



QUCHEH WORKING PAPER SERIES  
<http://www.quceh.org.uk/working-papers>

POPULATION AND POVERTY IN IRELAND ON THE  
EVE OF THE GREAT FAMINE

Alan Fernihough (Queen's University Belfast)  
Cormac Ó Gráda (University College Dublin)

Working Paper 2018-13

QUEEN'S UNIVERSITY CENTRE FOR ECONOMIC HISTORY  
Queen's University Belfast  
185 Stranmillis Road  
Belfast BT9 5EE

December 2018

# Population and Poverty in Ireland on the Eve of the Great Famine\*

Alan Fernihough<sup>†</sup>

Cormac Ó Gráda<sup>‡</sup>

## Abstract

The link between demographic pressure and economic conditions in pre-Famine Ireland has long interested economists. This paper re-visits the topic, harnessing the highly disaggregated parish-level data from the 1841 Census of Ireland. Using population per value adjusted acre as a measure of population pressure, our results indicate that on the eve of the Great Famine of 1846–50, population pressure was positively associated with both illiteracy rates and the prevalence of poor quality housing. But while our analysis shows that population pressure was one of the primary factors underpinning pre-Famine poverty, it also highlights the importance of geography and human agency. A counterfactual computation indicates that had Ireland’s population stayed at its 1800 level, this would have led to only modest improvements in literacy and housing.

**JEL-Classification:** N33, B30, J11,

**Keywords:** Famine, Malthus, Population, Ireland

---

\*We are grateful to Morgan Kelly, Bruce Campbell, Liam Kennedy, Alan de Bromhead, Peter Solar, Áine Doran, and seminar and conference participants at the Cambridge Group for the History of Population and Social Structure seminar series and the 2018 Irish Economic Association annual conference, for useful comments, advice and help with data. Funding from ESRC grant ES/N017323/1 is gratefully acknowledged.

<sup>†</sup>Queen’s Management School, Queen’s University Belfast, Belfast BT9 5EE, United Kingdom. Email: a.fernihough@qub.ac.uk.

<sup>‡</sup>University College Dublin.

# 1 Introduction

Despite two brutal World Wars and countless incidents of civil unrest and political turmoil, the populations of Western Europe's nation states have grown rapidly since the mid-19th century. Ireland is very much the outlier in this cohort. Even today, the population of the island of Ireland is approximately 6.7 million, some 1.8 million individuals shy of the pre-Famine population of 8.5 million. Research examining the link between economic growth and demography emphasises how economic improvements occurred in tandem with rapid population growth after 1850 (Galor and Weil, 2000; Hansen and Prescott, 2000; Clark, 2007). This paper addresses why Ireland's aberrant population trajectory contradicted this conventional pathway.

It is widely accepted that the Great Famine of the mid-19th century not only explains why the population of Ireland fell so dramatically during 1846–50 but also why it continued to fall for the proceeding century (O'Rourke, 1995). Although many regions across Western Europe were affected by the potato blight of the mid-19th century (Ó Gráda et al., 2007), none experienced anything like the devastation that befell Ireland. What made Ireland so vulnerable was poverty: a large portion of the island's inhabitants lived at a subsistence level and lacked any wealth or tangible financial assets. Economic backwardness and the failure of the population to recover in the post-famine period suggest that Ireland's pre-Famine malaise was, at least in part, caused by overpopulation and thus that Ireland would have been in a less precarious position in 1845 had the population been lower.

The evidence for Irish poverty on the eve of the Great Famine is both quantitative and qualitative. Data on literacy and housing quality in the 1841 population census point to Ireland's laggard status in the United Kingdom; wage data imply that Irish agricultural workers were paid about half as much as their English peers; and Mokyr (1985, p.11) reckoned that income per capita in Britain was over 2.5 times that of Ireland on the eve of the famine. The near-complete dependence of the bottom third or so of the population on a single crop—the potato—corroborates, although the potato's high nutritional content meant that the Irish poor were healthier and better fed than their low incomes would predict (Ó Gráda, 1994, pp.14–5, 86–90). Impressions reported by contemporary travellers—‘such various gradations of misery as he could not have supposed possible to exist, even among the most barbarous nations’ (Wakefield, 1812, II, pp.838); ‘shocking

for humanity to contemplate, and beyond the efforts of private beneficence to relieve’ (Inglis, 1835, II, pp.322–4); ‘everywhere people are in rags, and wearing traces of hunger and disease in their pale faces’ (Eötvös, 2015, pp.5 (1837))—broadly support the quantitative evidence. There is much more in that vein (compare de Beaumont (1839), Kohl (1844), and Thackeray (1911, (1837))). True, such impressions should be treated with some caution insofar as they tended to refer to short trips made during the summer months when food and employment were scarcest, and the tendency of travellers to stick to the main roads made them easy targets for beggars. But the sheer volume and unanimity of their bleak commentary is striking. Irish poverty on the eve of the Great Famine, in both relative and absolute terms, is not in doubt.

The conventional view amongst the British political establishment and commentators was that the proximate cause of Irish poverty was overpopulation (Gray, 2006). While the establishment of a ‘Royal Commission of Inquiry into the Conditions of the Poorer Classes in Ireland’ in 1833 represented a partial acknowledgement by the Westminster government that Irish poverty needed to be addressed, fears that a Poor Law along English lines might actually exacerbate Irish destitution were common. Still, in its broad outlines the Irish Poor Law Act of 1838 mirrored the English legislation of 1834 (see Black (1960, pp. 86–112) and Gray (2009)).

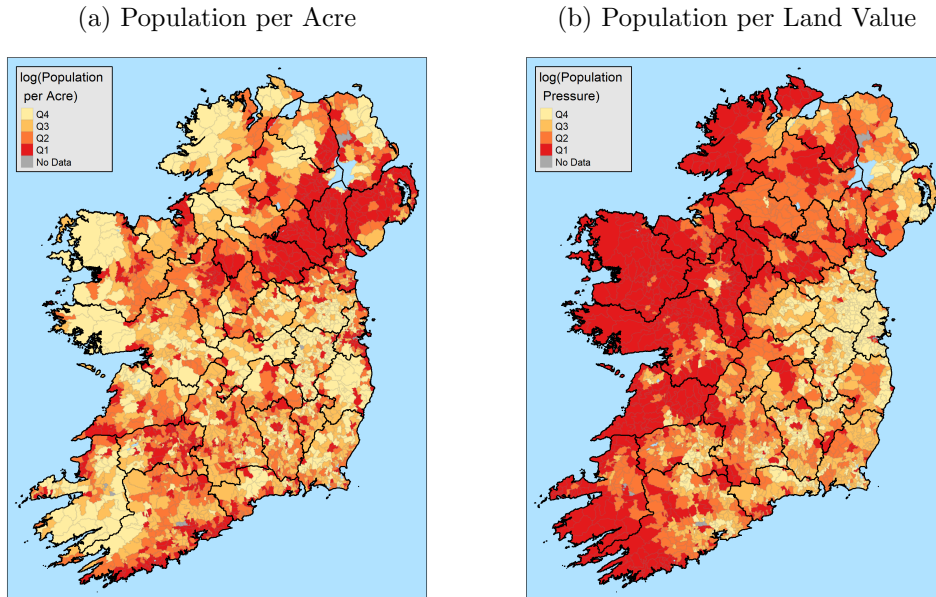
The historical consensus that overpopulation was the root cause of pre-Famine Irish poverty was endorsed in Connell’s classic account (1950), which saw the introduction of the potato as a major catalyst for 18th and early 19th century Irish population growth. That growth, however, did not grind to a halt as the potato’s value as a means of enriching the pre-existing diet was exhausted; instead, underpinned by early marriage and high marital fertility, it culminated in a dependence on the potato unmatched anywhere else, with ultimately apocalyptic consequences. This view was famously challenged by Mokyr (1985) (see also Mokyr (1980)), who made the very salient point that previous analyses connecting Irish overpopulation and poverty lacked hard empirical evidence. Mokyr’s own analysis combined newly-constructed county-level data (so  $n = 32$ ) in a linear regression model in which income per capita (or a suitable proxy measure) was modelled as a function of population density (rural population per cultivated acre) and other measures.<sup>1</sup> Mokyr

---

<sup>1</sup>Please see Figure 9 in Appendix A for a map of Ireland’s administrative boundaries.

thus invoked cross-section data as proxies for points in time. Surprisingly, his results failed to support the overpopulation hypothesis: Irish counties with greater population pressure did not have lower income per capita. Mokyr’s result remained robust across a number of different modelling specifications, which led him to conclude that economic historians should focus more on alternative reasons for pre-Famine Ireland’s endemic poverty.

Figure 1: Population Density Map

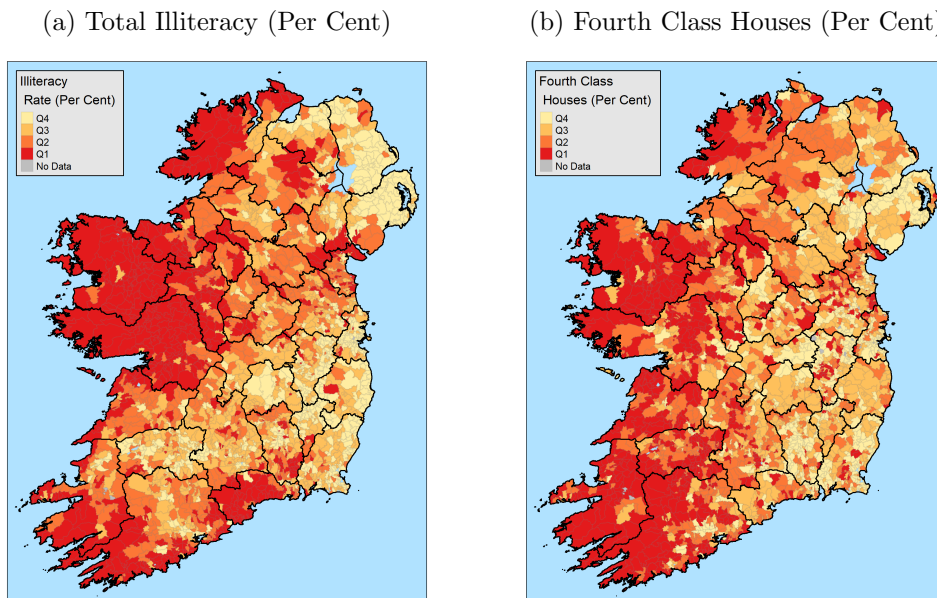


There can be little doubt but that there was more than one reason for Irish poverty, but we believe that, more than three decades on, Mokyr’s approach to testing Malthus can be improved upon in several ways. In attempting to do this, like Mokyr we rely on the returns from the 1841 Irish population census (BPP, 1843). However, our unit of analysis is the civil parish instead of the county, which increases sample size from 32 to 2,387. The much higher resolution parish-level returns improve the accuracy of our estimates and allow for greater nuance between different specifications and model subsamples (compare Brown and Guinnane (2007)). Another important difference, anticipated in McGregor (1989), is that we use a measure of population pressure that relies on a better measure of land quality than those proposed by Mokyr (1985, pp. 47–8). Here we adjust for quality by using population divided by Poor Law valuation. The Poor Law valuation was a measure of the fertility of land and buildings, and formed the basis for the tax payable toward the support of the poor under the Irish Poor Law of 1838. Essentially a land value—it was assessed with an eye to what might be considered a fair rent—it is a more accurate representation

of population pressure than e.g. mean height above sea level or the proportion of land under cultivation. The maps in Figure 1 highlight the importance of adjusting for land quality. In Map (a) much of the west of Ireland seems ‘under’-populated, but this outcome is reversed in Map (b) which adjusts for quality using the Poor Law valuation.

Measuring poverty is fraught with difficulty even in modern developed economies. This difficulty is compounded by data constraints in historical samples. We choose two measures of poverty in our analysis: illiteracy and the prevalence of low-quality (fourth class) housing. The illiteracy data in the 1841 census, which are based on self-reporting and refer to the population aged five and above, reflect the strong regional variation in standards of living: whereas less than one in four in County Antrim could neither read nor write, the proportion in County Mayo was four in five (BPP, 1843, p. lix).<sup>2</sup> The census defined fourth class houses as ‘all mud cabins having only one room’; the percentage ranged from 24.7 in County Down to 66.7 in County Kerry (BPP, 1843, pp. lvi–lvii). Figure 2 illustrates the variation in these variables at parish level, highlighting the desirability of higher resolution analysis. At the county level our two measures of poverty are strongly correlated with the income per capita measure used by Mokyr (which is available only this level).

Figure 2: Poverty Measures from the Census of Ireland 1841



The results from our regression analysis confirm that population pressure (defined as

---

<sup>2</sup>Our analysis considers those who can ‘read only’ to be illiterate. The ‘read only’ group accounts for 16.7 per cent of the population surveyed on literacy. Our results are not sensitive to this assumption.

our population to Poor Law land value variable) is positively associated with illiteracy and poor quality housing. To ease interpretation and comparison of the relative strength of coefficients, all of our regressors enter as standardised variables (using a  $Z$ -score transformation); in this way we can directly compare the influence of population pressure variable to a host of other covariates. This comparison reveals population pressure to be highly influential as in most model specifications the associated coefficient is relatively large in magnitude. Geographical remoteness also appears to be important, as variables measuring longitude (a proxy measure for distance to the larger industrial market in Britain), distance to Ireland’s two main cities, Dublin and Belfast, and the distance to the nearest town all have explanatory power. Institutional factors, captured by a variable measuring the proportion of individuals in a parish whose main source of income was ‘vested means’ also appears to play a role.

Comparing the coefficient on the population pressure variable to others in our model tells us about the relative importance of population pressure. However, it does not inform us of its importance in an absolute sense. To remedy this and to aid interpretation we conduct a back-of-the-envelope calculation that compares our estimates against a ‘no 19th century population growth’ benchmark. In other words, we use the estimated coefficient value to predict the absolute level of illiteracy and poor housing had the Irish population stayed at the 5 million level reached c.1800 (Daultrey et al., 1982). This exercise suggests that in the absence of post-1800 population growth, there would have been a 2 percentage point decrease in illiteracy. Given that on average 73 per cent of parish populations could not read and write, this implies a counterfactual reduction in illiteracy of about 4 per cent. An equivalent exercise for poor housing suggests a counterfactual poor housing rate of 34.8 percentage points against the observed 36 per cent. Clearly we cannot know that our counterfactual scenarios would have manifested themselves had the condition of no population growth been met. Still, such outcomes are hardly consistent with the view that Ireland on the eve of the Great Famine would have been substantially better off had she not experienced rapid population growth in the 18th and early 19th centuries. This finding squares with Mokyr’s conclusion that economic historians should look beyond simple Malthusian models to understand Ireland’s pre-Famine economy.

As mentioned, one advantage of having a larger number of observations is that it is pos-

sible to explore nuances in these data. Interestingly, the relationship connecting population pressure and poverty is robust to the inclusion of a large number of additional explanatory variables.<sup>3</sup> Similarly, we still find a population pressure effect when Poor Law Union fixed effects are included in our modelling specification. This suggests that a substantial part of the poverty-population pressure relationship occurs at a local level.

The potential endogeneity of the population pressure variable is of concern and we take a number of steps to counter this concern. The first, and most obvious, threat comes from omitted variable bias. For example, variation in both poverty and population pressure could be a function of factors such as geographic remoteness or land tenure institutions. To alleviate this concern we include several control variables and find that our population pressure coefficient is reasonably stable.

Reverse causality represents is another potential source of bias. Our regression approach assumes that population pressure has an effect on two dimensions of poverty. However, what if population pressure is itself caused by poverty, and what would such a relationship imply for our model estimates? We probe this issue by introducing cross-sectional variation in both migration and fertility. Whilst migration was a key demographic feature of the famine and post-famine periods, it was also reasonably common in the immediate pre-Famine decades. Indeed, Malthus himself proclaimed before a parliamentary inquiry that migration caused by Irish population growth would be ‘most fatal to the happiness of the labouring classes in England’ (BPP, 1827, p. 312). Fertility differences could also have played a role. In Connell’s account the rapid growth of the Irish population was preceded by the introduction of the potato which pushed down the age of marriage, and thus stimulated increased fertility amongst the poorest in Irish society. Alternatively, a parish in which population pressure was high but fertility low might be seen as one in which demographic adjustment was taking place through the preventive check.

If economic conditions had a substantial effect on migration or fertility then our population pressure coefficient might be biased. To measure the potential for this bias we include both a variable measuring fertility and another measuring pre-Famine migration. If population pressure does not have a causal effect on poverty, then its associated coefficient should tend towards the null once these variables are included. The limitation of this

---

<sup>3</sup>We also account for spatial autocorrelation using a spatial filtering approach (Murakami and Griffith, 2018). We elaborate upon this in Section 4.



approach is that the coefficient will shrink towards the null regardless of the potential for simultaneity, so there is an argument that controlling for these population growth factors should dilute the main effect. With this limitation in mind one can treat the population pressure coefficient as a lower bound in these specifications. Interestingly, both fertility and migration are somewhat influential, but the population pressure coefficient is remarkably stable across a number of regression models. Finally, we also match our parishes to the 1821 census and include a variable measuring population growth. Again, including this variable does not majority impact our estimates and supports the notion that population pressure is having a causal effect on poverty.

Another potential threat to our identification of population pressure effects is the possible endogeneity of Poor Law values. While the main focus of the valuation was land value, the valuation also included some buildings such as first-class houses, which might distort our estimates. Additionally, the land values used by the Poor Law Commission were not reported by landholders and agents in a uniform manner (as detailed in BPP, 1844, pp. iv–xiii). We assess the potential impact of endogenous Poor Law values by performing an instrumental variables analysis, wherein exogenous variation in terrain ruggedness is used as an instrument for Poor Law values. The results of this analysis are very much in line with our main results, supporting the assumption that the Poor Law value variable is a suitable proxy for land quality. Supplementary analyses also reveal that the population pressure effect was spatially homogeneous and approximately linear.

The rest of the paper proceeds as follows. Part 2 describes the role of Malthusian models in historical analysis generally and in historical writing on the Irish economy on the eve of the Great Famine of the 1840s. Part 3 describes the data used in our analysis of Irish poverty. Part 4 presents the results of our analysis, while Part 5 concludes.

## 2 Malthusian Models in Economic History

The enduring appeal of the Malthusian model lies in its simplicity. Any improvement in living standards will lead to population growth through an increase in the birth rate (the preventive check) and a reduction in the death rate (the positive check). But, unless such population growth is sustained by technological progress, the law of diminishing returns will

drive down income per capita, thus cancelling out all of the initial improvement. Whence the slow growth or stagnation associated with preindustrial societies. In this model the role of the law of diminishing returns rests on the supply of land being given: any increase in population necessitates the use of poorer quality land with lower yields. This process is known as the Malthusian Trap.

An industrial revolution is possible only after improvements in agricultural productivity have facilitated the transfer of labour from farm to factory, enabling population and income per capita to grow in tandem. This marks the escape from the Malthusian Trap. Non-Malthusian forces now lend a helping hand; population growth, instead of being a hindrance, may bolster economic fortunes either through Smithian mechanisms such as increased specialization, trade, and the expansion of markets, or population-driven technological change. This shift from Malthusian to post-Malthusian world represents the most important transition in economic history since the Neolithic revolution. How it happened has been the focus of research since the mid-nineteenth century. In recent years, several prominent theories and models of the industrial revolution have sought to explain how the Malthusian threat was gradually overcome in the 19th and 20th centuries (Galor and Weil, 2000; Hansen and Prescott, 2000; Clark, 2007).

When does an economy qualify as Malthusian? One common approach to testing the Malthusian model is to estimate the short-run response of demographic variables to wage and price shocks, using time-series data. Work in this tradition on England include Lee (1981), Nicolini (2007), Crafts and Mills (2009), and Kelly and Ó Gráda (2012, 2014). Studies of other regions in the past include Weir (1984) on France; Fernihough (2013) on Northern Italy; Pfister and Fertig (2010) on Germany; and Lagerlöf (2015) on Sweden. The empirical literature using variation in real wages or economic shocks along the time dimension appears to provide, at least, partial support for the Malthusian model. In England the disappearance of both check mechanisms and of major famines appears to have coincided with the beginnings of sustained economic growth in the seventeenth century. Similarly, the persistence of the check mechanisms in less-industrialized regions of Europe into the nineteenth century is confirmed, but the disappearance of famine in Italy and in France preceded the beginnings of sustained economic growth in those countries (Alfani and Ó Gráda, 2018).

Other research has used cross-sectional data to evaluate the usefulness of Malthusian models. Ashraf and Galor (2011) used the Maddison Historical Statistics database to estimate the relationship between land productivity, technological progress, population, and income per capita. Their results are in keeping with the conventional Malthusian approach: prior to the early modern period regions with higher land productivity had a technological advantage were more densely populated but no richer in terms of income per capita. Verpoorten (2012) also employs cross-sectional data, albeit in a modern setting, to estimate the effect of population density on the death toll during the Rwandan genocide of 1994. He finds that more densely populated administrative areas suffered disproportionately in this ‘neo-Malthusian’ context.

Reliable time-series data permitting an empirical study similar to those just described are lacking for pre-Famine Ireland. Undeterred by this gap, Mokyr (1985) applied a cross-sectional approach. Mokyr’s analyses, highly innovative in their day, regressed either income or wages on population pressure (defined as rural population per cultivated area, but using a range of proxies to control for land quality). Whilst the population pressure variable appeared to influence economic conditions in a number of Mokyr’s regressions, this relationship was far from robust and led Mokyr to conclude that pre-Famine economy should be viewed outside the narrow and simplistic scope of the Malthusian framework.

The results of Mokyr’s analysis were questioned in McGregor (1989), who argued that incorporating a better measure of land quality yielded results more in keeping with more traditional historical accounts of pre-Famine Ireland. Whilst Mokyr and McGregor rely on data collected at the county level ( $n = 32$ ), we use the more abundant parish-level data collected in the 1841 census ( $n = 2,387$ ). Fotheringham et al. (2013) and Kelly and Ó Gráda (2015) have also used recently used more disaggregated data—at district electoral division ( $n > 3,000$ ) and baronial ( $n > 300$ ) levels, respectively—to estimate alternative versions of the Malthusian model. But their focus was on population change between 1841 and 1851, not pre-Famine poverty. The aim of this paper is to examine the deeper roots of the famine, rather than its outcome. Using the higher resolution parish-level data is only of benefit if considerable spatial variation exists within the units of analysis used by existing studies. This appears to be the case here. Figure 1 highlights important intra-county variations hidden by county-level analysis. The presence of low population pressure

in livestock grazing areas in Roscommon and east Galway are good examples.

### 3 Data

The 1841 Census of Ireland (BPP, 1843) is the most comprehensive data source on demographic and economic conditions in pre-Famine Ireland. Whilst nearly all of the individual census returns were destroyed in a fire in the Public Records Office in 1922, many aggregated variables survive in the published census report and these data are also provided at a high spatial resolution (as shown in Figures 1 and 2).

The reliability of the 1841 census has been questioned on certain points and it is commonly accepted that both the 1821 and 1841 censuses undercounted the population (Lee (1968); Mokyr (1985, p. 31)). The 1841 census has also been criticised for inaccuracies in the reporting of age distributions used to infer fertility Tucker (1970) and also the reporting of agricultural statistics related to land use (Bourke, 1959/60). Others, however, have defended the census’s accuracy and, in particular, for its suitability for the kind of cross-sectional regression analysis performed here (Almquist, 1979). For the most part we avoid intercensal comparisons; by focusing instead on time-invariant comparisons across civil parishes in 1841, we avoid the potential pitfalls resulting from changes in the quality of coverage. The one exception is a variable measuring population growth between 1821 and 1841.

Table 1 displays the summary of statistics the variables used. We merge the census data with the spatial boundaries provided by a GIS shapefile source.<sup>4</sup> Merging these two data sources means we can further merge these data to geographic data from raster graphics and append numerous additional variables to these data that were not provided in the census report. The use of GIS data affords us the opportunity to visualize the geographical variance in these data through high-resolution maps. One slight disadvantage of merging the 1841 census data to the ‘townlands.ie’ source was that civil parish boundary changes meant that a number of parishes listed in the 1841 report had to be joined to a larger contiguous parish. This merging explains why our parish-level analysis is based on 2,387 observations rather than the 3,311 listed in the official report. The Census reports

---

<sup>4</sup>Source: <https://www.townlands.ie/>.

Table 1: Summary Statistics: 1841 Parish-Level Data

Statistic	N	Mean	St. Dev.
Fourth Class Houses (Per Cent)	2,387	36.235	18.356
Illiteracy Rate (Per Cent)	2,387	72.986	9.633
Male Illiteracy Rate (Per Cent)	2,387	63.738	10.975
Female Illiteracy Rate (Per Cent)	2,387	82.091	9.278
Population, 1821	1,599	2,857.169	2,774.767
Population, 1841	2,387	3,259.678	3,734.376
Pop. Growth 1821–41 (Per Cent)	1,519	21.628	32.983
Area (Acres)	2,387	8,641.286	10,609.290
Poor Law Land Value per Acre	2,387	0.732	0.997
Pop. (1841) Pressure (Person per Land Value)	2,387	0.802	1.346
Pop. (1841) Density (Person per Acre)	2,387	0.438	0.493
Child (0–4) per Families Ratio	2,387	0.723	0.119
Coast within 5km	2,387	0.170	0.376
Occupation: Agriculture (Per Cent)	2,387	74.851	14.282
Occupation: Manufactures/Trade (Per Cent)	2,387	16.756	10.897
Occupation: Other (Per Cent)	2,387	8.393	7.160
Roman Catholic (Per Cent)	2,387	85.761	21.334
Parish Town Population (Per Cent)	2,387	9.697	17.613
Distance to Nearest Town (km)	2,387	9.051	5.666
Vested Means (Per Cent)	2,387	2.082	1.966
Pre-1846 Migration	2,387	0.714	1.750
Terrain Ruggedness Index	2,387	9.162	9.700
Distance to City (km)	2,387	121.693	70.758
Distance to 17th Century Road (km)	2,387	8.717	11.309
Longitude	2,387	−7.597	1.055
Latitude	2,387	53.171	0.868
Sex Ratio	2,387	0.991	0.089

the population and other demographic data on a parish-level basis. Nearly a thousand parishes contain rural and ‘urban’ areas which are reported separately. The ‘urban’ section of each parish relates to a village or multiple villages and/or a town or towns within the civil parish. Essentially these denote the non-agricultural portion of the parish. In our data we consider each parish’s population and other demographic statistics to be comprised of all individuals living in the parish.<sup>5</sup>

In 1841 over 36 per cent of Ireland’s houses were deemed to belong to the lowest, or fourth, class. As noted earlier, a typical fourth-class house was a windowless one-room mud cabin. Other houses were described as first-, second-, or third-class, depending on the number of rooms, windows, and building material. Poor-quality housing is our first measure of poverty in pre-Famine Ireland. The second is a measure of illiteracy at parish level. Table 1 reports three different measures of illiteracy: full, male, and female. Illiteracy here is defined as the proportion of the population aged five and above without the ability to both read and write. We distinguish between male and female literacy as we believe they potentially measure slightly different elements of child investment. Male literacy is more likely to reflect parent’s investing in their son’s education so as to boost their human capital and thus employment prospects in later life. This was less of a concern for females, as thus variation in female literacy is more likely to reflect education as a consumption good rather than an investment good amongst parents (Reis, 2005).

The two measures of poverty employed here overlap. Figure 2 shows that, broadly speaking (but with some interesting exceptions such as much of Waterford in the southeast), areas where literacy was high contained fewer fourth-class houses and vice versa. The strong association between the two variables underlines their relevance in measuring poverty across Ireland in 1841, as reflected in a strong bivariate correlation coefficient of 0.5. The correlation between our outcome measures implies that they should be modelled as a joint multivariate distribution, something we examine in Section ??.

Irish civil parishes contained an average of just over three thousand people in 1841. In Section 1 we introduced the concept of ‘population per value adjusted acre’ as the most appropriate measure of population pressure. More conventional accounts, such as that used in Mokyr (1985) use population per cultivated acre as a measure of population pressure.

---

<sup>5</sup>In a separate analysis we considered how our results change when the parish population consists only of individuals living in the rural portion. The results, which were qualitatively similar, are available upon request.

However, we believe that population per value adjusted acre to be a superior measure because all land is not equal. In two equally populated parishes, the parish with poorer land faces greater population pressure. We use the land values reported in the 1851 Census (BPP, 1856) as our measure of land quality.<sup>6</sup> Known as the ‘Poor Law Value’, this measure determined the rate to be paid by the landowner for local poor relief. Furthermore, these values only relate to the rural portion of each parish, with urban areas such as villages and towns excluded (and thus the value of the properties within). Finally, on rural land the value of all houses and structures whose annual rent was less than £5 was omitted from this overall value, with all houses and structures above this threshold only included at two thirds of their quoted value.

Between 1847 and 1864 a “Griffith’s Valuation” was published for each of Ireland’s 32 counties. These surveys were diligently constructed under the stewardship of Sir Richard Griffith, a geologist employed by the Westminster government to provide a database of land values throughout Ireland. Given that a number of counties in Ireland had yet to be fully surveyed by Griffith’s department, the new Poor Law of 1838 had to rely on an alternative measure of land valuation. For this the Dublin-based Poor Law commission transferred this responsibility to the Board of Guardians of the 130 Poor Law Unions. Whilst each Poor Law Union was told to assess land on the basis of what one might consider a fair rent, the interpretation and implementation of this instruction varied from union to union. Furthermore, the value of rent also depended on the local price of produce. In theory, two estates with the same quality and quantity of land could have different land values if the price of agricultural goods differed in local markets. Differences in land valuation could also emerge as a result of market potential. For example, a parish further from a town or city would have a lower land value regardless of land quality.

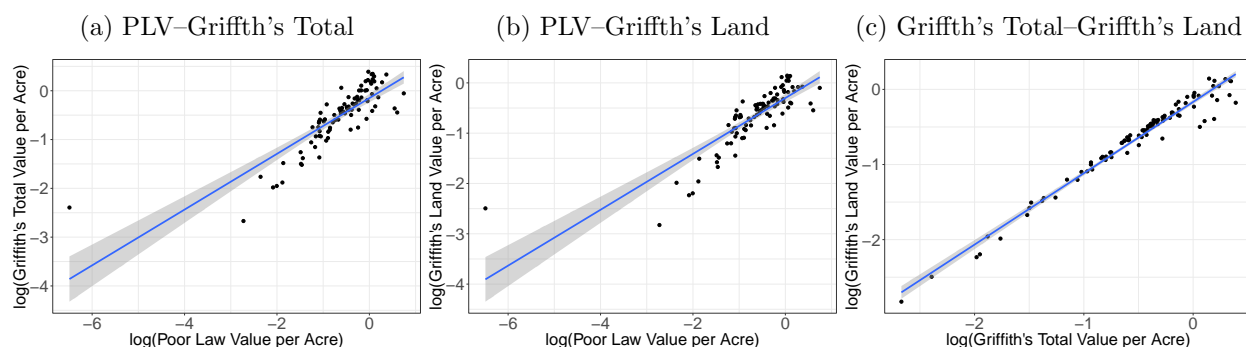
The Poor Law Value provides an exceptionally high resolution snapshot of land quality in Ireland. However, as evident in the previous paragraph, we must be willing to accept a number of caveats surrounding the application of this measure for use as a deflator in the construction of our population pressure variable. Fortunately, we can address these limitations in a number of ways. One way in which we can address this shortcoming is to compare the land values reported in the census to those published in Sir Richard Griffith’s

---

<sup>6</sup>Whilst these values are reported for 1851 they are unaffected by the famine as the valuations used for Poor Law rates were applied from 1838 onwards.

comprehensive survey.<sup>7</sup> Ideally, we would use the Griffith's land values as an alternate measure of land quality, however collected and matching between these two data sources would be extremely time consuming. Instead we sampled 100 parishes at random from our census data, of which we were able to match 95 to a corresponding entry in the Griffith's valuation ledger. The correlations between our data and Griffith's are shown in Figure 3. The first panel displays the relationship between the logged Poor Law value (PLV) per acre and logged Griffith's total value per acre. The bivariate correlation coefficient of 0.83 points alleviates concerns one might have about the use of the PLV. Griffith's survey reports both the value of land and buildings separately, something we explore in the middle and right-hand plots of Figure 3. Plot (b) of Figure 3 replicates plot (a) with the exception that Griffith's land value is used instead of the total value that includes buildings. The difference in this instance appears to be negligible (the correlation coefficient increases marginally to 0.84), as one might expect seeing as buildings only accounted for 16 per cent of total value in our sample of 95 parishes. The final plot, plot (c), emphasizes this point as one can see the correlation between the total land value and land (excluding buildings) reported in Griffith's survey is nearly one-for-one (the correlation coefficient here is 0.98).

Figure 3: Comparisons Between the Reported Poor Law Land Value and Griffith's Land and Total Value



The comparisons shown in Figure 3 reveal no major discrepancies in our use of the Poor Law values. We can further tackle the concern that systematic differences across Poor Law Unions bias our econometric results by superimposing the Poor Law Union boundaries upon our data and adding Poor Law Union fixed effects. Likewise, concerns that the Poor Law values varied by remoteness and geographic proximity to market towns can also be

<sup>7</sup>The ledger returns from Griffith's Survey were taken from the following site: <http://www.askaboutireland.ie/griffith-valuation>.



addressed with the inclusion of a number of additional explanatory variables such as distances to towns or cities, latitude, and longitude. Finally, Figure 4 (adapted from Ó Gráda (1999, p.140)) alleviates concerns that substantial regional price differences in agricultural produce existed. Another reassuring point worth emphasising is that the small price differences observed in Figure 4 might themselves be a function of land quality. Overall, we believe that the Poor Law valuation provides a plausible portrayal of land quality in pre-Famine Ireland although in our regression analyses we also examine the potential impact that misreported and/or endogenous land values might have for our population pressure estimates.

Figure 4: Regional Variation in Potato Prices (1840–6)

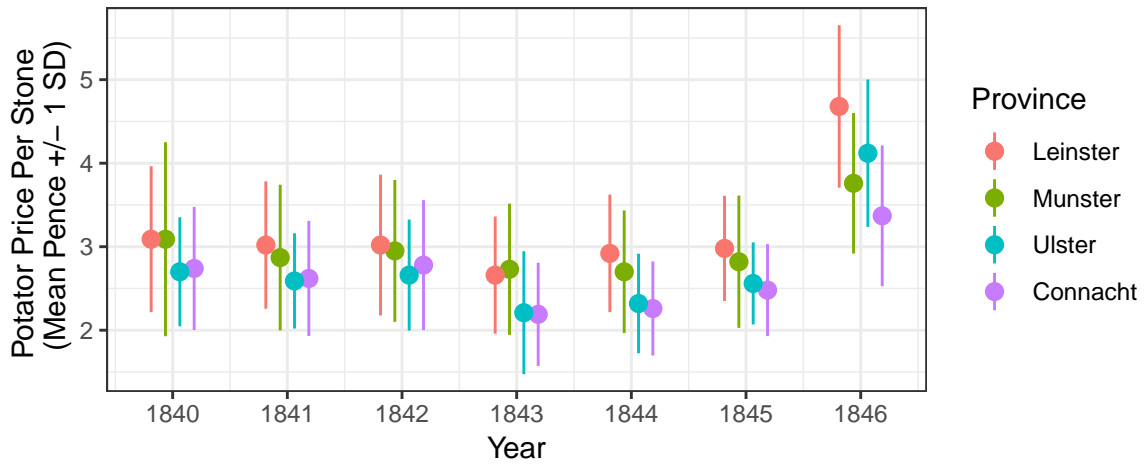


Table 1 also features a number of additional variables that we use as explanatory co-variables in our regression models. The East-West division is apparent in both Figures 1 and 2. Both figures strongly suggest that geographic remoteness and market potential play important roles, a topic stressed in development economics (Redding and Sturm, 2008). To capture this potentially offsetting factor our dataset includes longitude and latitude. Another measure of remoteness/market potential is the distance between each parish and various points of economic importance. With this in mind we include a measure of the distance to a major city (the minimum distance to either Belfast or Dublin) and also a measure of the distance to nearest town (towns here are designated as civil parishes with an urban population above 1,000 people). We also include a variable that measures the minimum distance to a 17th century road as detailed in William Petty’s General Map of

Ireland and ‘Coast within 5km’ as a measure of coastal access.<sup>8</sup> Table 1 shows that 17 per cent of parishes were located within 5km of a coast.

The economic structure of each parish is captured with three variables measuring the proportion of household heads (assumed to be the chief breadwinner) employed in either agriculture, manufactures or trade, or an “other” occupation (mainly the service sector). Ireland’s pre-Famine economy was overwhelmingly rural so it is no surprise that most household heads (around 75 per cent) were employed in the dominant agricultural sector. Manufactures and trade occupations account for a further 17 per cent of the economy, with “other” occupations representing the residual 8 per cent. This structure is further represented by the variable ‘Parish Town Population (per cent)’, a measure of the proportion of the parish population living in non-rural settlements of various sizes within the parish. The ‘Vested Means’ variable measures the proportion of families whose livelihoods are ‘chiefly dependent on vested means, professions, etc.’ as distinct from depending on either ‘the direction of labour’ or ‘their own manual labour’. Landowners who held over 50 acres of land fell in to this category alongside a small number of others with substantial wealth or capital. It represents the presence locally of an elite represented by white-collar occupations and rentiers, and hence a proxy for potential social capital, and non-agricultural employment. The ‘Vested Means’ variable, which is subject to a strong east-west gradient, can also be seen as a marker (albeit imperfect) for investment in schools, medical facilities, and other infrastructure. This variable also captures the “absentee landlord” effect, as it seems plausible to assume that the absence of landlords would be reflected in a reduction in the ‘Vested Means’ variable.<sup>9</sup>

We report the parish population in 1821 for 1,599 parishes. Boundary and name changes mean that we could not match all of the parishes exactly, although 67 per cent of parishes provides us with good coverage linking both censuses. ‘Pop. Growth 1821 to 1841’, a measure of parish population growth between 1821 and 1841, serves a double purpose in the estimation.<sup>10</sup> Besides serving as a check against reverse causation in the estimation (as

---

<sup>8</sup>The data for the roads is available at: <http://downsurvey.tcd.ie>.

<sup>9</sup>We also experimented with a variable describing the percentage of properties in an area auctioned in the Encumbered Estates Court in the wake of the Famine as a yardstick for the institutional quality of land management (Eiríksson and Ó Gráda, 1995). However this variable, available only at the more aggregated baronial level (Ireland’s 331 baronies were comprised of multiple civil parishes), lacked explanatory power and so we excluded it from the analysis.

<sup>10</sup>In our analysis we use 1,519 observations rather than 1,599. The reason for this is that some outliers existed

explained below), rapid population growth in a parish might also be seen as a by-product of absenteeism or landlord indebtedness. Table 1 also provides summary statistics for a terrain ruggedness index which we calculated based of a high-resolution elevation raster image using the method of Wilson et al. (2007).<sup>11</sup>

We add a proxy for variation in marital fertility on the eve of the Famine to our data by dividing the number of recorded children under the age of five in 1841 by the number of families resident in each parish. Another key population change indicator is migration. Here we capture pre-Famine migration by incorporating the recently released dataset from the Emigrant Industrial Savings Bank in New York (Wegge et al., 2017) to create a parish-level dataset of Irish migration. These data are available on a civil parish basis. Clearly, they tell only part of the Irish migration story as they omit Irish migration to both Britain and elsewhere in North America. However, given the pre-eminence of New York as a magnet for 19th century Irish emigrants this variable seems a plausible indicator of the extent to which migration might have influenced poverty. The parish sex ratio variable (the ratio of the male population aged five years or older to the equivalent female population) acts an alternative indicator of migration, as gender-selective migration has the potential to skew sex ratios. Finally, we also add a measure for the religion composition of each parish: the Roman Catholic share of the population. Unfortunately the 1841 or 1851 census did not survey religion. Data on religion for 1834 exist, but these have yet to be fully digitized and mapped. Thus, to incorporate religion in this we use data from the 1911 census, for which we have data on the religious make up of Ireland’s District Electoral Divisions (DED), and we superimpose these data on the existing parish data.<sup>12</sup> That a 60-year overlap exists between our data and this variable is of concern. However, the spatial variation of Ireland’s pre-independence demography did not change much over time.<sup>13</sup> Furthermore, Gregory and Cunningham (2016) show that the famine had an almost negligible impact on

---

in these data, so we trimmed the population growth variable at the 2.5th and 97.5th percentiles.

<sup>11</sup>These data are taken from the Global 30 Arc-Second Elevation Data Set (GTOPO30) and the raster cells are approximately 0.5km<sup>2</sup> (Source: [https://webmap.ornl.gov/ogcdown/wcsdown.jsp?dg\\_id=10003\\_1](https://webmap.ornl.gov/ogcdown/wcsdown.jsp?dg_id=10003_1)).

<sup>12</sup>To superimpose these data we convert the DED spatial polygon shapefile into a spatial point shapefile. We then assigned each parish a Catholic (per cent) based on where the DED spatial points lay. If one spatial point fell inside a parish’s boundaries then this was the assigned value used. In cases where multiple spatial points fell inside a parish we used the average of the Catholic variable. In the small number of cases where a parish did not contain a spatial point we used spatial kriging techniques to interpolate (Pebesma, 2004).

<sup>13</sup>E.g. the share of Catholics hardly changed across Ireland between 1861 and 1911 (Fernihough et al., 2015).

the percentage of Catholics in an area.<sup>14</sup>

## 4 Results

### 4.1 OLS Regressions

Our analysis estimates the following two linear regression models:

$$\text{Illiteracy}_i = \alpha_1 + \beta_1 \ln(\text{PopPres}_i) + \mathbf{X}\boldsymbol{\gamma}_1 + \epsilon_i \quad (1)$$

$$\text{Poor Housing}_i = \alpha_2 + \beta_2 \ln(\text{PopPres}_i) + \mathbf{X}\boldsymbol{\gamma}_2 + v_i \quad (2)$$

where poverty, as measured by either illiteracy or poor housing, is a function of population pressure ( $\text{PopPres}_i$ : population per value adjusted acre) and other factors. The matrix  $\mathbf{X}$  contains all additional explanatory variables and fixed effects, which we allow to vary depending on the model specification. To simplify the interpretation of our regression results we recast all of our explanatory variables using a  $Z$ -score transformation. This means that all of the coefficients can be interpreted as the effect of a one standard deviation movement in the variable in questions. This allows us to easily assess the relative magnitude of the coefficients.

Our two measures of poverty are expressed in percentage point terms. This means that we can interpret population pressure coefficients ( $\beta_1$  and  $\beta_2$ ) as the effect of a one standard deviation change in population pressure. To help put our estimates into context it is worth considering what a one standard deviation change in population pressure represents. Daultrey et al. (1982) estimated that Ireland's population in 1800 was approximately 5 million, by 1841 this number had risen dramatically to (at least) 8.2 million—a 63.5 per cent increase. We assume land quality, as measured by the land value variable, is fixed, so there was a 63.5 per cent rise in population pressure.<sup>15</sup> After creating a synthetic 1841 parish population dataset we found that moving from a population of 5 to 8.2 million is

---

<sup>14</sup>Figure 1 in Gregory and Cunningham (2016) is an  $x$ - $y$  scatterplot showing the relationship between the per cent Catholic in 1834 and the equivalent variable in 1861. The reported regression line ( $y = 0.06 + 0.99x$  with an  $R^2 = 0.98$ ) shows and almost one for one relationship between the two variables.

<sup>15</sup>Of course land quality is not fixed in the long run. Improved drainage and irrigation are just two methods by which land quality could be improved. However, there is little evidence to suggest that there was a dramatic improvement in land quality between 1800 and 1841 given the prevalence of rent-racking in pre-Famine Ireland.

equivalent to a 0.75 standard deviation movement in population pressure. Therefore one can multiply any of the coefficients by 0.75 to get an approximate ‘no post-1800 population growth’ counterfactual.

Table 2: OLS Regression of Illiteracy (Per Cent) on Population Pressure

	(1)	(2)	(3)	(4)	(5)
Intercept	36.235*** (0.302)	72.986*** (0.106)	76.309*** (3.063)	72.986*** (0.106)	73.435*** (0.124)
Population Pressure	3.840*** (0.411)	2.714*** (0.177)	2.767*** (0.187)	2.680*** (0.178)	2.913*** (0.224)
Fertility				0.199* (0.121)	0.142 (0.154)
Migration				−0.099 (0.124)	−0.082 (0.147)
Pop. Growth 1821–41					0.399*** (0.134)
Coast		−0.050 (0.133)	0.120 (0.144)	−0.054 (0.133)	0.071 (0.155)
Occupation: Manufactures/Trade (Per Cent)		−1.436*** (0.209)	−1.643*** (0.210)	−1.390*** (0.210)	−1.421*** (0.254)
Occupation: Other(Per Cent)		−0.737*** (0.152)	−0.631*** (0.147)	−0.707*** (0.153)	−0.948*** (0.180)
Catholic (Per Cent)		2.665*** (0.218)	1.995*** (0.232)	2.674*** (0.218)	2.899*** (0.247)
Parish Town Population (Per Cent)		−0.942*** (0.180)	−0.777*** (0.177)	−0.901*** (0.182)	−1.094*** (0.218)
Distance to Nearest Town (km)		0.355*** (0.126)	0.537*** (0.143)	0.338*** (0.126)	0.378** (0.150)
Vested Means (Per Cent)		−1.582*** (0.135)	−1.695*** (0.134)	−1.585*** (0.135)	−0.996*** (0.158)
Distance to City		4.015*** (0.929)	2.970 (2.000)	4.075*** (0.931)	2.666** (1.122)
Distance to 17th Century Road		−0.059 (0.160)	0.404 (0.267)	−0.064 (0.160)	−0.262 (0.208)
Sex Ratio		−0.543*** (0.126)	−0.399*** (0.122)	−0.548*** (0.126)	−0.843*** (0.151)
Latitude		2.742*** (0.379)	−1.555 (1.440)	2.796*** (0.380)	2.557*** (0.443)
Longitude		−0.754 (0.806)	−1.288 (1.676)	−0.716 (0.807)	−1.700* (0.940)
PLU FE	N	N	Y	N	N
Observations	2,387	2,387	2,387	2,387	1,519
R <sup>2</sup>	0.357	0.715	0.774	0.715	0.749
Adjusted R <sup>2</sup>	0.352	0.711	0.754	0.711	0.744
F Statistic	87.595***	218.725***	38.664***	203.880***	148.063***

*Notes:* All columns regress the illiteracy (per cent) variable on the indicated covariates. All covariates are expressed in terms of Z-scores and thus the coefficients are comparable. Residual autocorrelation was addressed in all models by transforming the outcome and all regressors via the eigenvector spatial filtering method detailed in Murakami (2018). \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

Table 2 presents our first set of regression results, where illiteracy (per cent) is the dependent variable. The regression results are the coefficient estimates alongside a corresponding standard errors. Estimation of equation (1) by Ordinary Least Squares assumes independence of the error term  $\epsilon_i$ . In this application such an assumption is untenable due to spatial autocorrelation. The unobserved component of illiteracy for parish

$i$  is more likely to be similar to the same component for parishes in closer proximity to  $i$  than further away.<sup>16</sup> To alleviate this issue we apply the spatial filtering approach advocated by Murakami and Griffith (2018). This methodology tackles the issue of spatial autocorrelation by extracting candidate eigenvectors from a spatial connectivity matrix that describes the spatial structure of the data and then using these selected eigenvectors as additional covariates in the regression model under consideration. In this application we use a spatial weights matrix where all parishes within 50km are assumed to be neighbours and the eigenvectors were extracted from the spatial connectivity matrix under the assumption that they pass the inclusion threshold of 0.25.<sup>17</sup> The advantage of this approach compared to the spatial lag regression model is that once the candidate eigenvectors have been extracted the model can be estimated and interpreted using the same framework as standard OLS, whereas one cannot apply OLS to estimate a spatial lag model and interpretation also requires one to account for both direct and indirect effects.

The first column in Table 2 reports the results from a simple (spatially filtered) bivariate regression of illiteracy on population pressure. The point estimate the point estimate is 3.8. Column (2) reports the results for our specification that includes a large number of control variables. As the variables are all measured as  $Z$ -scores the coefficients are comparable and the population pressure variable, with at point estimate of 2.7 ranks as one of the most important explanatory variables. Most of the other covariates also exhibit explanatory power and the signs of these coefficients are in line with our prior expectations. Geographically remote, agricultural parishes appear to fare worse in terms of illiteracy. The variable measuring the Roman Catholic proportion is also highly influential, something we believe partially captures illiteracy in Irish-speaking parishes. Furthermore, the presence of those whose income derived from vested means, the rentier class, is also significant.

Column (3) adds Poor Law Union (PLU) level fixed effects to the model specification. This was the administrative boundaries upon which the Poor Law value may have varied systematically due to differences the competency or integrity of the local evaluators. These fixed effects control for all variation between Ireland’s 130 PLUs meaning that the coeffi-

---

<sup>16</sup>An examination of the residuals when the above models are run using conventional OLS reveals reveals a substantial degree of spatial autocorrelation. Once we apply the spatial filtering method of Murakami and Griffith (2018) we find a marked reduction in residual spatial autocorrelation.

<sup>17</sup>These are two very typical choices for spatial connectivity and eigenvector threshold. Our results are not sensitive to our choice of spatial connectivity matrix or inclusion threshold.

cients here now represent variation at a local level. In other words, the variation driving these results occurs within each PLU, meaning the results can only be the result of local level differences. Despite the extra explanatory power the inclusion of the PLU-level fixed effects has the population pressure coefficient is effectively unchanged. Bar latitude, the results from column (2) remain largely intact.

The fourth column replicates the regression model in column (2) but adds variables that measure fertility (the child-family ratio) and migration. One potential concern is that the association between population pressure and illiteracy found in the first three columns reflects a feedback effect whereby economic backwardness influences population pressure. If the feedback relationship is positive then we expect that including the attributing sources of population growth differences will cause the population pressure coefficient to shrink towards the null. The downside of this approach is that including these variables as controls will reduce the coefficient towards the null, even when no such feedback mechanism exists. Expressed in a different way, the inclusion of the fertility variable controls for the channel through which a negative income-fertility gradient causes population pressure in poorer areas. Similarly, the inclusion of the migration variable reduces the possibility that illiteracy prevents emigration. Despite these concerns the coefficient on the population pressure appears to be robust to the aforementioned feedback effects. Although one might expect the coefficient to decline it is qualitatively unchanged.

Column (5) further probes the potential for simultaneity bias by including 1821 to 1841 population growth as an additional explanatory variable. Our rationale for including this variable is as before: if a substantial feedback effect exists then this should cause a reduction in the population pressure variable's coefficient. Unlike fertility and migration which are specific factors, population growth is general and incorporates fertility, migration, and mortality differences between parishes. One downside to using this variable is that we do not have the 1821 population total for every parish and thus must use a sample of the full dataset. Still, despite the change in data the results shown in the final column tally well with our previous findings. Once again the population pressure variable is influential and even increases in magnitude. The population growth variable exhibits a positive influence on illiteracy. That both fertility and population growth exhibit positive effects on illiteracy appears somewhat at odds with a conventional Malthusian interpretation of the Irish econ-

omy. One would expect that in an economy where population pressure was binding for the positive or preventive checks to be in operation. However, the coefficients on population growth and fertility in column (5) do not support this.

Our findings in columns (4) and (5) help alleviate concerns that the population pressure effect is due either wholly or in part by omitted variable bias or reverse causality. This relationship is robust to the inclusion of coincidental variables capturing geographic remoteness, market potential, and institutional differences. In all of the specifications we find population pressure to be important in relative terms compared with the other variables included in this analysis. We have not yet considered what the population pressure variable means in absolute terms. To do this we can perform our ‘no 19th century population growth’ counterfactual by multiplying the population pressure coefficient by  $-0.75$ . Our results indicate the effect is approximately  $-2$ . Considering that the illiteracy rate was around 73 per cent, this effect appears small in absolute magnitude.

In separate analyses we regress male and female literacy on our covariates to see if this creates any difference. However, the results are almost identical (see Appendix B). Table 3 replicates the regression analysis reported in Table 2 but uses the prevalence of fourth-class houses, our alternative measure of poverty/economic backwardness, as the outcome variable. Overall, the results are similar to those in Table 2. The population pressure coefficient is of a similar magnitude and a number of the other variables included are also relevant. Again geographic remoteness appears to play a role, as does the “vested means” variable. Column (1) implies that there is a strong relationship between population pressure and poor housing, although this relationship is greatly reduced once the various control variables are also included. The population pressure coefficient increases by a sizeable amount—from 3.0 to 3.5—once PLU fixed effects are added to the model specification.

The final two columns of Table 3 once again address the issue of reverse causality. Previously, fertility appeared to be slightly more important than migration. This finding is not repeated here, as the migration variable is the component of population change which appears to matter (as well as population growth between 1821 and 1841 itself). Parishes where more people were able to leave and cross the Atlantic to New York had a lower prevalence of poor housing. One could speculate that the option to migrate acted as a safety valve and reduced poverty, although this variable’s lack of influence in Table 2



Table 3: OLS Regression of Fourth Class Houses (Per Cent) on Population Pressure

	(1)	(2)	(3)	(4)	(5)
Intercept	36.235*** (0.302)	36.235*** (0.278)	54.870*** (8.251)	36.235*** (0.278)	36.968*** (0.345)
Population Pressure	3.840*** (0.411)	2.955*** (0.466)	3.532*** (0.504)	2.987*** (0.468)	1.896*** (0.623)
Fertility				-0.431 (0.317)	-0.538 (0.429)
Migration				-0.611* (0.325)	-1.119*** (0.409)
Pop. Growth 1821-41					1.317*** (0.372)
Coast		-0.329 (0.349)	-0.531 (0.388)	-0.338 (0.349)	-0.909** (0.431)
Occupation: Manufactures/Trade (Per Cent)		-1.335** (0.549)	-1.818*** (0.564)	-1.322** (0.553)	-1.270* (0.706)
Occupation: Other (Per Cent)		-0.908** (0.398)	-0.817** (0.395)	-0.978** (0.401)	-1.659*** (0.501)
Catholic (Per Cent)		0.115 (0.573)	0.133 (0.626)	0.109 (0.573)	0.512 (0.685)
Parish Town Population (Per Cent)		-1.600*** (0.472)	-1.457*** (0.476)	-1.518*** (0.477)	-0.771 (0.605)
Distance to Nearest Town (km)		1.522*** (0.330)	0.936** (0.386)	1.531*** (0.331)	2.072*** (0.416)
Vested Means (Per Cent)		-1.265*** (0.356)	-1.490*** (0.361)	-1.251*** (0.356)	-0.693 (0.440)
Distance to City		9.190*** (2.442)	3.375 (5.388)	8.863*** (2.445)	4.890 (3.115)
Distance to 17th Century Road		-0.115 (0.421)	0.215 (0.720)	-0.104 (0.421)	0.157 (0.578)
Sex Ratio		-0.099 (0.331)	-0.205 (0.329)	-0.088 (0.331)	-0.470 (0.420)
Latitude		-1.312 (0.997)	-9.665** (3.879)	-1.302 (0.999)	-0.494 (1.231)
Longitude		2.242 (2.117)	0.693 (4.515)	1.932 (2.120)	-1.554 (2.611)
PLU FE	N	N	Y	N	N
Observations	2,387	2,387	2,387	2,387	1,519
R <sup>2</sup>	0.357	0.457	0.548	0.458	0.473
Adjusted R <sup>2</sup>	0.352	0.451	0.508	0.452	0.462
F Statistic	87.595***	73.575***	13.700***	68.779***	44.537***

Notes: All columns regress the fourth class houses (per cent) variable on the indicated covariates. All covariates are expressed in terms of  $Z$ -scores and thus the coefficients are comparable. Residual autocorrelation was addressed in all models by transforming the outcome and all regressors via the eigenvector spatial filtering method detailed in Murakami (2018). \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

cautions against any concrete conclusions here. If we accept that the population pressure coefficient in this model is 3 we find a ‘no 19th century population growth’ counterfactual of:  $-0.75 \times 3 = 2.25$ . Again this result cautions against overstating the overall importance of population pressure despite the relative influence and strong fit of our regression model results.

## 4.2 IV Regressions

So far our analysis has assumed a specific linear function form. Poverty is a linear function of the logged ratio of population to land value. By the rules of logs the following specification is equivalent:

$$\text{Illiteracy}_i = \alpha_0 + \beta_1 \ln(\text{PopDens}_i) + \beta_2 \ln(\text{ValuePerAcre}_i) + \mathbf{X}\boldsymbol{\gamma}_1 + \epsilon_i \quad (3)$$

$$\text{Poor Housing}_i = \alpha_1 + \beta_1 \ln(\text{PopDens}_i) + \beta_2 \ln(\text{ValuePerAcre}_i) + \mathbf{X}\boldsymbol{\gamma}_2 + v_i \quad (4)$$

where the coefficient on population on both population density (per acre) and land value per acre is restricted. In Section 3 we introduced the land value variable and provided evidence that supported its use as a suitable measure to create our population per adjusted acre variable. In this section we use econometric procedure to further allay concerns that our previous results are affected by endogeneity bias that stems from the use of this variable. For example, illiteracy might force those with low human capital to farm land in a way that is not conducive to improving land quality. To remedy this, we instrument the land value measure with our measure of terrain ruggedness. For the purposes of our empirical strategy, the terrain ruggedness variable is exogenous, created by geography. Our identifying assumption is that terrain ruggedness can only influence poverty through its effect on land quality which we proxy using the land value per acre variable.

Table 4 shows the results for models that regress illiteracy and poor housing on population density, land values, and all of the other control variables as indicated in column (2) of Tables 2 and 3.

Column (1) of Table 4 shows the unrestricted model results. In this illiteracy is regressed on population density and value per acre instead of population pressure. The results here are consistent with our use of the population pressure variable, the coefficient on

Table 4: OLS and IV Regressions with Population Density and Land Values Per Acre

	Illiteracy (Per Cent)		Fourth Class Houses (Per Cent)	
	(1)	(2)	(3)	(4)
Population Density	2.223*** (0.196)	2.881*** (0.378)	1.804*** (0.514)	2.142** (0.982)
Value per Acre (OLS)	-2.968*** (0.206)		-3.652*** (0.540)	
Value per Acre (IV)		-4.446*** (0.753)		-4.412** (1.956)
Controls	Y	Y	Y	Y
Model	OLS	IV	OLS	IV
First-Stage F-Stat		194.377		194.377
Observations	2,387	2,387	2,387	2,387
R <sup>2</sup>	0.715	0.708	0.458	0.458
Adjusted R <sup>2</sup>	0.711	0.705	0.452	0.452
F Statistic	210.961***		71.286***	

*Notes:* Columns (1) to (2) regress illiteracy (per cent) and (3) to (4) regress fourth class houses (per cent) on the indicated covariates and the full set of control variables (as in column (2) of the previous regression tables). Residual autocorrelation was addressed in all models by transforming the outcome and all regressors via the eigenvector spatial filtering method detailed in Murakami (2018). \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

population pressure is positive and the value per acre variable is negative. Column (2) permits endogenous land values by instrumenting the value per acre variable with terrain ruggedness. The results in this column also list the first-stage  $F$ -statistic which is around 194, well over the oft-cited benchmark of 10 (Stock and Yogo, 2005). The IV results indicate that, if anything, endogenous land values actually increase the population pressure effect. However, this difference, moving from 2.2 to 2.9 is not substantial enough to alter any of the interpretation or analysis from the OLS results subsection. The value per acre is further away from the null, increasing from  $-3$  to  $-4.4$ , however there is also a marked increase in the standard error.

The final two columns of Table 4 repeat this exercise but use our measure of poor housing as the dependent variable. Again, the outcome validates our use of the population pressure variable as the coefficient estimates are in line with our expectations. Column (4) reports the results obtained from the IV regression. As in the first two columns, the coefficient signs are as we would expect: population density is associated with a greater proportion of fourth class houses, and value per acre is associated with fewer. Again, the IV estimates are larger in magnitude. This indicates that, if anything, endogenous land values are biasing the population pressure coefficient towards the null although it should

also be noted that the value per acre coefficient is also estimated with a larger degree of uncertainty so the larger standard error estimates rule out any conclusive inference on this issue. Furthermore, the value per acre coefficient is unanimously stronger than the population density value, which shows the importance of using this variable to capture differences in land quality.

### 4.3 Bayesian Multivariate Model with Heterogeneous Effects

Our results thus far have modelled the two poverty dependent variables separately and assumed that population pressure had a homogeneous effect. In this subsection we relax these restrictions by estimating a multivariate multilevel model that includes varying intercepts and population pressure slopes.

Figure 5: Poor Housing-Illiteracy Scatterplot, Correlation = 0.50

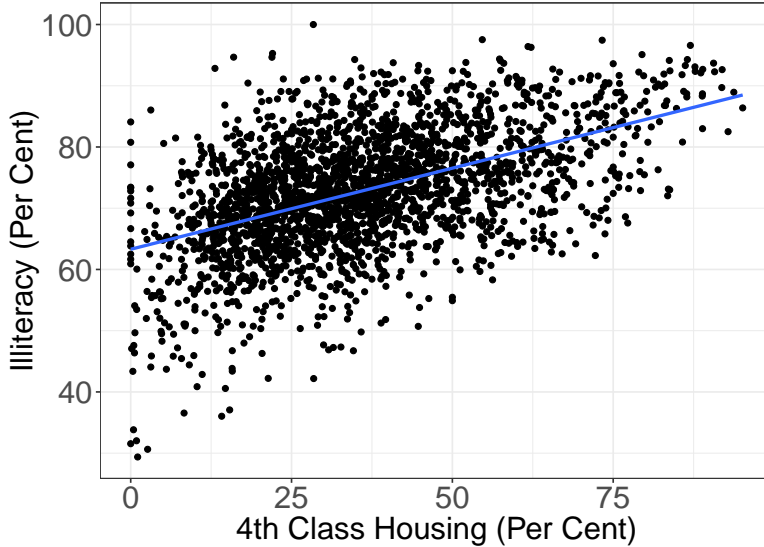


Figure 5 illustrates that a strong, albeit imperfect, correlation exists between poor housing and illiteracy. This implies that potential efficiency gains could be achieved by jointly modelling both illiteracy and poor housing. We blend this multivariate approach with a multilevel one, wherein the model equations take on a hierarchical structure with heterogeneous intercepts and population pressure effects. We estimate the following two

equations:

$$\text{Illiteracy}_i = \alpha_{1,\text{County}} + \beta_{1,\text{County}} \ln(\text{PopPres}_i) + \mathbf{X}\boldsymbol{\gamma}_1 + \epsilon_i \quad (5)$$

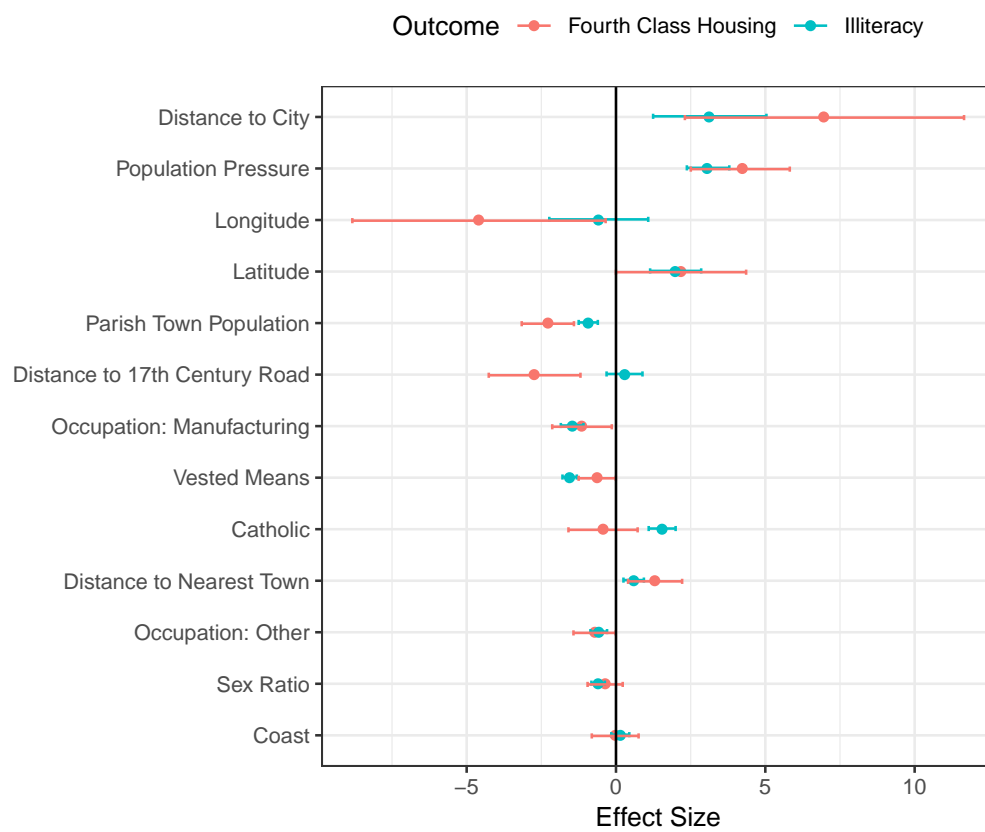
$$\text{Poor Housing}_i = \alpha_{2,\text{County}} + \beta_{2,\text{County}} \ln(\text{PopPres}_i) + \mathbf{X}\boldsymbol{\gamma}_2 + v_i \quad (6)$$

where the errors  $\epsilon_i$  and  $v_i$  correlate and we permit the intercepts ( $\alpha$ ) and slope ( $\beta$ ) terms to vary according to the county in which the parish  $i$  is situated. Additionally, we also model the correlation between the county-varying intercept and slope terms in each equation. If a positive correlation exists between  $\alpha_{1,\text{County}}$  and  $\beta_{1,\text{County}}$ , this would indicate that the population pressure effect was stronger in counties with greater illiteracy. The matrix  $\mathbf{X}$  contains the full set of additional covariates included in column (2) of Tables 2 and 3.

Given the hierarchical (where some parameters are nested) structure of the model, we adopt a Bayesian approach to inference (Gelman and Hill, 2006). Another advantage of using Bayesian over frequentist inference is that it provides the entire distributions for all parameters of interest encompassing the fixed model coefficients, group-specific parameters, and the variance and correlation coefficients. Please refer to Appendix C for the complete overview of the likelihood, linear model, and prior distributions used. Estimation was conducted using the `brms` package in R (Bürkner, 2017). This package implements the No-U-Turn Sampler (NUTS) from the language `Stan`. The NUTS sampler has been shown to converge much quicker than others used in Bayesian inference. Quick convergence in posterior draws was also achieved in this application as the 1,000 posterior draws we use (after discarding an initial 1,000 burn-in draws) displayed no sign of in-chain autocorrelation.

The coefficient plot in Figure 6 displays the posterior distributions of all the upper-level regression model slopes. As before, these variables are expressed in terms of  $Z$ -scores and have been spatially filtered. Figure 6 orders the variables according their collective influence on poor housing and illiteracy. For the majority of the explanatory variables the effect sizes are consistent across both modelling equations. The minimum distance to either Belfast or Dublin captured by the “Distance to City” variables appears to exhibit the strongest influence on the poor housing variable. The median posterior draw for this parameters is 7.5. The influence of geographical remoteness on fourth class housing is further underlined by the relative strength of the longitude variable. It appears that parishes further from

Figure 6: Median Effect Sizes alongside 95 Per Cent Credible Intervals for all Covariates in Model



Ireland’s two major commercial centres, and also a growing British market, fared worse compared to parishes in closer proximity. These results are echoed when illiteracy is the dependent variable. The model coefficients on the illiteracy are lower across the board, however this is expected as illiteracy is higher (73 per cent) than fourth class housing (36 per cent) so a one-unit change in any of the independent variables should have a bigger impact on the poor housing variable, even when they are proportionally equal.

Figure 6 also highlights the importance of population pressure. If anything these results point towards a larger population pressure than has been previously found. The median posterior draw for the population pressure coefficient on illiteracy and poor housing are 3.1 and 4.2 respectively. If we translate these effects into our “no 19th-century population growth counterfactual” we would expect illiteracy to fall by about 3 per cent ( $\frac{0.75 \times 3.1}{73}$ ) and fourth class housing by a more sizeable 9 per cent ( $\frac{0.75 \times 4.2}{36}$ ). Whilst once again we find population pressure to be an important determinant of poverty, the counterfactual exercise cautions against overemphasising its importance.

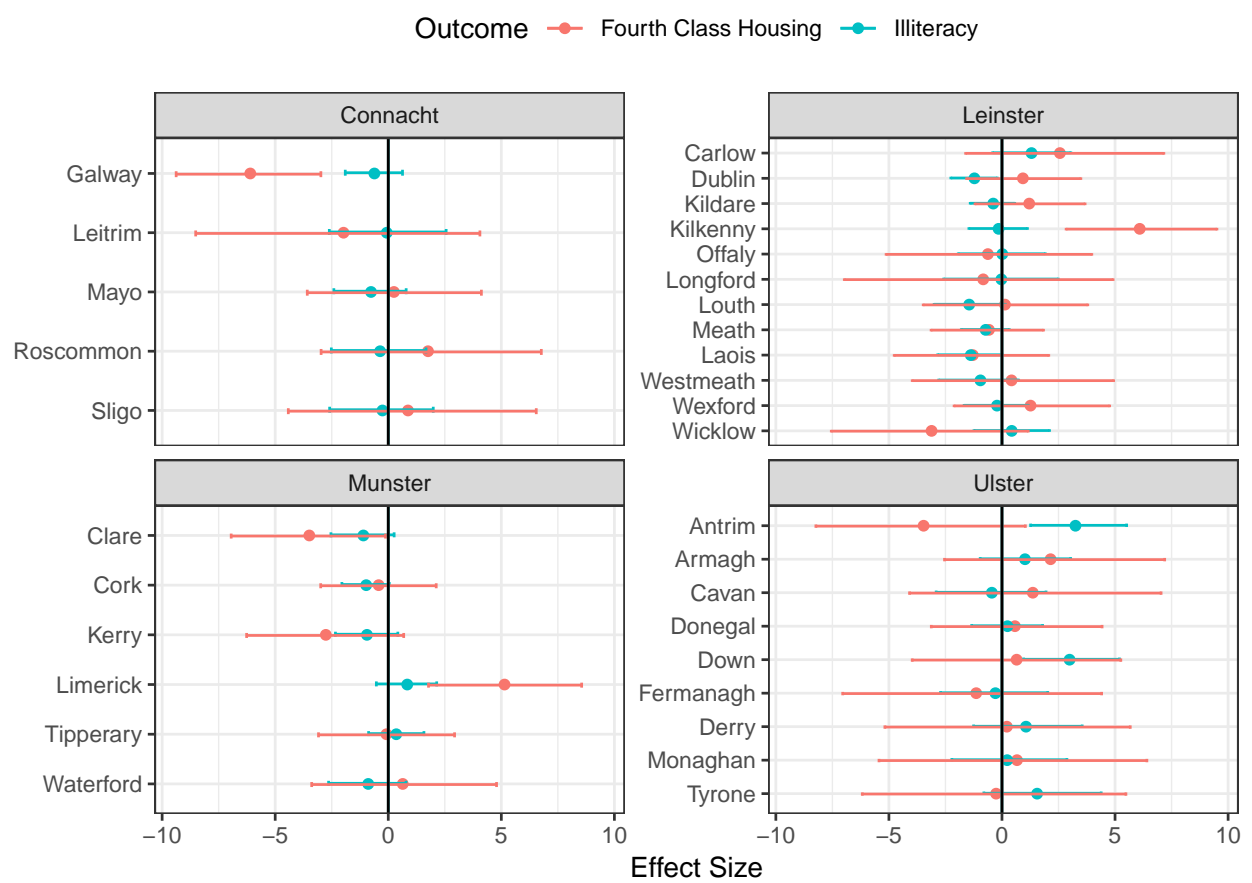
The multilevel structure of this model allows both the intercepts and slopes to vary by county.<sup>18</sup> In effect this allows each county to have a separate intercept and population pressure coefficient meaning we estimate an additional 64 model parameters. Thus, our model has two levels. The upper-level parameters are displayed in Figure 6. These upper-level parameters include both an intercept and population pressure slope. However, we know that these terms are not fixed as they vary by county. Figure 7 summarizes the posterior draws for  $\beta_{County}$ . All of these draws already account for the upper-level effects of 3.1 and 4.2 for illiteracy and fourth class housing respectively, so all of the values in Figure 9 are centred around the null (as indicated by the vertical black line).

Figure 9 is segmented by Ireland’s four provinces. Overall, it does not support the idea of substantial heterogeneity in the population pressure parameter. The 95 per cent credible interval of every county-varying illiteracy slope straddles the vertical black line indicating the null. Some counties’ slopes, such as those of Galway, Kilkenny, Limerick, and Antrim do seem to differ from the null, although this is probably due to random chance as there does not appear to be any systematic reason for these differences. It appears that once all of the additional explanatory variables and the spatial filtering are

---

<sup>18</sup>Again, please see Figure 9 in Appendix A for a guide on the geographical positioning of Ireland’s counties and provinces.

Figure 7: Varying Slopes by Province and County, Median Effects alongside 95 Per Cent Credible Interval





factored into the modelling procedure, few further elements influenced poverty. This view is emphasized by an inspection of the county-varying intercepts in both equations. The median posterior intercept draws for illiteracy and poor housing were 73 and 36 respectively, and the standard deviation of the county-varying intercepts was only 0.2 in both instances. Despite large spatial differences in both variables (as seen in Figure 2) only a fraction of these differences manifest themselves in the county-varying parameters examined in our multilevel analysis.<sup>19</sup>

The multilevel model also estimates the correlation between the random intercepts and slopes in each equation. This correlation would be positive if, for example, the population pressure effect was greater for counties with higher levels of illiteracy/fourth-class housing. However, our results also fail to indicate that such an effect is at play here. For fourth class housing the median intercept-slope correlation coefficient is 0.01. The equivalent value in the illiteracy equation is  $-0.46$ , indicating that the population pressure variable is stronger in parishes with lower illiteracy although the wide 95 per cent credible interval of 0.35 to  $-0.93$  cautions against any strong conclusion. One of the motivations for estimating this model using a multivariate structure is to allow the error terms from the two linear equations to be correlated. In Figure 5 we found a strong correlation ( $\rho = 0.5$ ) between the two variables. The median residual term correlation coefficient for the posterior draws from the multivariate model is considerably lower at 0.19 (0.15–0.23, 95 per cent C.I.). This difference emphasizes the importance of our explanatory variables and how little of the variation can be explained once they are included in the model.

The results in this subsection do not support the hypothesis that the population pressure variable was heterogeneous across Ireland’s 32 counties. One obvious avenue for further investigation is nonlinearity. Thus far we have assumed that all of the explanatory variables have a linear effect on the respective poverty outcomes. Figures 10 and 11 in Appendix D relax the linear restriction using a Generalised Additive Modelling approach (Wood, 2017). Both figures show that the population pressure effect is approximately linear. This pattern is common across most of explanatory variables, although a number of the distance-related variables, such as distance to city and longitude, do appear to kink but only at the extreme

---

<sup>19</sup>We also experimented with running the above multilevel model but without applying the spatial filtering procedure, and obtained similar results. Thus, the homogeneous intercept terms are not an artefact of spatial filtering.

ends of their distributions.

## 4.4 Fertility

Thus far our analysis has evaluated population's role in explaining poverty. In this subsection we reverse the question and ask what influenced population growth? To do this we regress fertility (the child-family ratio) on various explanatory variables. Our focus on fertility is justified on the basis that the rapid 18th century growth in the Irish population was caused by increased marital fertility (Connell, 1950). Mokyr (1985, ch.3) also addressed the issue of fertility and population growth, but failed to find any covariates that consistently predicted either variable. Panel (a) in Figure 8 displays the spatial distribution of our measure of marital fertility. The East-West difference in our other maps is also apparent here. Panel (b) is a scatterplot showing the bivariate relationship connecting fertility to population growth (1821–41). The (imperfect) correlation here of 0.23 is in line with expectations: parishes with higher fertility grew by more in the preceding twenty years. Was fertility, and thus population growth, randomly distributed and not systematically related to any other factor, as Mokyr's analysis suggests, or do the improved data used here offer an alternative perspective?

Figure 8: Child-Family Ratio 1841 and Population Growth 1821–41

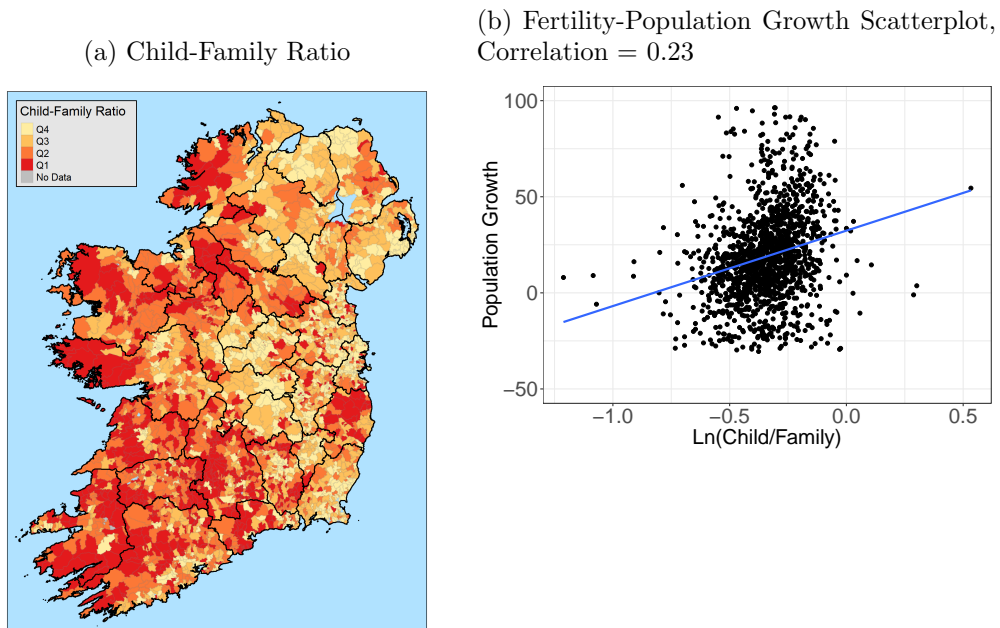


Table 5: Parish-Level Fertility OLS Regressions

	(1)	(2)
Intercept	−0.339*** (0.003)	−0.339*** (0.003)
Population Pressure	0.025*** (0.005)	0.023*** (0.006)
Coast	0.002 (0.004)	0.002 (0.004)
Occupation: Manufactures/Trade (Per Cent)	−0.029*** (0.006)	−0.028*** (0.006)
Occupation: Other(Per Cent)	−0.027*** (0.005)	−0.026*** (0.005)
Catholic (Per Cent)	−0.006 (0.007)	−0.009 (0.007)
Parish Town Population (Per Cent)	−0.018*** (0.005)	−0.017*** (0.005)
Distance to Nearest Town (km)	0.012*** (0.004)	0.012*** (0.004)
Vested Means (Per Cent)	0.004 (0.004)	0.005 (0.004)
Distance to City	−0.075*** (0.028)	−0.076*** (0.028)
Distance to 17th Century Road	0.004 (0.005)	0.004 (0.005)
Sex Ratio	0.005 (0.004)	0.006 (0.004)
Latitude	−0.035*** (0.011)	−0.039*** (0.012)
Longitude	−0.059** (0.024)	−0.057** (0.024)
Illiteracy (Per Cent)		0.012** (0.006)
Fouth Class Housing (Per Cent)		−0.008* (0.004)
Observations	2,387	2,387
R <sup>2</sup>	0.831	0.832
Adjusted R <sup>2</sup>	0.830	0.830
F Statistic	431.125***	402.238***

*Notes:* Columns (1) and (2) regress logged fertility (child-family ratio) on the indicated covariates. All covariates are expressed in terms of  $Z$ -scores and thus the coefficients are comparable. Residual autocorrelation was addressed in all models by transforming the outcome and all regressors via the eigenvector spatial filtering method detailed in Murakami (2018). \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

In Table 5 we regress parish-level fertility on the complete set of other regressors. Columns (1) and (2) distinguish between models that exclude and include the illiteracy and poor housing variables. In both columns population pressure has a significant effect on fertility, where a one standard deviation increase in population pressure increases the child-family ratio by around 2.5 per cent. If parishes were “overpopulated”, then steps were not being taken to correct this. This result is underlined by the inclusion of fourth-class housing and illiteracy variables, which do not alter the population pressure coefficient in any meaningful way. If population pressure was being corrected in response to poverty, then one would expect the inclusion of the poverty variables to attenuate the coefficient associated with population pressure.

The illiteracy and fourth class housing variables appear to have opposing effects in column (2). A one-SD increase in illiteracy is associated with a 1.2 per cent increase in fertility, whilst the equivalent one-SD movement in fourth-class housing reduces fertility by 0.8 per cent. The estimated coefficients for all other regressors are similar in both specifications. Rural parishes that were further away from market towns also had higher levels of fertility. However, this is contrasted with the large negative effect corresponding to the “Distance to City” variable, as parishes further from Belfast or Dublin experienced lower levels of fertility (although this effect is conditional on holding all other covariates such as longitude constant). It appears that a weak negative relationship between living standards and fertility existed and that this relationship was mediated by geographic and cultural factors.

Overall, Table 5 provides an interesting overview of Irish fertility prior to the famine. It appears that Mokyr’s analysis of population change and fertility relied on data that was too coarse to detect influential regressors. These results indicate that if poverty was underpinned by population pressure, this pressure had only a limited affect on fertility. A negative income-fertility gradient suggested by these regressions is also at odds with the conventional “preventive check” mechanism in Malthusian models. These conclusions are tentative, however, as our analysis is somewhat hampered by our data leading to an inability to disentangle economic from geographic, cultural or other social factors.

## 5 Conclusion

In the wake of the Great Irish Famine land agent William Steuart Trench informed his friend, the economist Nassau Senior (Senior, 1868):

It was an awful remedy. The country wore a delusive appearance of prosperity. Capital had been accumulating — rents had risen, and were well paid ... the value of property was increasing; but all this time the population was increasing more rapidly than the capital that was to maintain and employ it. ... Such were its numbers that it seemed irrevocably doomed to the potato... Nothing but the successive failures of the potato, its failure season after season, could have produced the emigration which will, I trust, give us room to become civilised.

That stark Malthusian interpretation of Irish backwardness on the eve of the Great Famine has often been re-echoed since, in Ireland and further afield (e.g. O'Brien (1921); Solow (1971, p. 196); Galbraith (1977, pp. 37–8); Caldwell (1998)), and it was the conventional wisdom when contested by Joel Mokyr in the early 1980s. Mokyr interpreted the outcome of his own research as casting ‘serious doubt on the simple and easy explanation that blames Irish poverty on excess population’. His work elicited widespread reaction and further econometric analyses (Mokyr, 1985; McGregor, 1989; Solar, 1989, 2017; O'Rourke, 1991; Fotheringham et al., 2013; Goodspeed, 2016; Kelly and Ó Gráda, 2015). In this paper we addressed the issue anew, using new data and new variables. The link between population pressure and living standards on the eve of the Famine is reasserted, although its quantitative impact turns out to be modest. We also provide evidence for a role for factors that might be interpreted as reflecting geography and human agency or institutions. First, we show that physical distance mattered: trying to eke out a living on an acre of any given quality of land was much tougher in the west of the country than in the east. Living in a remote area presumably meant not just distance from commodity markets, but also from government services, educational facilities, and exposure to new techniques and ideas. It also made the presence of a local elite, in the form of landed proprietors and professionals, less likely. Second, we show that the local presence of such a ‘leisure class’—an elite dependent on ‘vested means’ in the form of accumulated or inherited wealth—was associated with lower levels of poverty. It seems plausible to assume that such an elite could have added to

social capital, provided non-agricultural employment, or subsidized emigration. Absentee landlords could have been less likely to have provided these services. Third, we show that the effect of population per adjusted acre was (conditionally) homogenous and robust to the inclusion of a large number of control variables and a battery of econometric procedures. That this effect persists in the presence of Poor Law Union fixed effects is intriguing. It appears that the population pressure effect can be detected at a local level, proving both the importance and value of using highly disaggregated parish data. Finally, we find that fertility was not restricted in response to economic conditions. Our analysis shows that fertility was highest in rural areas further from the north and east coasts, as geographical explanatory factors appear to supersede economic ones. The limited role that economic variables play, alongside a result showing a strong positive conditional relationship between fertility and population pressure, demonstrate the absence of a preventive-check correction mechanism. Whilst Irish population pressure influenced a poverty, population growth was not slowing in response to Ireland's economic backwardness and there is no evidence that suggests this would have changed if not for the famine.

## Bibliography

- G. Alfani and C. Ó Gráda. The Timing and Causes of Famine: The European Experience. *Nature Sustainability*, 1(6):283–88, 2018.
- E. L. Almquist. Pre-Famine Ireland and the Theory of European Proto-Industrialization: Evidence from the 1841 Census. *Journal of Economic History*, 39(3):699–718, 1979.
- Q. Ashraf and O. Galor. Dynamics and Stagnation in the Malthusian Epoch. *American Economic Review*, 101(5):2003–41, 2011.
- R. D. C. Black. *Economic Thought and the Irish Question, 1817–1870*. Cambridge University Press, Cambridge, 1960.
- P. M. A. Bourke. The Extent of the Potato Crop in Ireland at the Time of the Famine. *Journal of the Statistical and Social Inquiry Society of Ireland*, 113, 1959/60.
- P. Boyle and C. Ó Gráda. Fertility Trends, Excess Mortality, and the Great Irish Famine. *Demography*, 23(4):543–62, 1986.

- Great Britain. *Select Committee on Emigration from United Kingdom Third Report, Minutes of Evidence, Appendix*, volume V. 1826–1827.
- Great Britain. *Report of the Commissioners Appointed to Take the Census of Ireland for 1841*, volume XXIV. 1843.
- Great Britain. *Select Committee on Townland Valuation of Ireland: Report Together with the Minutes of Evidence, Appendix and Index*, volume VII. 1844.
- Great Britain. *The Census of Ireland for the Year 1851, Part VI: General Report*, volume XXXI. 1856.
- P.-C. Bürkner. brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1):1–28, 2017. doi: 10.18637/jss.v080.i01.
- J. C. Brown and T. W. Guinnane. Regions and Time in the European Fertility Transition: Problems in the Princeton Project’s Statistical Methodology. *Economic History Review*, 60(3):574–95, 2007.
- J. C. Caldwell. Malthus and the Less Developed World: The Pivotal Role of India. *Population and Development Review*, 24(4):675–96, 1998.
- G. Clark. *A Farewell to Alms : A Brief Economic History of the World*. Princeton University Press, Princeton, 2007.
- K. H. Connell. *The Population of Ireland, 1750–1845*. Clarendon Press, Oxford, 1950.
- N. F. R. Crafts and T. C. Mills. From Malthus to Solow: How did the Malthusian Economy Really Evolve? *Journal of Macroeconomics*, 31(1):68–93, 2009.
- S. Daultrey, D. Dickson, and C. Ó Gráda. Hearth Tax, Household Size and Irish Population Change 1672–1821. *Proceedings of the Royal Irish Academy*, 82(6):125–81, 1982.
- G. de Beaumont. *Ireland Social Political and Religious*, volume 1–2. Bentley, London, 1839.
- A. Eiríksson and C. Ó Gráda. *Estate Records of the Irish Famine*. Irish Famine Network, Dublin, 1995.

- J. Eötvös. *Poverty in Ireland, 1837: A Hungarian's View*. Phaeton, Dublin, 2015. Original work published 1837.
- A. Fernihough. Malthusian Dynamics in a Diverging Europe: Northern Italy, 1650–1881. *Demography*, 50(1):311–32, 2013.
- A. Fernihough, C. Ó Gráda, and B. M. Walsh. Intermarriage in a divided society: Ireland a century ago. *Explorations in Economic History*, 56:1–14, 2015.
- A. S. Fotheringham, M. H. Kelly, and M. Charlton. The Demographic Impacts of the Irish Famine: Towards a Greater Geographical Understanding. *Transactions of the Institute of British Geographers*, 38(2):221–37, 2013.
- J. K. Galbraith. *The Age of Uncertainty*. Houghton Mifflin, Boston, 1977.
- O. Galor and D. N. Weil. Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond. *American Economic Review*, 90(4):806–28, 2000.
- A. Gelman and J. Hill. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, Cambridge, MA, 2006.
- T. B. Goodspeed. Microcredit and Adjustment to Environmental Shock: Evidence from the Great Famine in Ireland. *Journal of Development Economics*, 121:258–77, 2016.
- P. Gray. The Irish Poor Law and the Great Famine. XIV International Economic History Congress, Helsinki, 2006.
- P. Gray. *The Making of the Irish Poor Law, 1815–43*. Manchester University Press, Manchester, 2009.
- I. N. Gregory and N. A. Cunningham. ‘The Judgement of God on an Indolent and Unself-Reliant People’?: The Impact of the Great Irish Famine on Ireland’s Religious Demography. *Journal of Historical Geography*, 51:76–87, 2016.
- G. Hansen and E. Prescott. Malthus to Solow. *American Economic Review*, 92(4):1205–17, 2000.



- H. Inglis. *A Journey Throughout Ireland in 1834*. Whittaker, London, 1835.
- M. Kelly and C. Ó Gráda. Why Ireland Starved after Three Decades: The Great Famine in Cross-Section Reconsidered. *Irish Economic and Social History*, 42(1):53–61, 2015.
- M. Kelly and C. Ó Gráda. The Preventive Check in Medieval and Preindustrial England. *The Journal of Economic History*, 72(4):1015–35, 2012.
- M. Kelly and C. Ó Gráda. Living Standards and Mortality Since the Middle Ages. *The Economic History Review*, 67(2):358–81, 2014.
- J. G. Kohl. *Travels in Ireland*. Bruce and Wyld, London, 1844.
- N.-P. Lagerlöf. Malthus in Sweden. *The Scandinavian Journal of Economics*, 117(4):1091–133, 2015.
- J. Lee. Marriage and Population in Pre-Famine Ireland. *Economic History Review*, 21(2):283–95, 1968.
- R. D. Lee. Short-term Variation: Vital Rates, Prices and Weather. In E. A. Wrigley and R. S. Schofield, editors, *The Population History of England, 1541–1871. A Reconstruction*, pages 356–401. Edward Arnold, London, 1981.
- P. P. L. McGregor. Demographic Pressure and the Irish Famine: Malthus after Mokyr. *Land Economics*, 65(3):228–38, 1989.
- J. Mokyr. The Deadly Fungus: An Econometric Investigation into the Short-Term Demographic Impact of the Irish Famine, 1846–51. *Research in Population Economics*, 2:237–77, 1980.
- J. Mokyr. *Why Ireland Starved: A Quantitative and Analytical History of the Irish Economy, 1800–1850*. Allen & Unwin, London, second edition, 1985.
- D. Murakami and D. A. Griffith. Eigenvector Spatial Filtering for Large Data Sets: Fixed and Random Effects Approaches. *Geographical Analysis*, 2018. doi: 10.1111/gean.12156.
- E. Nicolini. Was Malthus Right? A VAR Analysis of Economic and Demographic Interactions in Pre-Industrial England. *European Review of Economic History*, 11(1):99–121, 2007.

- C. Ó Gráda. *Ireland: A New Economic History 1780–1939*. Oxford University Press, Oxford, 1994.
- C. Ó Gráda. *Black ‘47 and Beyond: the Great Irish Famine in History*. Princeton University Press, Princeton, 1999.
- C. Ó Gráda, R. Paping, and E. Vanhaute. *When the Potato Failed: Causes and Effects of the Last European Subsistence Crisis, 1845–1850*. CORN Publication Series: Comparative Rural History of the North Sea Area. Isd, 2007. ISBN 9782503519852.
- G. O’Brien. *The Economic History of Ireland between the Union and the Famine*. Longmans, Green, London, 1921.
- K. H. O’Rourke. Did the Great Irish Famine Matter? *Journal of Economic History*, 51(1):1–22, 1991.
- K. H. O’Rourke. Emigration and Living Standards in Ireland since the Famine. *Journal of Population Economics*, 8(4):407–21, 1995.
- E. J. Pebesma. Multivariable Geostatistics in S: The gstat Package. *Computers & Geosciences*, 30:683–91, 2004.
- U. Pfister and G. Fertig. The Population History of Germany: Research Strategy and Preliminary Results. MPIDR Working Papers WP-2010-035, Max Planck Institute for Demographic Research, Rostock, Germany, 2010.
- S. J. Redding and D. M. Sturm. The Costs of Remoteness: Evidence from German Division and Reunification. *American Economic Review*, 98(5):1766–97, 2008.
- J. Reis. Economic Growth, Human Capital Formation and Consumption in Western Europe before 1800. In R. C. Allen, T. Bengtsson, and M. Dribe, editors, *Living Standards in the Past: New Perspectives on Wellbeing in Asia and Europe*, pages 195–225. Oxford University Press, Oxford, 2005.
- N. W. Senior. *Journals, Essays, and Conversations Relating to Ireland*. 2 Vols. Longmans, Green, London, 1868.

- P. M. Solar. The Great Irish Famine was no Ordinary Subsistence Crisis. In E. M. Crawford, editor, *Famine: The Irish Experience 900–1900*, pages 112–31. John Donald, Edinburgh, 1989.
- P. M. Solar. Occupation, Poverty and Social Class in Pre-Famine Ireland, 1740–1850. In E. F. Biagini and M. E. Daly, editors, *The Cambridge Social History of Modern Ireland*, pages 25–37. Cambridge University Press, Cambridge, 2017.
- B. L. Solow. *The Land Question and the Irish Economy 1870–1903*. Harvard University Press, Cambridge, 1971.
- J. H. Stock and M. Yogo. Testing for Weak Instruments in Linear IV Regression. In James Stock and D.W.K. Andrews, editors, *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg*, pages 80–108. Cambridge University Press, New York, 2005.
- W. M. Thackeray. Irish Sketchbook. In *The Works of William Makepeace Thackeray*, volume XV. Scribner, New York, 1911. Original work published 1842.
- G. S. L. Tucker. Irish Fertility Ratios before the Famine. *Economic History Review*, 23(2): 267–84, 1970.
- M. Verpoorten. Leave None to Claim the Land: A Malthusian Catastrophe in Rwanda. *Journal of Peace Research*, 49(4):547–63, 2012.
- E. Wakefield. *An Account of Ireland, Statistical and Political. 2 vols.* Longman, London, 1812.
- S. A. Wegge, T. Anbinder, and C. Ó Gráda. Immigrants and Savers: A Rich New Database on the Irish in 1850s New York. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 50(3):144–55, 2017.
- D. R. Weir. Life Under Pressure: France and England 1670–1870. *Journal of Economic History*, 44(1):27–47, 1984.
- M. F. J. Wilson, B. O’Connell, J. C. Guinan C. Brown, and A. J. Grehan. Multiscale Terrain Analysis of Multibeam Bathymetry Data for Habitat Mapping on the Continental Slope. *Marine Geodesy*, 30(1–2):3–35, 2007.

S. N. Wood. *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC, 2 edition, 2017.

## A Ireland's Administrative Borders

Figure 9: Ireland's Administrative Borders

(a) Counties



(b) Provinces



## B Male and Female Illiteracy

Table 6: OLS Regression of Male Illiteracy (Per Cent) on Population Pressure

	(1)	(2)	(3)	(4)	(5)
Intercept	63.738*** (0.162)	63.738*** (0.128)	67.915*** (3.678)	63.738*** (0.128)	64.177*** (0.152)
Population Pressure	3.690*** (0.220)	2.807*** (0.215)	2.806*** (0.224)	2.733*** (0.215)	2.910*** (0.274)
Fertility				0.489*** (0.146)	0.313* (0.189)
Migration				-0.057 (0.149)	-0.023 (0.180)
Pop. Growth 1821–41					0.432*** (0.164)
Coast		0.347** (0.161)	0.450*** (0.173)	0.341** (0.161)	0.558*** (0.190)
Occupation: Manufactures/Trade (Per Cent)		-2.034*** (0.253)	-2.297*** (0.252)	-1.946*** (0.255)	-1.891*** (0.311)
Occupation: Other(Per Cent)		-0.651*** (0.184)	-0.519*** (0.176)	-0.576*** (0.185)	-0.895*** (0.220)
Catholic (Per Cent)		2.935*** (0.264)	2.239*** (0.279)	2.953*** (0.264)	3.143*** (0.302)
Parish Town Population (Per Cent)		-0.899*** (0.218)	-0.690*** (0.212)	-0.838*** (0.220)	-1.016*** (0.266)
Distance to Nearest Town (km)		0.262* (0.152)	0.591*** (0.172)	0.227 (0.152)	0.264 (0.183)
Vested Means (Per Cent)		-1.359*** (0.164)	-1.548*** (0.161)	-1.369*** (0.164)	-0.886*** (0.193)
Distance to City		4.604*** (1.126)	2.300 (2.402)	4.798*** (1.126)	3.133** (1.370)
Distance to 17th Century Road		0.279 (0.194)	0.548* (0.321)	0.268 (0.194)	0.079 (0.254)
Sex Ratio		-0.261* (0.153)	-0.166 (0.147)	-0.274* (0.152)	-0.574*** (0.185)
Latitude		2.537*** (0.460)	-2.859* (1.729)	2.643*** (0.460)	2.431*** (0.541)
Longitude		-0.876 (0.976)	-2.630 (2.013)	-0.732 (0.976)	-1.842 (1.149)
PLU FE	N	N	Y	N	N
Observations	2,387	2,387	2,387	2,387	1,519
R <sup>2</sup>	0.481	0.677	0.749	0.679	0.711
Adjusted R <sup>2</sup>	0.478	0.673	0.726	0.675	0.705
F Statistic	146.426***	183.174***	33.666***	171.617***	121.734***

Notes: All columns regress the male illiteracy (per cent) variable on the indicated covariates. All covariates are expressed in terms of  $Z$ -scores and thus the coefficients are comparable. Residual autocorrelation was addressed in all models by transforming the outcome and all regressors via the eigenvector spatial filtering method detailed in Murakami (2018). \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

Table 7: OLS Regression of Female Illiteracy (Per Cent) on Population Pressure

	(1)	(2)	(3)	(4)	(5)
Intercept	82.091*** (0.135)	82.091*** (0.103)	84.047*** (3.050)	82.091*** (0.103)	82.558*** (0.120)
Population Pressure	3.091*** (0.183)	2.641*** (0.173)	2.741*** (0.186)	2.639*** (0.174)	2.931*** (0.217)
Fertility				-0.038 (0.118)	0.006 (0.149)
Migration				-0.154 (0.120)	-0.157 (0.143)
Pop. Growth 1821-41					0.367*** (0.130)
Coast		-0.426*** (0.130)	-0.203 (0.143)	-0.429*** (0.130)	-0.404*** (0.150)
Occupation: Manufactures/Trade (Per Cent)		-0.901*** (0.204)	-1.076*** (0.209)	-0.886*** (0.205)	-0.956*** (0.246)
Occupation: Other(Per Cent)		-0.799*** (0.148)	-0.721*** (0.146)	-0.806*** (0.149)	-0.991*** (0.174)
Catholic (Per Cent)		2.425*** (0.213)	1.773*** (0.231)	2.426*** (0.213)	2.687*** (0.239)
Parish Town Population (Per Cent)		-1.000*** (0.175)	-0.866*** (0.176)	-0.972*** (0.177)	-1.192*** (0.211)
Distance to Nearest Town (km)		0.466*** (0.123)	0.510*** (0.143)	0.464*** (0.123)	0.494*** (0.145)
Vested Means (Per Cent)		-1.764*** (0.132)	-1.809*** (0.133)	-1.762*** (0.132)	-1.089*** (0.153)
Distance to City		3.142*** (0.906)	3.553* (1.992)	3.089*** (0.908)	2.015* (1.084)
Encumbered Estates (Per Cent)		-0.387** (0.156)	0.289 (0.266)	-0.386** (0.156)	-0.594*** (0.201)
Distance to 17th Century Road		-0.088 (0.123)	0.097 (0.122)	-0.087 (0.123)	-0.437*** (0.146)
Sex Ratio		2.862*** (0.370)	-0.333 (1.434)	2.878*** (0.371)	2.622*** (0.428)
Latitude		-0.877 (0.785)	0.038 (1.669)	-0.932 (0.787)	-1.731* (0.909)
PLU FE	N	N	Y	N	N
Observations	2,387	2,387	2,387	2,387	1,519
R <sup>2</sup>	0.499	0.708	0.758	0.708	0.743
Adjusted R <sup>2</sup>	0.496	0.704	0.737	0.704	0.738
F Statistic	157.483***	211.386***	35.434***	196.845***	143.183***

Notes: All columns regress the Female illiteracy (per cent) variable on the indicated covariates. All covariates are expressed in terms of  $Z$ -scores and thus the coefficients are comparable. Residual autocorrelation was addressed in all models by transforming the outcome and all regressors via the eigenvector spatial filtering method detailed in Murakami (2018). \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

## C Multilevel Model Structure and Priors

$$\begin{bmatrix} \text{Illiteracy}_i \\ \text{Poor Housing}_i \end{bmatrix} \sim \text{MVNormal} \left( \begin{bmatrix} \mu_{1,i} \\ \mu_{2,i} \end{bmatrix}, \mathbf{S}_1 \right)$$

$$\mathbf{S}_1 \sim \begin{bmatrix} \sigma_{\mu_1} & 0 \\ 0 & \sigma_{\mu_2} \end{bmatrix} \mathbf{R}_1 \begin{bmatrix} \sigma_{\mu_1} & 0 \\ 0 & \sigma_{\mu_2} \end{bmatrix}$$

$$\mathbf{R}_1 = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix}$$

$$\mu_{1,i} = \alpha_{1,\text{County}} + \beta_{1,\text{County}} \ln(\text{PopPres}_i) + \mathbf{X}\gamma_1$$

$$\mu_{2,i} = \alpha_{2,\text{County}} + \beta_{2,\text{County}} \ln(\text{PopPres}_i) + \mathbf{X}\gamma_2$$

$$\begin{bmatrix} \beta_{1,\text{County}} \\ \beta_{2,\text{County}} \end{bmatrix} \sim \text{MVNormal} \left( \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}, \mathbf{S}_2 \right)$$

$$\mathbf{S}_2 \sim \begin{bmatrix} \sigma_{\beta_1} & 0 \\ 0 & \sigma_{\beta_2} \end{bmatrix} \mathbf{R}_2 \begin{bmatrix} \sigma_{\beta_1} & 0 \\ 0 & \sigma_{\beta_2} \end{bmatrix}$$

$$\mathbf{R}_2 = \begin{bmatrix} 1 & \rho_2 \\ \rho_2 & 1 \end{bmatrix}$$

$$\begin{bmatrix} \alpha_{1,\text{County}} \\ \alpha_{2,\text{County}} \end{bmatrix} \sim \text{MVNormal} \left( \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}, \mathbf{S}_3 \right)$$

$$\mathbf{S}_3 \sim \begin{bmatrix} \sigma_{\alpha_1} & 0 \\ 0 & \sigma_{\alpha_2} \end{bmatrix} \mathbf{R}_3 \begin{bmatrix} \sigma_{\alpha_1} & 0 \\ 0 & \sigma_{\alpha_2} \end{bmatrix}$$

$$\mathbf{R}_3 = \begin{bmatrix} 1 & \rho_3 \\ \rho_3 & 1 \end{bmatrix}$$

$$\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \dots, \gamma_{12} \sim \text{Normal}(0, 10)$$

$$\sigma_{\mu_1}, \sigma_{\mu_2}, \sigma_{\alpha_1}, \sigma_{\alpha_2}, \sigma_{\beta_1}, \sigma_{\beta_2} \sim \text{HalfCauchy}(0, 2)$$

$$\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3 \sim \text{LKJcorr}(2)$$



## D Generalized Additive Model Result Plots

We model nonlinearity with a Generalised Additive Model (GAM) approach (Wood, 2017).

Thus, our two modelling equations, which we estimate separately, are:

$$\text{Illiteracy}_i = \alpha_1 + f_1[\ln(\text{PopPres}_i)] + \left( \sum_{k=2}^{14} f_k[X_k] \right) + \epsilon_i \quad (7)$$

$$\text{Poor Housing}_i = \alpha_2 + g_1[\ln(\text{PopPres}_i)] + \left( \sum_{k=2}^{14} g_k[X_k] \right) + v_i \quad (8)$$

where the 14 explanatory variables now effect the outcome through the various  $f$  and  $g$  functions. In this application the  $f$  and  $g$  functions are penalized regression splines which we estimate. The value  $X_k$  represents the  $k$ -th column of the matrix of additional independent variables. In this instance we use the same explanatory variables listed in Tables 2 and 3 and the results of the GAM model are illustrated in Figures 10 and 11.

Figure 10: Illiteracy Housing GAM Results

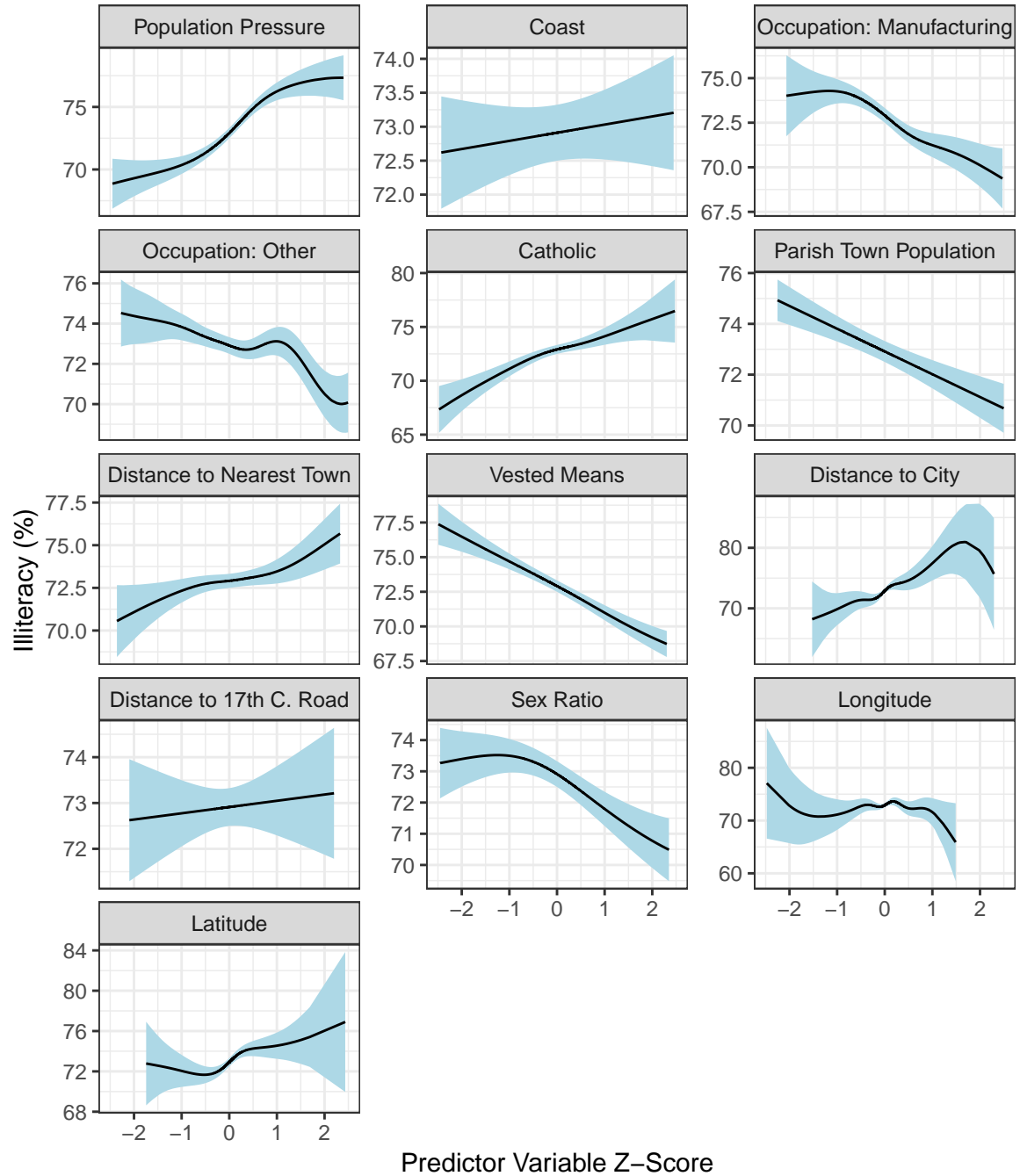


Figure 11: Fourth Class Housing GAM Results

