correlations for the two variables. The data suggest that weather did not change much across England in the last 200 years. In any given season, cells that were on average colder (wetter) in 1800-1830, are still so in 1960-1990. Moreover, the correlation between the two periods of average temperature (precipitation) is always above 99% (98%).

Table 3: Correlation between weather in 1801-1830 and weather in 1961-1990.

	Temperature	Precipitation
Winter	99.78%	99.48%
Spring	99.45%	98.68%
Summer	99.50%	99.13%
Fall	99.95%	98.69%
Observations	135	135

Notes. The first column reports the correlation for temperature and the second column for precipitation. All correlations are significant at < 0.001 level.

[10] Luterbacher et al. (2004) and Xoplaki et al. (2005) describe the construction of temperature records, and Pauling et al. (2006) describe the construction of precipitation data.

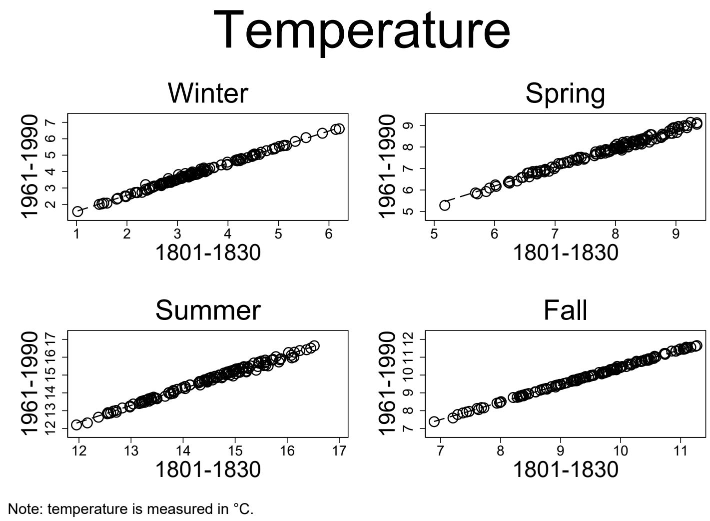


Figure 3: Average temperature by season.

Notes. The figure plots average temperature across England and Wales in the period 1801-1830 (on the x-axes) against the average temperature in the period 1961-1990 (on the y-axes) for the four seasons of the year. Source: Luterbacher et al. (2004) and Xoplaki et al. (2005).

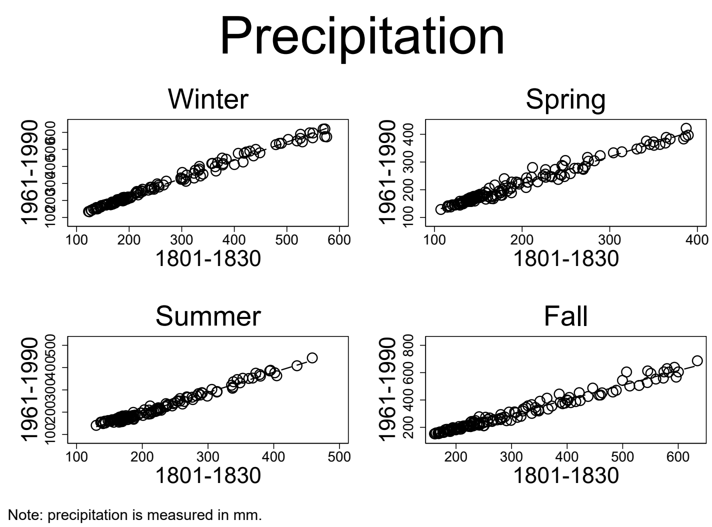


Figure 4: Average precipitation by season.

Notes. The figure plots average precipitation across England and Wales in the period 1801-1830 (on the x-axes) against the average temperature in the period 1961-1990 (on the y-axes) for the four seasons of the year. Source: Pauling et al. (2006).

One possible concern with this analysis is that historical weather data are estimated rather that observed. Moreover, data are available only for separate seasons, not for separate months. To address this concern we perform a second test, using the historical series maintained by the Hadley Centre at the UK Meteorological Office (Alexander and Jones, 2000; Met Office Hadley Centre, 2001). The office collects monthly precipitation records across England and Wales since 1700. Thus, it allows us to compare monthly records obtained from actual observations. We use these data to compare the average monthly precipitation during 1801-1830 with the average monthly precipitation in the years 1961-1990. Figure 5 plots these averages for the two periods along with their 95 percent intervals.

The graph confirms that precipitation did not change much in England over the last 200 years. Average yearly precipitation is not significantly different in 1961-90 relative to the 30 years leading up to the Swing riots. Unfortunately, precipitation is the only weather variable for which the Hadley Centre preserves historical records. Moreover, these records are admittedly noisy, and are available only for the whole England. Nevertheless, the analysis of these records, together with the previous analysis, suggest that weather in 1961-1990 is a valid proxy for weather at the beginning of 1800.

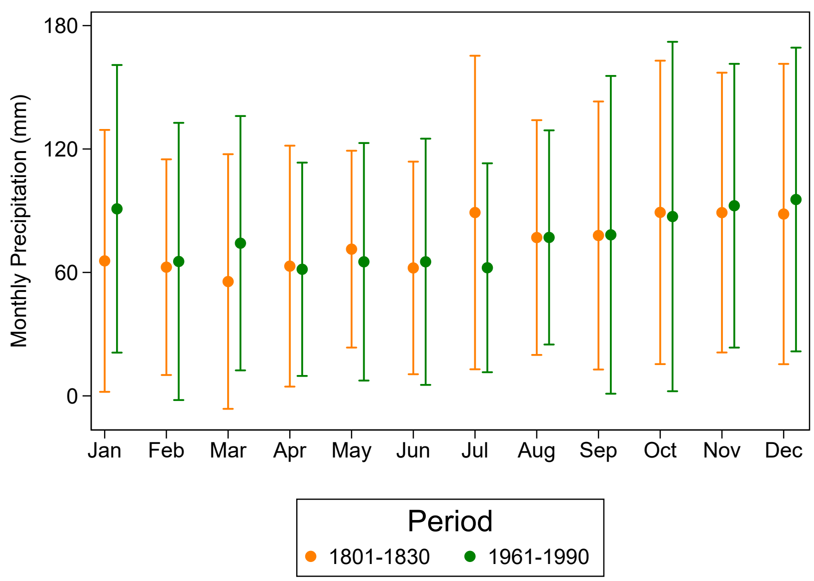


Figure 5: Precipitation by month.

Notes. The figure plots the average monthly precipitation across England and Wales over the period 1801-1830 (in orange) and over the period 1961-1990 (in green). The bar identify 95 percent intervals. The average yearly precipitation in 1801-1830 was 891mm: this is not significantly different from the average yearly precipitation in 1961-1990, which was 915m (difference: 23,96 mm, s.e.: 24.72). Source: Met Office Hadley Centre (2001): http://www.metoffice.gov.uk/hadobs/hadukp/.

B Additional results

B.1 Additional figures

The figures in this section provide additional background on the temporal distribution of Swing riots and on our measure of threshing machine adoption.

Figure 6 plots frequency of Swing riots by month, differentiating between machine attacks and other forms of unrest. The graph uses information on the date of the riot and the type of event from Holland (2005).

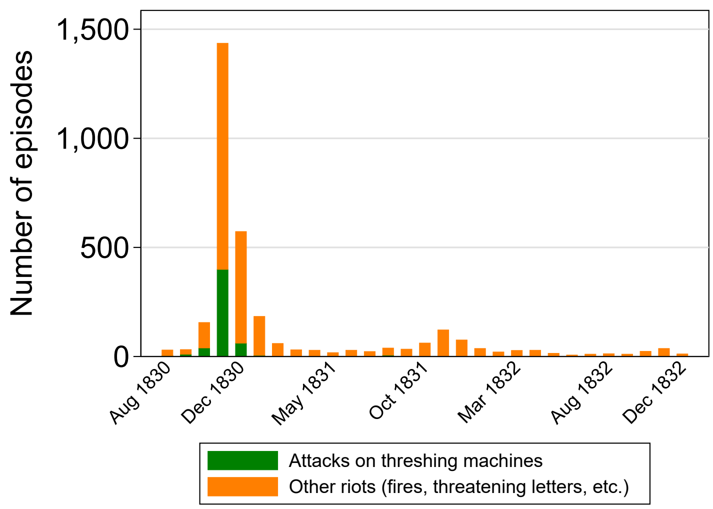
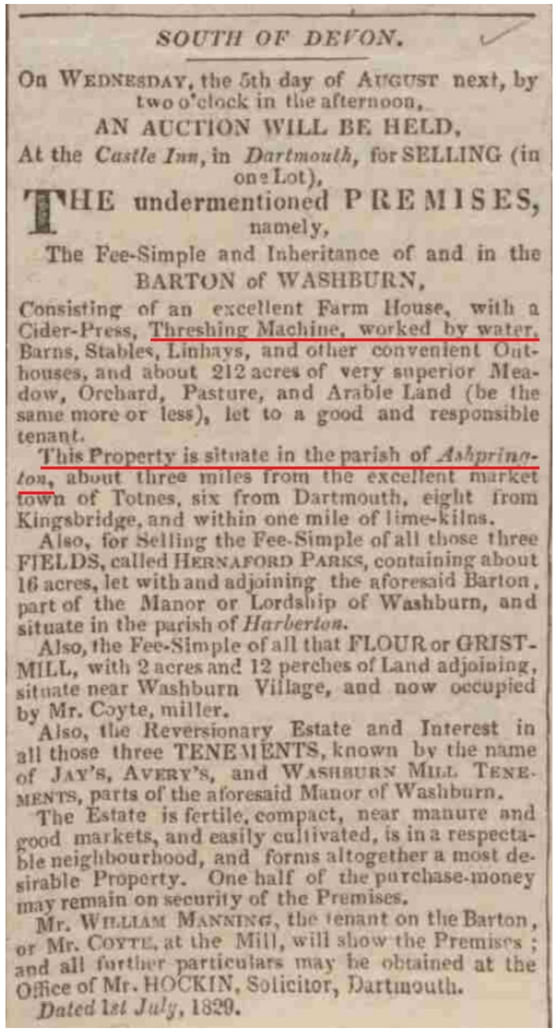


Figure 6: Swing riots over time.

Notes. In green: attacks on threshing machines. In orange: all other riots associated to Swing: including threatening letters and arson attacks. Source: Holland (2005).

Figure 7 represents the typical advertisement in our database of threshing machine adoption: it publicizes a farm on sale in the parish of Ashprington (Devon). The ad lists a 'threshing machine' among its assets (highlighted). Figure 8 is a different type of advertisement: it is published by a *manufacturer* of threshing machines and lists names and location of clients who purchased one of these machines in the past. We classify each of the parishes of the clients as having one threshing machine. This was the only article of this kind we found in the newspapers of British Library and Findmypast (2016).



SOUTH OF DEVON.

On WEDNESDAY, the 5th day of AUGUST next, by two o'clock in the afternoon,

AN AUCTION WILL BE HELD,

At the *Castle Inn*, in *Dartmouth*, for SELLING (in one Lot),

THE undermentioned PREMISES, namely,

The Fee-Simple and Inheritance of and in the BARTON of WASHBURN,

Consisting of an excellent Farm House, with a Cider-Press, Threshing Machine, worked by water, Barns, Stables, Linhays, and other convenient Out-houses, and about 212 acres of very superior Meadow, Orchard, Pasture, and Arable Land (be the same more or less), let to a good and responsible tenant.

This Property is situate in the parish of *Ashprington*, about three miles from the excellent market town of Totnes, six from Dartmouth, eight from Kingsbridge, and within one mile of lime-kilns.

Also, for Selling the Fee-Simple of all those three FIELDS, called HERNAFORD PARKS, containing about 16 acres, let with and adjoining the aforesaid Barton, part of the Manor or Lordship of Washburn, and situate in the parish of *Harberton*.

Also, the Fee-Simple of all that FLOUR or GRIST-MILL, with 2 acres and 12 perches of Land adjoining, situate near Washburn Village, and now occupied by Mr. Coyte, miller.

Also, the Reversionary Estate and Interest in all those three TENEMENTS, known by the name of JAY'S, AVERY'S, and WASHBURN MILL TENEMENTS, parts of the aforesaid Manor of Washburn.

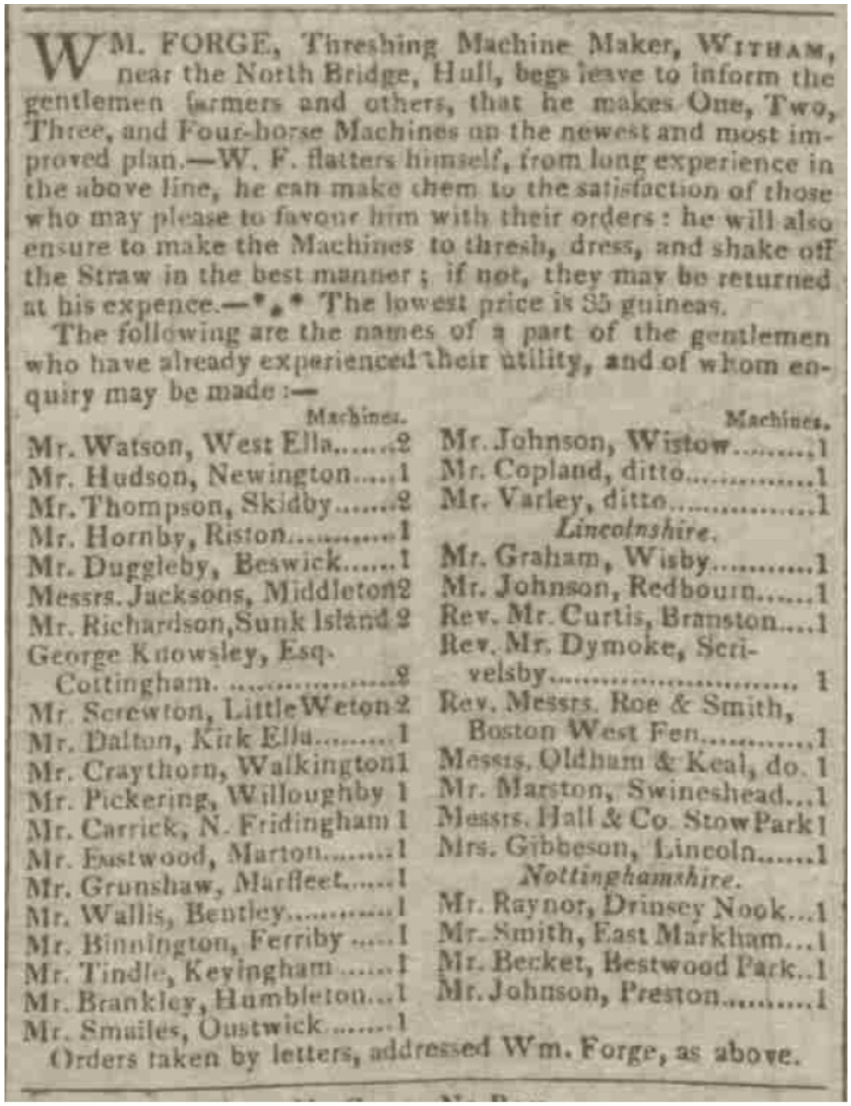
The Estate is fertile, compact, near manure and good markets, and easily cultivated, is in a respectable neighbourhood, and forms altogether a most desirable Property. One half of the purchase-money may remain on security of the Premises.

Mr. WILLIAM MANNING, the tenant on the Barton, or Mr. COYTE, at the Mill, will show the Premises; and all further particulars may be obtained at the Office of Mr. HOCKIN, Solicitor, Dartmouth.

Dated 1st July, 1829.

Figure 7: Example of an advertisement for a 'threshing machine'

Notes. On July the 1st, 1829, the Sherborne Mercury advertised the sale of a farm in the parish of Ashprington (Devon). We count this advertisement as an indication that threshing machines are used in this parish because the farm includes a 'threshing machine' among the assets that went on sale. Source: British Library and Findmypast (2016).



WM. FORGE, Threshing Machine Maker, WITHAM, near the North Bridge, Hull, begs leave to inform the gentlemen farmers and others, that he makes One, Two, Three, and Four-horse Machines on the newest and most improved plan.—W. F. flatters himself, from long experience in the above line, he can make them to the satisfaction of those who may please to favour him with their orders: he will also ensure to make the Machines to thresh, dress, and shake off the Straw in the best manner; if not, they may be returned at his expence.—*** The lowest price is 35 guineas.

The following are the names of a part of the gentlemen who have already experienced their utility, and of whom enquiry may be made:—

	Machines.
Mr. Watson, West Ella	2
Mr. Hudson, Newington	1
Mr. Thompson, Skidby	2
Mr. Hornby, Riston	1
Mr. Duggleby, Beswick	1
Messrs. Jacksons, Middleton	2
Mr. Richardson, Sunk Island	2
George Knowsley, Esq. Cottingham	2
Mr. Screwton, Little Weton	2
Mr. Dalton, Kirk Ella	1
Mr. Craythorn, Walkington	1
Mr. Pickering, Willoughby	1
Mr. Carrick, N. Fridingham	1
Mr. Eastwood, Marton	1
Mr. Grunshaw, Marfleet	1
Mr. Wallis, Bentley	1
Mr. Binnington, Ferriby	1
Mr. Tindle, Keyingham	1
Mr. Brankley, Humbleton	1
Mr. Smailes, Oustwick	1

	Machines.
Mr. Johnson, Wistow	1
Mr. Copland, ditto	1
Mr. Varley, ditto	1
Lincolnshire.	
Mr. Graham, Wisby	1
Mr. Johnson, Redbourn	1
Rev. Mr. Curtis, Branston	1
Rev. Mr. Dymoke, Scrivelsby	1
Rev. Messrs. Roe & Smith, Boston West Fen	1
Messrs. Oldham & Keal, do.	1
Mr. Marston, Swineshead	1
Messrs. Hall & Co. Stow Park	1
Mrs. Gibbeson, Lincoln	1
Nottinghamshire.	
Mr. Raynor, Drinsey Nook	1
Mr. Smith, East Markham	1
Mr. Becket, Bestwood Park	1
Mr. Johnson, Preston	1

Orders taken by letters, addressed Wm. Forge, as above.

Figure 8: Example of an advertisement.

Notes. On February the 2nd, 1808, the *Stamford Mercury* published the notice of William Forge, a threshing machine maker, who advertised his product by suggesting to contact one of his past customers. We code each of the parishes listed above as parishes in which at least one threshing machine is in operation. Source: British Library and Findmypast (2016).

B.2 Threshing machines and the labor market

Manual threshing was a winter activity, and employed men for most of that season (Hobsbawm and Rudé, 1969). Do we find evidence of greater winter unemployment as a result of threshing machine adoption? Checkland (1974) reports total and unemployed workers in winter and summer for some 600 parishes in 1832. We use this data to compute the average difference in unemployment between winter and summer. In Table 4 we regress this difference against our measure of threshing machine adoption. Unemployment was on average 5.5% higher in winter than in summer. In parishes with a threshing machine, this difference was 2% higher. The result holds unconditionally (col. 1), with controls (cols. 2-3) and with controls and region fixed effects (col. 4). These results confirm that threshing machines brought technological unemployment during the winter season.

Table 4: Threshing machines and the labor market.

	Unemployment: winter - summer			
	(1)	(2)	(3)	(4)
No. of threshers	0.025	0.021	0.022	0.019
	[0.007]	[0.007]	[0.007]	[0.008]
log 1801 density		0.021	0.014	0.012
		[0.006]	[0.006]	[0.006]
Share of agricultural workers in 1801			-0.017	-0.023
			[0.016]	[0.016]
log 1801 sex ratio			-0.032	-0.031
			[0.032]	[0.031]
log distance to Elham			-0.033	-0.022
			[0.009]	[0.014]
log distance to newspaper			0.011	0.013
			[0.006]	[0.006]
Region fixed effects (5)	No	No	No	Yes
R^2	0.010	0.032	0.081	0.091
Mean dependent variable	0.055	0.055	0.055	0.055
Observations	574	574	574	574

Notes: Threshing machines and the labor market. The dependent variable in is winter unemployment rate minus summer unemployment rate. Robust standard errors in parentheses.

B.3 Type of unrest

In this section we use rich data on types of unrest from Holland (2005) to better understand the relationship between threshing machines and riots. We break down riots into two categories: attacks on threshing machines, and other type of revolt. We then estimate Equation (1) in the main text with these two measures. Cols. 1-2 of Table 5 report results for machine attacks and cols. 3-4 for other types of unrest. That counties with more machines witnessed more attacks on threshers is not too surprising: what is crucial is that these machines spelled higher probabilities for other types of unrest. For both variables, there is a robust correlation between machines and riots. This implies that threshing machines worked as a catalyst of general unrest: the more of them there were, the more protests occurred that were not directly aimed at the machines.

Table 5: Basic correlations: type of unrest.

No. of	Threshers attacked		Other riots	
	(1)	(2)	(3)	(4)
No. of threshers	0.097	0.087	0.292	0.266
	[0.029]	[0.029]	[0.054]	[0.054]
log 1801 density	0.007	0.006	0.094	0.093
	[0.004]	[0.004]	[0.015]	[0.016]
Share of agricultural workers in 1801	0.031	0.027	-0.095	-0.083
	[0.016]	[0.016]	[0.036]	[0.036]
log 1801 sex ratio	-0.038	-0.032	-0.144	-0.161
	[0.012]	[0.013]	[0.036]	[0.037]
log distance to Elham	-0.077	-0.048	-0.248	-0.169
	[0.012]	[0.021]	[0.023]	[0.035]
log distance to newspaper	-0.001	-0.001	0.023	0.020
	[0.005]	[0.005]	[0.016]	[0.017]
Region fixed effects (5)	No	Yes	No	Yes
R^2	0.023	0.026	0.051	0.058
Mean share	0.053	0.053	0.255	0.255
Observations	9674	9674	9674	9674

Notes: Threshers and type of unrest. Estimates of Equation (1) in the main text. Dep. var. is: Cols. 1-2: number of attacks on threshing machines (1830-32); Cols. 3-4: number of 1830-32 riots that did target a threshing machine. Robust standard errors in brackets.

B.4 Machine adoption and unrest over time

In Table 6 we look at machine adoption and unrest over time. This analysis puts the Swing riots in context within the English revolts during the early 19$^{\text{th}}$ century. To measure pre-1830 unrest, we digitized new data from 1750-1829 newspapers. We search for the words 'arson' and 'machine attack' within the universe of articles published in one of the 60 regional newspaper printed in those years. The search yielded a total of 6,392 articles for 'arson' and 15,986 articles for 'machine attack.' We read in full each of the 'arson' articles and a 35% random sample of the 'machine attack' articles. We first determine whether an article describes an episode of civil unrest. If it does, we manually geo-locate the event on the map of England. The final database contains 610 episodes of arson and 69 attacks on machines. For the year 1830-32 we use data from Holland (2005): to maintain comparability with the pre-1830 episodes we only use episodes classified as 'arson' or 'machine attack' (results with all episodes are qualitatively similar).

With this data, we estimate:

$$\text{Riots}_{pt} = \beta_p + \sum_{t=pre1800}^{1830} \beta_{1t}\text{Machines}_{pt} + \beta_2 density_{pt} + \beta_X X_{pt} + \chi_{rt} + e_{pt} \tag{1}$$

The unit of observation is a parish-decade, and we pool all years before 1800 into a single time period. Riots_{pt} is the number of episodes of unrest in the parish-decade and Machines_{pt} is the number of machines we observe in that parish up to that decade. We control for the usual set of covariates: log of density, log sex ratio, share of agricultural workers are from the decadal censuses and are time-varying. Log distance to Elham and log distance to a town with a newspaper do not vary overtime and we interact them with year dummies. No Census exists before 1801: for the decade before 1800, we use 1801 demographic variables interacted with pre-1800 year dummy. In all regressions we control for parish fixed effects, and in the most demanding specification we include 5 regions × year fixed effects: χ_{rt}.

Estimates of Equation (1) are in Table 6. Col. 1 includes only year and parish fixed effects; col. 2 adds density; col. 3 all other controls and col. 4 region × year fixed effects. In all specifications, we find that early threshing machines had a negative but small and insignificant correlation with unrest. This is inconsistent with the idea that early adopters introduced threshing machines in response to violent workers, and may instead suggest that farmers were wary of adopting labor-saving technologies in areas where rioting was likely. By the 1820s, however, we observe a positive and significant association between the existing stock of threshing machines and riots.

These results are consistent with threshing machines leading to a progressive deterioration of living conditions in the countryside. However, because adoption is endogenous, coefficients can not be interpreted causally. We extend the identification strategy to the panel dataset at the end of Section B.5.

Table 6: Correlation between machine adoption and arsons and machine attacks overtime.

	Riots (1780-1832)			
	(1)	(2)	(3)	(4)
Threshers in the 1800s	-0.076	-0.076	-0.073	-0.071
	[0.088]	[0.088]	[0.087]	[0.088]
Threshers in the 1810s	-0.003	-0.004	-0.005	-0.003
	[0.036]	[0.036]	[0.036]	[0.036]
Threshers in the 1820s	0.164	0.163	0.157	0.157
	[0.068]	[0.068]	[0.067]	[0.068]
Threshers in 1830-32	0.224	0.222	0.216	0.201
	[0.054]	[0.054]	[0.054]	[0.054]
Parish & year fixed effects	Yes	Yes	Yes	Yes
log density	No	Yes	Yes	Yes
Other controls	No	No	Yes	Yes
Region (5) $\times$ year fixed effects	No	No	No	Yes
R^2	0.257	0.258	0.271	0.273
Observations	48636	48636	48636	48636

Notes: Threshing machine adoption and pre-1830 riots. Table reports estimates of Equation (1). Dependent variable is number of arsons or attacks on machines (both agricultural and industrial) between 1758 and 1832. The omitted category is pre-1800: for these years we sum all episodes of unrest and keep a single observation for every parish. Data source is British Library and Findmypast (2016) for pre-1830 events and Holland (2005) for 1830-32: see text for details. Standard errors clustered at parish level in brackets.

B.5 Validity of the instrument: additional results

Here, we discuss additional evidence supporting the validity of our IV strategy. First, we show that heavy soils predict prevalence of non-wheat farming. Second, we discuss the balance of the instrument. Finally, we show the correlation of heavy soils and pre-1830 unrest.

Our central claim is that soil heaviness predicts threshing machine adoption because it makes wheat cultivation less attractive compared to cultivation of other cereals. In the early 1800s, the second most cultivated cereal in England was oat. Thereofre, in Table 7, we ask whether areas with heavier soils have on average more oat than wheat. We use data from Brunt and Cannon (2013) on quantity and value of wheat and oat sold in 174 British market towns in the 1820s. For every parish, we construct the average ratio of wheat to oat sold in those years in the closest market town. Cols. 1-4 look at relative values and cols. 5-8 at relative quantities; cols. 1 and 5 present unconditional correlations, cols. 2 and 6 add an index of relative weather suitability between the two crops, cols. 3 and 7 the usual set of controls and cols. 4 and 8 include 5 region fixed effects. We cluster standard errors at the level of the market town. Across specifications, we find that a higher share of heavy soil is associated with a lower wheat-oat ratio. These results confirm that heavy soils make wheat cultivation unattractive relative to the leading alternative at the time (oat).

Table 8 presents the balance of the instrument with respect to observable characteristics. In col. 1 we show the coefficient of the share of soil that is heavy in simple bi-variate regressions. Dependent variables are listed on the left of the table. Coefficients are non-standardized and col. 3 reports the mean value of each dependent variable (standardized beta-coefficient are displayed in panel (b) of Figure 2 in the main text). Col. 2 of Table 8 reports the coefficients of the share of heavy soil in a regression in which we control for cereal suitability (see Section A.3 for details). Heavy soil remains uncorrelated with all variables except distance to Elham, indicating that even conditional on the most important determinant of 1800 agriculture, the instrument is not associated with potential causes of Swing.

Table 9 present estimates from the following panel regression:

$$\text{Riots}_{pt} = \gamma_p + \sum_{t=pre1800}^{1830} \gamma_{1t} \cdot \text{Share heavy}_p + \gamma_2 \text{density}_{pt} + \gamma_X X_{pt} + \chi_{rt} + v_{pt} \qquad (2)$$

where we substitute Machines in Equation (1) with the share of heavy soil interacted with year dummies. This regression asks *when* the association between riots and soils emerged. Table 9 gives the answer. Col. 1 includes only parish and year fixed effects; col. 2 adds density and the cereal suitability index interacted with year dummies; col. 3 adds all other controls and col. 4 region × year fixed effects. Across specifications, the correlation between wheat suitability and riots is 0 during the first two decades of 1800. In the 1820s the relationship turns negative (though remains insignificant). In the 1830s however, we find a significant correlation between soils and riots. These results suggest that soil characteristics started to predict riots after they became a relevant factor for threshing machine adoption.

Table 7: Sanity check: do light soils predict wheat prevalence?

	log wheat / oat value sold 1820-30				log wheat / oat quantity sold 1820-30			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share of area in parish whose soil is heavy	-0.072	-0.069	-0.067	-0.051	-0.073	-0.071	-0.068	-0.053
	[0.032]	[0.029]	[0.030]	[0.020]	[0.034]	[0.030]	[0.031]	[0.022]
Wheat / oat suitability index		0.074	0.140	0.054		0.053	0.138	0.065
		[0.123]	[0.132]	[0.102]		[0.130]	[0.139]	[0.108]
log 1801 density			-0.007	-0.011			-0.007	-0.011
			[0.009]	[0.009]			[0.009]	[0.009]
Share of agricultural workers in 1801			0.009	-0.010			0.006	-0.012
			[0.012]	[0.010]			[0.013]	[0.011]
log 1801 sex ratio			-0.020	0.009			-0.015	0.013
			[0.014]	[0.012]			[0.015]	[0.012]
log distance to Elham			-0.022	0.025			-0.027	0.022
			[0.025]	[0.017]			[0.026]	[0.019]
log distance to newspaper			0.017	0.016			0.015	0.016
			[0.016]	[0.016]			[0.017]	[0.017]
Region fixed effects (5)	No	No	No	Yes	No	No	No	Yes
R^2	0.022	0.024	0.034	0.101	0.019	0.020	0.028	0.082
Mean dependent variable	1.216	1.216	1.216	1.216	1.153	1.153	1.153	1.153
Observations	9562	9562	9562	9562	9562	9562	9562	9562

Notes: wheat and oat sales and soil characteristics. Col. 1-4: dependent variable is the ratio of the values sold of wheat to oat. Col. 5-8: dependent variable is the ratio of the quantities sold of wheat to oat. Market data is from Brunt and Cannon (2013). Standard errors clustered at the level of the closest market town ($G = 174$) in brackets.

Table 8: Balance table.

	Coefficient of heavy soil:			
	Unconditional	Conditional on cereal suitability	Mean dep. variable	Observations
Poor rates per capita 1800	-0.029	-0.032	0.695	1251
	[0.033]	[0.033]		
log distance to newspaper	-0.016	-0.035	2.951	9674
	[0.020]	[0.021]		
log distance to Elham	0.001	0.108	5.325	9674
	[0.017]	[0.016]		
Share agricultural workers 1801	0.005	0.006	0.386	9674
	[0.007]	[0.008]		
Share trade workers 1801	0.005	0.006	0.117	9674
	[0.004]	[0.004]		
Share other workers 1801	-0.009	-0.012	0.497	9674
	[0.008]	[0.008]		
log 1801 density	-0.044	-0.004	3.646	9674
	[0.028]	[0.029]		
log 1801 sex ratio	0.008	0.004	-0.025	9674
	[0.006]	[0.006]		
Share of land cultivated with cereals 1801	-0.003	-0.001	0.837	3859
	[0.006]	[0.006]		
Riots (1758-1829)	0.002	-0.005	0.067	9674
	[0.014]	[0.014]		

Notes: Balance of heavy soils relative to pre-existing characteristics. Col. 1: coefficients of separate bi-variate regressions. Dependent variable is listed on the left; explanatory variable is share of heavy soil. Col. 2: coefficients of separate regressions. Dependent variable is listed on the left; explanatory variables are share of heavy soil and cereal suitability index. Only the coefficient of share of heavy soil is reported. Robust standard errors in brackets.

Table 9: Correlation between arsons and machine attacks and heavy soil overtime.

	Riots (1780-1832)			
	(1)	(2)	(3)	(4)
Heavy soil × 1800s	0.000	0.000	-0.000	0.000
	[0.001]	[0.002]	[0.001]	[0.001]
Heavy soil × 1810s	0.005	0.002	0.002	0.003
	[0.004]	[0.005]	[0.004]	[0.005]
Heavy soil × 1820s	-0.008	-0.014	-0.014	-0.007
	[0.015]	[0.014]	[0.014]	[0.015]
Heavy soil × 1830-32	-0.125	-0.138	-0.116	-0.115
	[0.019]	[0.019]	[0.018]	[0.019]
Parish & year fixed effects	Yes	Yes	Yes	Yes
log density & cereal suitability index	No	Yes	Yes	Yes
Other controls	No	No	Yes	Yes
Region (5) × year fixed effects	No	No	No	Yes
R^2	0.254	0.256	0.269	0.272
Observations	48636	48636	48636	48636

Notes: Heavy soils and pre-1830 riots. Table reports estimates of Equation (2). Dependent variable is number of arsons or attacks on machines (both agricultural and industrial) between 1758 and 1832. The omitted category is pre-1800: for these years we sum all episodes of unrest and keep a single observation for every parish. Data source is British Library and Findmypast (2016) for pre-1830 events and Holland (2005) for 1830-32: see text for details. Standard errors clustered at parish level in brackets.

B.6 Plausible exogeneity test

To illustrate the robustness of our IV results to limited violations of the exclusion restriction, we perform the test proposed by Conley, Hansen and Rossi (2012). In this exercise, we allow heavy soils to have a direct effect on riots and then re-estimate the IV coefficient of threshing machines. We let the direct effect take any value between 0 and the coefficient of the reduced form: for each of these direct effects, we calculate the union of the 95% confidence intervals of the IV coefficient. In Figure 9 we plot these confidence intervals (y-axis) against the assumed direct effect of the instrument (x-axis). Panel (a) show results for the model with all controls and panel (b) adds 5 region fixed effects. The blue vertical lines flag the value of the reduced form coefficients.

To read the results of this test, we compare the reduced form coefficients to the value of the direct effect where the union of confidence intervals crosses the 0. We find that the direct effect of heavy soils on riots would have to account for between 74% and 78% of the overall reduced form effect before the estimated coefficient becomes insignificant. Because heavy soils are uncorrelated with other determinants of unrest (see Figure 2 in the main text-panel (b) and Table 8) we consider such large direct effects unlikely.

(a) Controls (b) Controls and region fixed effects

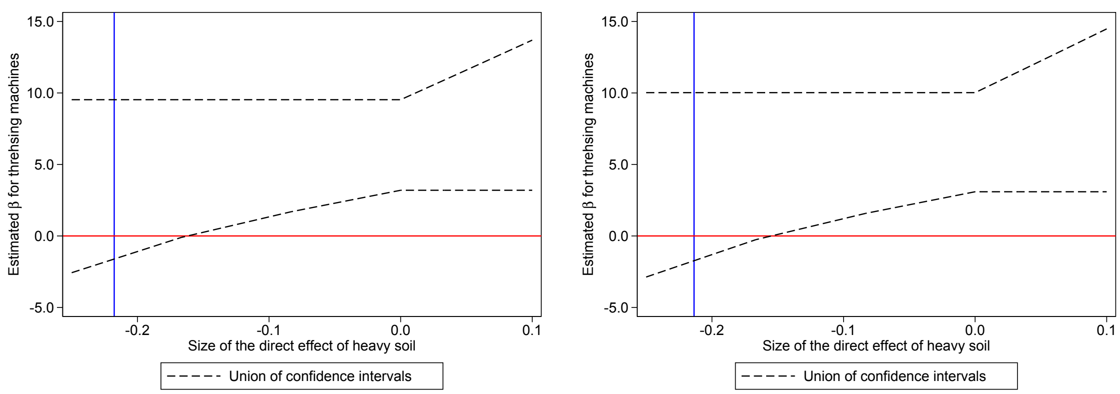


Figure 9: Plausible exogeneity test

Notes. Robustness: effect of violation of exclusion restriction (Conley, Hansen and Rossi, 2012). Union of confidence intervals of the IV estimates (y-axis) when the exclusion restriction is violated (x-axis). Panel (a): regression includes all controls as in col. 9 of Table 2 in the main text. Panel (b): regression includes all controls and 5 region fixed effects as in col. 10 of Table 2 in the main text. Blue vertical lines: point estimate of the reduced form coefficient (cols. 6-7 of Table 2 in the main text).

B.7 Aggravating circumstances: full results

In this section we show the full tables for Figure 3 in the main text. Table 10 reports OLS estimates of Equation (1) in the main text when we split the sample according to the distance to the closest industrial town. Table 11 reports OLS estimates of the same equation when we split the sample according to the level of enclosed commons in 1820.

Table 10: Aggravating circumstances: distance to closest industrial town.

	Distance to industrial town					
	All	Distant	Close	All	Distant	Close
No. threshers	0.389	0.556	0.158	0.353	0.469	0.168
	[0.071]	[0.109]	[0.064]	[0.071]	[0.108]	[0.063]
log 1801 density	0.101	0.143	0.081	0.099	0.162	0.078
	[0.018]	[0.035]	[0.017]	[0.018]	[0.037]	[0.017]
Share of agricultural workers in 1801	-0.065	0.003	-0.141	-0.056	-0.011	-0.124
	[0.044]	[0.067]	[0.054]	[0.043]	[0.066]	[0.054]
log 1801 sex ratio	-0.181	-0.111	-0.187	-0.193	-0.096	-0.206
	[0.042]	[0.069]	[0.055]	[0.043]	[0.071]	[0.056]
log distance to Elham	-0.325	-0.381	-0.289	-0.217	-0.208	-0.335
	[0.029]	[0.046]	[0.037]	[0.045]	[0.057]	[0.080]
log distance to newspaper	0.022	0.026	0.025	0.019	0.056	0.024
	[0.018]	[0.024]	[0.027]	[0.019]	[0.027]	[0.031]
Region fixed effects (5)	No	No	No	Yes	Yes	Yes
R^2	0.057	0.084	0.039	0.064	0.107	0.041
Mean dependent variable	0.308	0.309	0.306	0.308	0.309	0.306
p-value Close = Distant			0.002			0.016
Observations	9674	4785	4889	9674	4785	4889

Notes: Aggravating circumstances: distance to closest industrial town. Dependent variable: number of Swing riots. The table reports results after splitting the sample according to the distance to the closest industrial town. Col. 1 and 4: baseline results (full sample); Col. 2 and 5: results for 4785 parishes above the median parish in terms of distance to industrial town; Col. 3 and 6: results for 4889 parishes below median parish. See Appendix A.2 for details. Robust standard errors in brackets.

Table 11: Aggravating circumstances: enclosures.

	Share land enclosed					
	All	High	Low	All	High	Low
No. threshers	0.462	0.555	0.240	0.398	0.498	0.188
	[0.085]	[0.112]	[0.113]	[0.085]	[0.112]	[0.114]
log 1801 density	0.169	0.122	0.220	0.176	0.147	0.214
	[0.022]	[0.033]	[0.029]	[0.022]	[0.033]	[0.029]
Share of agricultural workers in 1801	0.017	-0.122	0.155	0.009	-0.113	0.133
	[0.057]	[0.082]	[0.079]	[0.056]	[0.081]	[0.076]
log 1801 sex ratio	-0.194	-0.103	-0.240	-0.161	-0.067	-0.209
	[0.051]	[0.075]	[0.069]	[0.053]	[0.079]	[0.069]
log distance to Elham	-0.228	-0.340	-0.219	0.037	-0.013	0.057
	[0.037]	[0.065]	[0.046]	[0.064]	[0.083]	[0.108]
log distance to newspaper	0.049	-0.016	0.105	0.019	-0.057	0.105
	[0.022]	[0.032]	[0.030]	[0.025]	[0.037]	[0.036]
Region fixed effects (5)	No	No	No	Yes	Yes	Yes
R^2	0.040	0.051	0.039	0.052	0.065	0.047
Mean dependent variable	0.345	0.373	0.317	0.345	0.373	0.317
p-value Low = High			0.048			0.053
Observations	6715	3350	3365	6715	3350	3365

Notes: Aggravating circumstances:enclosures and unrest. Dependent variable is number of Swing riots in all columns. The table reports results after splitting the sample according to the 1820 level of enclosures. Columns 1, and 4: baseline results (full sample); columns 2 and 5: results for 3307 parishes above the median parish in terms of enclosures; columns 3 and 6: results for 3408 parishes below median parish. See Appendix A.2 for details. Robust standard errors in brackets.

B.8 Productivity of threshing machines

In this section, we quantify the productivity of threshing machines relative to manual labor. Contemporary observers were aware that threshing machines were markedly more productive (Donaldson, 1794; Batchelor, 1813, p.210).[11] However, there exists no systematic analysis of productivity for the machines in use in 1800, nor are we aware of any attempt to determine the productivity of machines operated with different power sources.

We source information on machine productivity from the county surveys of the *General View of Agriculture*. Sir John Sinclair commissioned the *General Views* as president of the Board of Agriculture in the 1790s, and professional agronomists prepared these documents under the supervision of Arthur Young. Separate volumes cover each county, and the commission surveyed most counties twice: once in 1790s and a second time in the 1810s. We collect all editions covering English counties: a total of 38 separate volumes. All of the *General Views* published in the 1810s, and few of those that appeared in the 1790s contain a chapter on threshing machines. We read these chapters in full, and collect all information that is useful to determine the productivity of these machines. The officials who prepared these chapters toured the English countryside and took detailed notes of every threshing machine they found. A typical entry in this chapter lists owner and location of the machine, as well as material and shape of each different component of the machine. It also reports the mode of operation, the number of men, women and children required to move it and the average quantity of wheat that the machine could thresh in a given amount of time.

We find 121 separate machines in the *General Views*. To calculate productivity we require information on wheat threshed per unit of time, number of people needed to operate the machine, and the main source of power for the machine. Under these constraints, we are able to calculate productivity for 23 horse-powered machines, 3 water-powered machines, and a single machine operated by hand. We show the productivities on Figure 10, where we contrast them with the average productivity of a worker threshing with a flail, as estimated by Clark (1987). Our data is too sparse to provide precise measures of relative productivity. However, the differences are stark. Horse-powered threshing machines may have been 5 times more productive than manual threshing, and water-powered threshing machines more than 10 times more productive. The estimates also suggest that threshing machines operated with human force did not save as much as other types of machines, and did not offer labor savings.[12] Available information also suggest that water-power threshing machines were significantly more productive than horse-powered, possibly by a factor of two.

[11] In the 1794 General View of Banffshire, Donaldson notes: "Threshing-mills have also been introduced of late, and the advantages of them seem to be so well known and established, that there is no doubt of their soon coming into general use" (Donaldson, 1794, p.20).

[12] We only found two hand-powered threshing machines, both in Berkshire (Mavor, 1813, p.133-135). On the first, the informant observes that: "This machine in its present form is evidently more curious than useful. Without horses it is impossible to produce a saving." About the second, he notes: "The only saving Mr. Tull finds in its use is in making reed for thatching." Available information allows to estimate productivity only for one of these two machines.

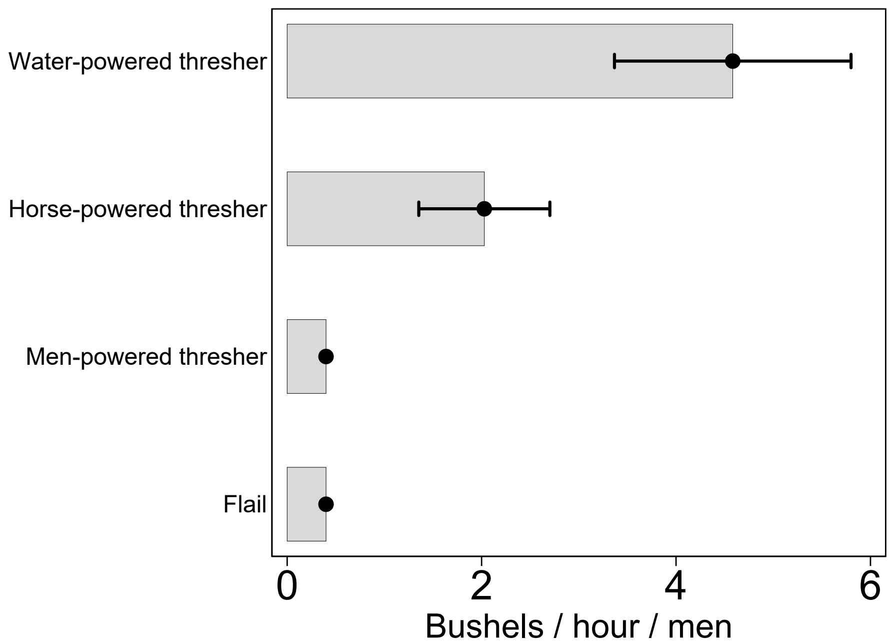


Figure 10: Threshing machine productivity relative to manual threshing.

Notes. Data for threshing machine comes from the county surveys of the *General View of Agriculture.* Sample size is 3 water-powered threshing machines, 24 horse-powered threshing machines and 1 men-powered threshing machine. We only consider wheat threshed, and convert all quantities into bushels. We assume an 8-hours day of work when the surveys report average grains threshed per day. When farmers used women or children to operate these machines we assume that both women and children cost half of what a man does. This is likely to bias productivity downwards, as figures from the Poor Law Report suggest that on average a woman (child) was paid 37.5% (25%) of what men were paid. Average productivity of manual threshers comes from Clark (1987) who uses primary sources to estimate average productivity of English threshers in 1800s.

C Robustness

In this section we show the robustness of our results.

C.1 Alternative specifications and estimation methods

In our baseline results, we control for 1801 Census variables and use OLS to document the effect of threshing machine adoption on riots. This specification has two limitations. First, it does not consider enclosures nor temporary weather shocks as potential causes of Swing. Second, it does not take into account the discrete nature of the dependent variable. We deal with these concerns in Table 12.

In cols. 1-2 of Table 12 we control for 1820 enclosure and abnormal weather conditions in 1830. Point estimates are barely effected and remain highly significant. We do not include these controls in the baseline specification because enclosures are available only for two-thirds of the sample, and historical weather has very high spatial correlation which may bias standard errors downwards.

Col. 3-4 of Table 12 we estimate Poisson regressions. With parish controls (col. 3) or with controls and region fixed effects (col. 4), results remain robust. Finally, in col. 5-8 we look at the extensive margin of riots, and use as a dependent variable a dummy for having at least one incident in 1830-32. Col. 5-6 report results from a linear probability model: in this specification threshers strongly predict riots. In col. 7-8, we use probit estimation to account for the dichotomous nature of the dependent variable. With or without region fixed effects, we always find significant results.

Table 12: Robustness to different estimation methods.

	No. of Swing riots				=1 if Swing riot			
	(1) OLS	(2) OLS	(3) Poisson	(4) Poisson	(5) LPM	(6) LPM	(7) Probit	(8) Probit
No. threshers	0.438	0.396	0.576	0.460	0.108	0.089	0.383	0.295
	[0.084]	[0.085]	[0.060]	[0.058]	[0.016]	[0.016]	[0.050]	[0.049]
log 1801 density	0.182	0.180	0.218	0.192	0.036	0.035	0.144	0.138
	[0.022]	[0.022]	[0.030]	[0.032]	[0.005]	[0.005]	[0.019]	[0.019]
Share of agricultural workers in 1801	0.049	0.029	-0.258	-0.279	-0.047	-0.043	-0.232	-0.242
	[0.058]	[0.056]	[0.172]	[0.163]	[0.014]	[0.014]	[0.070]	[0.071]
log 1801 sex ratio	-0.159	-0.158	-0.529	-0.553	-0.054	-0.059	-0.254	-0.268
	[0.053]	[0.053]	[0.107]	[0.109]	[0.018]	[0.019]	[0.085]	[0.091]
log distance to Elham	0.036	0.146	-0.699	-0.376	-0.113	-0.055	-0.454	-0.197
	[0.078]	[0.079]	[0.037]	[0.062]	[0.007]	[0.010]	[0.024]	[0.034]
log distance to newspaper	0.022	0.016	0.063	0.090	0.002	0.003	-0.004	0.010
	[0.023]	[0.025]	[0.054]	[0.063]	[0.005]	[0.006]	[0.024]	[0.027]
Abnormal precipitation in spring 1830	-0.011	-0.011						
	[0.003]	[0.003]						
Abnormal precipitation in summer 1830	0.002	0.003						
	[0.002]	[0.002]						
Abnormal temperature in fall 1830	-0.817	-0.321						
	[0.856]	[0.965]						
Share of land enclosed (1820)	0.006	0.004						
	[0.004]	[0.004]						
Region fixed effects (5)	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.047	0.054			0.067	0.083		
Mean share	0.345	0.345	0.308	0.308	0.145	0.145	0.145	0.145
Observations	6715	6715	9674	9674	9674	9674	9674	9674

Standard errors in brackets

Notes: Robustness: alternative estimation methods. Col. 1-4: dependent variable is number of Swing riots. Col. 5-8: dependent variable is a dummy for at least one Swing riot. Col. 1-2 and 5-6: OLS regressions. Col. 3-4: Poisson regression. Col. 7-8: Probit regression. Robust standard errors in brackets.

C.2 Spatial autocorrelation

In Section III, we base inference on conventional robust standard errors that do not account for spatial correlation in the explanatory variable. However, the geographic distribution of machines and riots, as well as soil suitability, suggest some spatial correlation. Here, we show that accounting for spatial correlation has no effect on the significance of our results.

We control for spatial correlation in two ways. First, we compute standard errors with the formula proposed by Conley (1999).[13] We experiment with three different cutoffs: 20, 50 and 100 km. Second, we estimate standard errors in a non-parametric way, and estimate cluster-robust standard errors. We consider 3 different levels of clustering: closest market town, closest city that publishes a newspaper and county. This creates respectively 174, 60 and 54 clusters.

Table 13 reports the results. OLS results remain strong and significant when we introduce Conley standard errors or clustering. Similarly, first stage, reduced form and IV results survive when we account for spatial correlation: spatially robust standard errors tend to be larger than conventional robust standard errors, but all estimates remain significant at the 2.8 percent level or better. All in all, these results suggest that spatial autocorrelation is not responsible for the significance of our findings.

[13]We estimate these standard errors with the code `acreg` of Colella et al. (2019).

Table 13: Robustness: standard errors robust to spatial autocorrelation.

No. of	Swing riots		threshers		Swing riots			
	(1) OLS	(2) OLS	(3) FS	(4) FS	(5) RF	(6) RF	(7) 2SLS	(8) 2SLS
No. threshers	0.389	0.353					6.361	6.557
Huber-Ecker-White robust s.e.	[0.071]	[0.071]					[1.616]	[1.768]
Conley (1999) s.e.: cutoff = 20km	[0.075]	[0.074]					[2.265]	[2.403]
Conley (1999) s.e.: cutoff = 50km	[0.095]	[0.088]					[2.924]	[2.968]
Conley (1999) s.e.: cutoff = 100km	[0.110]	[0.094]					[3.066]	[3.231]
Clustered s.e.: closest market town (174)	[0.085]	[0.075]					[2.527]	[2.660]
Clustered s.e.: closest town with newspaper (60)	[0.083]	[0.081]					[2.490]	[2.747]
Clustered s.e.: county (56)	[0.096]	[0.090]					[2.878]	[2.735]
Share of area in parish whose soil is heavy			-0.034	-0.033	-0.218	-0.214		
Huber-Ecker-White robust s.e.			[0.008]	[0.008]	[0.026]	[0.027]		
Conley (1999) s.e.: cutoff = 20km			[0.011]	[0.011]	[0.039]	[0.037]		
Conley (1999) s.e.: cutoff = 50km			[0.015]	[0.013]	[0.049]	[0.042]		
Conley (1999) s.e.: cutoff = 100km			[0.017]	[0.014]	[0.062]	[0.050]		
Clustered s.e.: closest market town (174)			[0.013]	[0.012]	[0.046]	[0.041]		
Clustered s.e.: closest town with newspaper (60)			[0.014]	[0.013]	[0.050]	[0.041]		
Clustered s.e.: county (56)			[0.016]	[0.013]	[0.052]	[0.042]		
log 1801 density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
parish characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region fixed effects (5)	No	Yes	No	Yes	No	Yes	No	Yes
Mean dependent variable	0.308	0.308	0.308	0.062	0.062	0.308	0.308	0.308
Observations	9671	9671	9671	9671	9671	9671	9671	9671

Notes: Robustness: correction for spatial correlation. Point estimates from Table 2 in the main text. Standard errors underneath estimates. Row 1: heteroschedastic-robust standard errors. Rows 2-4: standard error corrected with the formula of Conley (1999). Cutoff is 20 (row 2) 50 (row 3) and 100 Km (row 4). Rows 5-7: cluster-robust standard errors. Clustering at: closest market town (row 5), closest city with a newspaper (row 6) and county (row 7). Col. 1-3: OLS estimates of Equation (1) in the main text. Col. 4-5: first stage estimates of Equation (3) in the main text. Col. 6-7: reduced form estimates of Equation (4) in the main text. Col. 8-10: IV estimates of Equation (1) in the main text, using share of heavy soil as instrument.

C.3 County fixed effects and nearest neighbor matching

All our results are robust to introducing 54 county fixed effects or estimating treatment effects based on nearest neighbor matching. The robustness of our results to the inclusion of county fixed effects reinforces our conclusions since counties are small, relatively homogeneous geographical units. Because we find that threshers cause more riots even within these small areas, we conclude that unobservables are unlikely to drive our results.

Table 14 presents results with county fixed effects. The first 4 columns report the basic correlation between riots and threshing machines. Whether we estimate OLS or a Poisson regression (col. 1-2), or we take a dummy for the presence of Swing and estimate a linear probability model or a Probit (col. 3-4), we always find strong correlations between riots and threshers. We report first stage, reduced form and IV in col. 5-7 of the same table: also these results remain strong after the inclusion of county fixed effect.

Table 15, panel (a) estimates the average treatment effect of threshers on riots with nearest neighbor matching. Treatment is the presence of at least one thresher: we match each treated parish based on latitude and longitude. We report results when we find a single match (col. 1 and 4), 3 (col. 2 and 5) or 5 matches (col. 3 and 6). In col. 4-6, we also force matched parishes to lie within the same county. In all specifications we find that threshers are a significant predictor of unrest.

Table 15, panel (b) uses nearest neighbor matching with heavy soil as treatment. Treated parishes are all those in the top quartile in the distribution of heavy soils. We always match on latitude and longitude, and col. 4-6, we also require matched parishes to lie within the same county. Results confirm that parishes with heavy soils have significantly fewer riots.

Counties constitute small geographical units with relatively forms of agricultural cultivation. Moreover, close parishes share many unobserved characteristics that may bias our estimates. Because we find that threshers cause more riots even within these fine geographical units, we conclude that unobservables are unlikely to drive our results.

Table 14: Robustness: county fixed effects.

	Swing riots		=1 if Swing		Threshers	Swing riots	
	(1) OLS	(2) Poisson	(3) LPM	(4) Probit	(5) FS	(6) RF	(7) 2SLS
No. of threshers	0.324	0.374	0.080	0.273			4.423
	[0.069]	[0.063]	[0.015]	[0.051]			[1.703]
Share of area in parish whose soil is heavy					-0.028	-0.122	
					[0.009]	[0.030]	
Cereal suitability index					-0.139	-0.365	0.251
					[0.045]	[0.158]	[0.350]
log 1801 density	0.147	0.350	0.051	0.232	0.020	0.155	0.065
	[0.021]	[0.036]	[0.005]	[0.023]	[0.004]	[0.022]	[0.042]
Share of agricultural workers in 1801	-0.082	-0.346	-0.049	-0.267	-0.033	-0.090	0.054
	[0.044]	[0.160]	[0.014]	[0.075]	[0.011]	[0.045]	[0.083]
log 1801 sex ratio	-0.143	-0.434	-0.044	-0.236	-0.003	-0.137	-0.124
	[0.044]	[0.118]	[0.019]	[0.095]	[0.014]	[0.045]	[0.070]
log distance to Elham	-0.067	-0.141	-0.035	-0.127	-0.006	-0.062	-0.037
	[0.114]	[0.114]	[0.027]	[0.080]	[0.013]	[0.114]	[0.126]
log distance to newspaper	0.047	0.159	0.006	0.030	0.008	0.051	0.014
	[0.026]	[0.064]	[0.007]	[0.030]	[0.007]	[0.026]	[0.041]
County fixed effects (54)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.101		0.117		0.052	0.096	
Mean DV	0.308	0.308	0.145	0.154	0.062	0.308	0.308
F-test excluded instrument					9.3		
Rubin-Anderson test (p)							0.000
Observations	9674	9674	9674	9100	9674	9674	9674

Notes: Robustness: County fixed effect. Col. 1-2 and 6-7: dependent variable is number of Swing riots. Col. 3-4: dependent variable is a dummy for at least one Swing riot. Col. 5: dependent variable is number of threshers. Col. 1 and 3: OLS regressions. Col. 2: Poisson regression. Col. 4: Probit regression. Col. 5: first stage estimates of Equation (3) in the main text. Col. 6: reduced form estimates of Equation (4) in the main text. Col. 7: IV estimates of Equation (1) in the main text, using share of heavy soil as instrument. Robust standard errors in brackets.

Table 15: Nearest neighbor matching.

Panel (a): treatment = thresher	No. of Swing riots					
ATT	0.401	0.429	0.385	0.419	0.414	0.376
	[0.084]	[0.069]	[0.068]	[0.081]	[0.069]	[0.068]
Panel (b): treatment = heavy soil	**No. of Swing riots**					
ATT	-0.102	-0.079	-0.081	-0.111	-0.090	-0.095
	[0.039]	[0.028]	[0.026]	[0.041]	[0.029]	[0.027]
Number of matches	1	3	5	1	3	5
Matched within county? (54)	No	No	No	Yes	Yes	Yes
Observations	9674	9674	9674	9674	9674	9674

Notes: Robustness: nearest neighbor matching. Dependent variable is number of Swing riots. Panel (a): treated parishes have at least one thresher. Panel (b): treated parishes have share of heavy soil in the top quartile of the distribution. Col. 1-3: matching on latitude and longitude. Col. 4-6: matching on latitude, longitude and county (exact). Number of matches: 1 (col. 1 and 4), 3 (col. 2 and 5) and 5 (col. 3 and 6).

C.4 Sample restrictions

Part of the information we use to track machine adoption comes from historical newspapers. These newspapers come from 60 towns and cities, and they were more likely to contain advertisements for farm sales near the place of publication. Similarly, part of the riot data come from newspapers, and may be more likely to report unrest in the same surrounding villages. To control for this possible confounding mechanism, we include the distance to the closest newspaper in all our regressions. Additionally, here we show that all our results survive if we restrict the sample to parishes within 30 kilometers from the closest newspaper. We report our estimates in Table 16. This table shows estimates for OLS (columns 1-3), first stage (columns 4-5), reduced form (columns 6-7) and IV (columns 8-10). These estimates confirm that none of our results is driven by the potentially uneven coverage of English parishes offered by 1800 newspapers.

A second concern involves the timing of the riots. While Holland (2005) records episodes that happened until the end of 1832, most of the protests took place during the winter of 1830-31, and the most violent part of the revolt was over by the spring of 1831. Including later unrest episodes may introduce noise. To address this concern, we replicate the whole analysis after excluding all episodes that happened after April 1831.[14] Results in Table 17 confirm that the specific definition of riots is not driving our results.

A third concern has to do with the urban nature of some of the parishes in our sample. Around 3.4 percent of the English parishes have a share of workers employed in agriculture below 10 percent. These places were mostly urban, and in 1801 they were home to about 40 percent of the English population. Because threshing machines affected agricultural workers and Swing was mostly a rural uprising, it is useful to evaluate whether our results hold when we remove urban parishes from the sample. Table 18 reports results for parishes with agricultural share greater than 10 percent: coefficients are similar to our baseline estimates.

A final concern with our results is that they may reflect the contrast between English and Welsh parishes. English parishes specialized in cereal production and bore the brunt of the Swing riots. In contrast, pastoral agriculture was more common in Wales, and the riots left this region almost untouched. We already showed that all results are robust to including 54 county fixed effects. Table 19 shows that excluding the 949 Welsh parishes from our regressions further strengthens our results.

[14] This excludes 619 episodes, leaving 2421 riots.

Table 16: Robustness: sample excludes parishes farther than 30 Km from a town with a newspaper.

	Number of Swing riots		Number of threshers		Number of Swing riots			
	(1) OLS	(2) OLS	(3) FS	(4) FS	(5) RF	(6) RF	(7) 2SLS	(8) 2SLS
Threshers	0.399	0.367					8.631	9.509
	[0.083]	[0.083]					[2.919]	[3.672]
Share of area in parish whose soil is heavy			-0.029	-0.025	-0.249	-0.241		
			[0.009]	[0.010]	[0.032]	[0.032]		
log 1801 density	0.101	0.098	0.012	0.010	0.103	0.097	0.002	-0.001
	[0.021]	[0.022]	[0.004]	[0.004]	[0.022]	[0.022]	[0.047]	[0.052]
Cereal suitability index			0.057	0.101	0.203	0.481	-0.289	-0.483
			[0.043]	[0.044]	[0.131]	[0.137]	[0.450]	[0.621]
Share of agricultural workers in 1801	-0.116	-0.109	-0.026	-0.033	-0.127	-0.120	0.095	0.196
	[0.050]	[0.049]	[0.012]	[0.012]	[0.050]	[0.049]	[0.131]	[0.167]
log 1801 sex ratio	-0.215	-0.226	-0.036	-0.022	-0.229	-0.236	0.082	-0.031
	[0.050]	[0.051]	[0.017]	[0.017]	[0.052]	[0.053]	[0.176]	[0.177]
log distance to Elham	-0.311	-0.228	-0.004	0.059	-0.326	-0.240	-0.291	-0.802
	[0.032]	[0.050]	[0.004]	[0.008]	[0.036]	[0.052]	[0.053]	[0.216]
log distance to newspaper	0.039	0.030	-0.001	-0.004	0.048	0.041	0.057	0.076
	[0.029]	[0.030]	[0.008]	[0.008]	[0.030]	[0.030]	[0.073]	[0.080]
Region fixed effects (5)	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.053	0.060	0.006	0.030	0.050	0.059	-4.313	-5.196
Mean dependent variable	0.337	0.337	0.063	0.063	0.337	0.337	0.337	0.337
F-test excluded instrument			9.3	7.1				
Rubin-Anderson test (p)							0.000	0.000
Observations	7396	7396	7396	7396	7396	7396	7396	7396

Notes: Robustness: sample excludes all parishes further than 30 Km from a city that publishes at least 1 newspaper. Col. 1-3: OLS estimates of Equation (1) in the main text. Col. 4-5: first stage estimates of Equation (3) in the main text. Col. 6-7: reduced form estimates of Equation (4) in the main text. Col. 8-10: IV estimates of Equation (1) in the main text, using share of heavy soil as instrument. Robust standard errors in brackets.

Table 17: Robustness: sample excludes riots after april 1831.

	Number of Swing riots		Number of threshers		Number of Swing riots			
	(1) OLS	(2) OLS	(3) FS	(4) FS	(5) RF	(6) RF	(7) 2SLS	(8) 2SLS
Threshers	0.313	0.279					4.922	5.008
	[0.063]	[0.063]					[1.294]	[1.402]
Share of area in parish whose soil is heavy			-0.034	-0.033	-0.168	-0.163		
			[0.008]	[0.008]	[0.024]	[0.024]		
log 1801 density	0.075	0.074	0.015	0.013	0.076	0.073	0.004	0.007
	[0.015]	[0.015]	[0.004]	[0.004]	[0.015]	[0.015]	[0.027]	[0.027]
Cereal suitability index			0.050	0.044	0.193	0.300	-0.052	0.079
			[0.032]	[0.032]	[0.082]	[0.085]	[0.192]	[0.193]
Share of agricultural workers in 1801	-0.018	-0.017	-0.015	-0.022	-0.026	-0.024	0.049	0.086
	[0.040]	[0.039]	[0.010]	[0.010]	[0.040]	[0.039]	[0.065]	[0.071]
log 1801 sex ratio	-0.174	-0.174	-0.024	-0.011	-0.180	-0.184	-0.062	-0.128
	[0.038]	[0.039]	[0.014]	[0.014]	[0.039]	[0.040]	[0.080]	[0.077]
log distance to Elham	-0.285	-0.181	-0.006	0.070	-0.299	-0.186	-0.268	-0.536
	[0.025]	[0.040]	[0.004]	[0.007]	[0.027]	[0.041]	[0.033]	[0.109]
log distance to newspaper	0.017	0.013	-0.000	0.000	0.018	0.016	0.020	0.016
	[0.015]	[0.017]	[0.005]	[0.006]	[0.015]	[0.017]	[0.029]	[0.032]
Region fixed effects (5)	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.049	0.055	0.006	0.032	0.046	0.053	-1.753	-1.793
Mean dependent variable	0.244	0.244	0.062	0.062	0.244	0.244	0.244	0.244
F-test excluded instrument			17.7	15.9				
Rubin-Anderson test (p)							0.000	0.000
Observations	9674	9674	9674	9674	9674	9674	9674	9674

Notes: Robustness: only riots between August 1830 and April 1831. Col. 1-3: OLS estimates of Equation (1) in the main text. Col. 4-5: first stage estimates of Equation (3) in the main text. Col. 6-7: reduced form estimates of Equation (4) in the main text. Col. 8-10: IV estimates of Equation (1) in the main text, using share of heavy soil as instrument. Robust standard errors in brackets.

Table 18: Robustness: sample excludes urban parishes.

	Number of Swing riots		Number of threshers		Number of Swing riots			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	FS	FS	RF	RF	2SLS	2SLS
Threshers	0.375	0.327					6.547	6.652
	[0.076]	[0.076]					[1.884]	[2.031]
Share of area in parish whose soil is heavy			-0.029	-0.028	-0.191	-0.186		
			[0.008]	[0.008]	[0.026]	[0.026]		
log 1801 density	0.120	0.121	0.015	0.013	0.126	0.125	0.025	0.035
	[0.017]	[0.017]	[0.005]	[0.005]	[0.017]	[0.017]	[0.043]	[0.042]
Cereal suitability index			0.055	0.053	0.077	0.197	-0.284	-0.155
			[0.032]	[0.032]	[0.086]	[0.088]	[0.252]	[0.255]
Share of agricultural workers in 1801	0.028	0.028	-0.009	-0.015	0.023	0.025	0.085	0.127
	[0.044]	[0.044]	[0.011]	[0.011]	[0.045]	[0.044]	[0.082]	[0.088]
log 1801 sex ratio	-0.099	-0.098	-0.007	0.003	-0.099	-0.103	-0.050	-0.122
	[0.042]	[0.042]	[0.015]	[0.016]	[0.042]	[0.043]	[0.105]	[0.107]
log distance to Elham	-0.309	-0.168	-0.005	0.063	-0.315	-0.164	-0.282	-0.583
	[0.026]	[0.039]	[0.004]	[0.007]	[0.028]	[0.041]	[0.037]	[0.136]
log distance to newspaper	0.029	0.023	0.000	0.001	0.029	0.025	0.027	0.020
	[0.014]	[0.016]	[0.005]	[0.005]	[0.014]	[0.016]	[0.032]	[0.037]
Region fixed effects (5)	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.057	0.067	0.004	0.027	0.052	0.064	-2.827	-2.890
Mean dependent variable	0.272	0.272	0.058	0.058	0.272	0.272	0.272	0.272
F-test excluded instrument			13.9	12.4				
Rubin-Anderson test (p)							0.000	0.000
Observations	8747	8747	8747	8747	8747	8747	8747	8747

Notes: Robustness: sample excludes all parishes with less than 10% of agricultural workers in 1801. Col. 1-3: OLS estimates of Equation (1) in the main text. Col. 4-5: first stage estimates of Equation (3) in the main text. Col. 6-7: reduced form estimates of Equation (4) in the main text. Col. 8-10: IV estimates of Equation (1) in the main text, using share of heavy soil as instrument. Robust standard errors in brackets.

Table 19: Robustness: sample excludes Welsh parishes.

	Number of Swing riots		Number of threshers		Number of Swing riots			
	(1) OLS	(2) OLS	(3) FS	(4) FS	(5) RF	(6) RF	(7) 2SLS	(8) 2SLS
Threshers	0.397	0.366					6.571	7.781
	[0.073]	[0.073]					[1.608]	[2.474]
Share of area in parish whose soil is heavy			-0.040	-0.030	-0.260	-0.231		
			[0.009]	[0.009]	[0.030]	[0.030]		
log 1801 density	0.107	0.105	0.016	0.015	0.108	0.105	0.005	-0.009
	[0.019]	[0.020]	[0.004]	[0.004]	[0.019]	[0.020]	[0.037]	[0.047]
Cereal suitability index			0.070	0.064	0.209	0.440	-0.248	-0.057
			[0.040]	[0.040]	[0.119]	[0.128]	[0.319]	[0.390]
Share of agricultural workers in 1801	-0.061	-0.059	-0.010	-0.021	-0.068	-0.068	-0.003	0.094
	[0.050]	[0.049]	[0.012]	[0.012]	[0.051]	[0.049]	[0.089]	[0.111]
log 1801 sex ratio	-0.214	-0.220	-0.036	-0.019	-0.221	-0.230	0.013	-0.083
	[0.048]	[0.049]	[0.016]	[0.016]	[0.049]	[0.050]	[0.121]	[0.135]
log distance to Elham	-0.312	-0.219	-0.001	0.070	-0.320	-0.233	-0.315	-0.776
	[0.030]	[0.046]	[0.004]	[0.007]	[0.033]	[0.049]	[0.041]	[0.179]
log distance to newspaper	0.041	0.027	0.008	0.004	0.045	0.027	-0.005	-0.004
	[0.022]	[0.023]	[0.006]	[0.007]	[0.022]	[0.023]	[0.049]	[0.058]
Region fixed effects (5)	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.051	0.057	0.007	0.031	0.048	0.055	-2.463	-3.477
Mean dependent variable	0.344	0.344	0.067	0.067	0.344	0.344	0.344	0.344
F-test excluded instrument			18.9	10.9				
Rubin-Anderson test (p)							0.000	0.000
Observations	8591	8591	8591	8591	8591	8591	8591	8591

Notes: Robustness: sample excludes all Welsh parishes. Col. 1-3: OLS estimates of Equation (1) in the main text. Col. 4-5: first stage estimates of Equation (3) in the main text. Col. 6-7: reduced form estimates of Equation (4) in the main text. Col. 8-10: IV estimates of Equation (1) in the main text, using share of heavy soil as instrument. Robust standard errors in brackets.

References

Alexander, Lisa V., and Philip D. Jones. 2000. "Updated precipitation series for the U.K. and discussion of recent extremes." *Atmospheric Science Letters*, 1(2): 142–150.

Batchelor, Thomas. 1813. *General View of the Agriculture of the County of Bedford: Drawn Up for the Consideration of the Board of Agriculture and Internal Improvement.* Sherwood, Neely and Jones.

British Geological Survey. 2009. "Soil Parent Material model - 1 Kilometer resolution [database]."

British Library and Findmypast. 2016. "The British Newspaper Archive."

Brunt, Liam, and Edmund Cannon. 2013. "The Truth, the Whole Truth, and Nothing but the Truth: The English Corn Returns as a Data Source in Economic History, 1770–1914." *European Review of Economic History*, 17(3): 318–339.

Checkland, Sidney G. 1974. *The Poor Law Report of 1834.* London: Penguin Books.

Clark, Gregory. 1987. "Productivity Growth Without Technical Change in European Agriculture Before 1850." *Journal of Economic History*, 47(2): 419–432.

Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli, and Mathias Thoenig. 2019. "Inference with Arbitrary Clustering." *Working Paper.*

Conley, T.G. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics*, 92(1): 1–45.

Conley, Timothy G., Christian B. Hansen, and Peter E. Rossi. 2012. "Plausibly Exogenous." *Review of Economics and Statistics*, 94(1): 260–272.

Donaldson, James. 1794. *General View of the Agriculture of the County of Banff.* Thomas Ruddiman.

Gonner, E. C. K. 1912. *Common Land and Inclosure.* MacMilland and Co.

Griffin, Carl J. 2012. *The Rural War: Captain Swing and the Politics of Protest.* Manchester: Manchester Univ. Press.

Hijmans, Robert J, Susan E Cameron, Juan L Parra, Peter G Jones, and Andy Jarvis. 2005*a*. "Very high resolution interpolated climate surfaces for global land areas." *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 25(15): 1965–1978.

Hijmans, Robert J, Susan E Cameron, Juan L Parra, Peter G Jones, and Andy Jarvis. 2005*b*. "WorldClim 1.4 [database]."

Hobsbawm, Eric J., and George F. E. Rudé. 1969. *Captain Swing.* London: Verso.

Holland, Michael. 2005. *Swing Unmasked: The Agricultural Riots of 1830 to 1832 and Their Wider Implications.* Milton Keynes: Family and Community Historical Research Society Publishing.

Kain, Roger JP, and Richard R. Oliver. 2001. *The Historic Parishes of England and Wales: An Electronic Map of Boundaries Before 1850 with a Gazetteer and Metadata.* Twayne Publishers.

Luterbacher, J., D. Dietrich, E. Xoplaki, M. Grosjean, and H. Wanner. 2004. "European Seasonal and Annual Temperature Variability, Trends, and Extremes Since 1500." *Science*, 303(5663): 1499–1503.

Mavor, William Fordyce. 1813. *General View of the Agriculture of Berkshire.* Sheerwood, Neely and Jones.

Met Office Hadley Centre. 2001. "Met Office Hadley Centre observations datasets: HadUKP [database]."

Moulton, Brent R. 1990. "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units." *Review of Economics and Statistics*, 72(2): 334–338.

Pauling, Andreas, Jürg Luterbacher, Carlo Casty, and Heinz Wanner. 2006. "Five Hundred Years of Gridded High-Resolution Precipitation Reconstructions Over Europe and the Connection to Large-Scale Circulation." *Climate Dynamics*, 26(4): 387–405.

Southall, H.R., and N. Burton. 2004. *GIS of the Ancient Parishes of England and Wales, 1500-1850.* Colchester, Essex: UK Data Archive.

Southall, H.R., P. Ell, D. A. Gatley, and I. Gregory. 2004. *Great Britain Historical Database: Census Data: Parish-Level Population Statistics, 1801-1951.* Colchester, Essex: UK Data Archive.

Turner, M. E. 2005. *Crop Returns for England, 1801.* UK Data Service.

Wigton-Jones, Evan. 2019. "Legacies of Inequality: The Case of Brazil." *Working Paper.*

Xoplaki, E., J. Luterbacher, H. Paeth, D. Dietrich, N. Steiner, M. Grosjean, and H. Wanner. 2005. "European Spring and Autumn Temperature Variability and Change of Extremes Over the Last Half Millennium." *Geophysical Research Letters*, 32(15).