HISTORICAL METHODS, Fall 2005, Volume 38, Number 4

Analyzing Spatiotemporal Change by Use of National Historical Geographical Information Systems

Population Change during and after the Great Irish Famine

IAN N. GREGORY
PAUL S. ELL
Centre for Data Digitisation and Analysis
Queen's University Belfast

Abstract. Several countries have made large investments in building historical Geographical Information Systems (GIS) databases containing census and other quantitative statistics over long periods of time. Making good use of these databases requires approaches that explore spatial and temporal change. The authors use a variety of visualization and spatial analysis techniques to explore population change in Ireland during and after the Great Famine of the late 1840s. Importantly, the techniques allow differences over space and time to be explored, thus stressing the diversity between places, rather than making all places appear the same, a common criticism of many statistical approaches. The authors demonstrate the potential of these techniques to explore geographical and temporal variations in large quantitative GIS datasets.

Keywords: census, Geographical Information Systems (GIS), Great Irish Famine, spatial analysis

The past few years have seen a rapid rise in interest in the potential use of Geographical Information Systems (GIS) for historical research. The field has become known as historical GIS. Its evolution is evident in a number of ways, including the publication of books on the subject (Gregory 2003; Knowles 2002); special editions of journals such as *Social Science History* (Knowles 2000) and *History and Computing* (Ell and Gregory 2001); the organization of a conference, *History and Geography: The Role of Geographical Information in Historical Scholarship*, specifically on the subject;[1] and numerous conference sessions and presentations, for example, the Social Science History Association,[2] the Association of American Geographers,[3] and the Royal Geographical Society with the Institute of British Geographers[4] annual conferences.

Scholars working on GIS projects have been particularly effective in raising funding to create large national historical GISs containing census and related data. The Great Britain Historical GIS (GBHGIS; Gregory 2005; Gregory et al. 2002), the U.S. National Historical GIS (NHGIS; McMaster and Noble 2005),[5] the China Historical GIS (Bol and Ge 2005),[6] and the Historical GIS of the Belgian Historical Structure (De Moor and Wiedemann 2001; Vanhaute 2005) are among the best-developed examples, and a variety of others have also been proposed. Summaries of the leading national historical GISs are provided by Ian Gregory (2002a) and Anne Knowles (2005). With the exception of the China Historical GIS, which is more broadly focused, these systems are basically databases that contain long runs of nineteenth- and twentieth-century census and related statistics linked to polygons[7] that represent the administrative areas for which these data were published. They are expensive resources to build. The GBHGIS that covers the period from the early nineteenth century to the 1970s cost well over half a million pounds, mainly provided by the Economic and Social Research Council (ESRC); the NHGIS is funded by a $5 million grant from the National Science Foundation.

In spite of this investment, there is little clear idea of how these systems can best be used to enhance our understanding of the past. By the nature of the data that they contain, any analysis using these systems is likely to emphasize changing human geographies as recorded in quantitative sources such as the census. Although this subject has received attention in the past (see, e.g., Darby et al. 1979; Lawton 1968; Lee 1991), studies in this area have always been hampered by the complexities of analyzing change over space and time in the face of changing administrative boundaries.

In this article, we report on a pilot project that uses GIS and spatial statistical analysis techniques to explore geographical change in Ireland during and after the Great Famine, covering the period from 1841 to 1861. We aim to demonstrate both the utility and limitations of using GISs to explore changing patterns over time and space. The time period is deliberately short, as our aim is to focus on techniques and potential rather than results. However, it should be apparent that the techniques described here allow us to explore much longer periods of time, thus letting us make significant advances in the historiography of how places develop over time. It should also be noted that we focus on a single approach to illustrate the utility of the techniques used. In reality, these techniques should be used not in isolation but in conjunction with more conventional and perhaps nonquantitative historical approaches.

GISs and Historical Research

There are multiple approaches to defining a GIS (Chrisman 1999), but one common approach is based on the functionality that the software offers. A British government inquiry into the use of geographical information produced a widely quoted definition in which a GIS is defined as "a system for capturing, storing, checking, integrating, manipulating, analysing and displaying data which are spatially referenced to the Earth" (Department of the Environment 1987, 132). Other authors list slightly different characteristics, but it is widely agreed that the key abilities of the GIS are that it allows a database of spatially referenced data to be created and the data in it to be manipulated, integrated, analyzed, and displayed.

GIS software achieves this by combining conventional data, termed *attribute data*, which can be statistics, text, images, sound, or a wide variety of other formats, with a coordinate-based reference to the place where each item of data is located on the Earth's surface, known as *spatial data*. This combined structure makes the GIS unique because it is able to answer questions not only about *what* something is but also about *where* it is located.

The GIS is thus a database technology that allows the spatial component of data to be explicitly handled. Although the potential of the GIS to contribute to historical research has been widely acknowledged (Baker 2003; Holdsworth 2002, 2003), there has yet to be a clear identification of exactly how the GIS will fit into the discipline of history. Gregory, Karen Kemp, and Ruth Mostern (2001) argued that there are three ways in which the GIS can help the historian: (1) by using location to discover, manage, and integrate historical resources; (2) by using its mapping abilities to visualize data and the results of analyses; and (3) to perform explicitly spatial analyses of historical data. Most historians primarily associate the GIS with the second of these, visualizing spatially referenced historical information through the use of maps. Mapping is undoubtedly one of the key advantages of the GIS and is also one of the easiest, for once a GIS database has been created, maps can immediately be produced from it. Maps have enormous value in suggesting spatial patterns inherent within a dataset and, in some cases, between two or more datasets. They are, however, quite limited tools for geographical analyses, analogous perhaps to graphs in statistical analyses. Choropleth maps, the type of mapping most readily available from national historical GISs, stress spatial patterns but in so doing tend to overemphasize sparsely populated rural areas at the expense of dense urban ones, which may shrink to near invisibility. They usually focus on only a single variable that is heavily simplified, using around five arbitrarily selected class intervals, and they show only a single point in time or perhaps the change between two snapshots. If the GIS is to produce new knowledge, it has to be an integral part of a research process that stresses the importance of location and spatial relationships. The map goes only part of the way to achieving this.

Statistical analysis of quantitative data is, of course, well established. When spatially referenced data are used, spatial analysis techniques that use both the spatial and attribute components of the data may be used. Manfred Fischer, Henk Scholten, and David Unwin (1996, 5) defined spatial analysis in two ways. On one level, they said that it is concerned with "the use of quantitative (mainly statistical) procedures and techniques to analyse patterns of points, lines, areas and surfaces depicted on maps or defined by co-ordinates in two- or three-dimensional space." At a higher level, they broadened the definition to state that spatial analysis is an approach that places emphasis "on the indigenous features of geographical space, on spatial choice processes and on their implications for the spatio-temporal evolution of complex spatial systems." Spatial analysis as a subdiscipline precedes the GIS by many years (Gatrell 1985), but the clear synergies between the GIS and spatial analysis have reawakened interest in the use of spatial analysis within quantitative geography (Fotheringham, Brunsdon, and Charlton 2000; O'Sullivan and Unwin 2003). Spatial analysis is frequently criticized for being far better at describing patterns (the first part of the definition by Fischer, Scholten, and Unwin) than explaining the processes that cause them. Thus it is argued that spatial statistics leads to an overemphasis on geometry and forgets the importance of people (see, e.g., Sack 1980). There is undoubtedly truth in this; however, being able to first identify patterns and then investigate how they change over time does provide a way into understanding the processes that create these patterns. As an acknowledgment of this, much recent spatial analytical work in human geography has focused on using exploratory techniques that analyze data to suggest relationships and identify outliers, as opposed to confirmatory analysis that attempts to test hypotheses (Bailey and Gatrell 1995). There is still a further step, namely, interpreting how patterns have been formed. Clearly, technological and methodological innovation may help, but interpretation will always require good historical scholarship.

Time tends not to be explicitly incorporated into GIS software (Peuquet 1999), but strategies can be developed to handle it, particularly by comparing snapshots at different dates (Gregory 2003). Thus, the GIS gives us, in theory at least, a framework from which the researcher can explore the temporal, spatial, and attribute components of a research question in an integrated manner. Potentially, this framework enables us to overcome a traditional stumbling block that is far from unique to historical research, namely, that most approaches to a problem concentrate on one or at best two of these three components. Gail Langran and Chrisman (1988) argued that to accurately represent one of these three aspects it has been traditional to control a second and fix a third. They illustrate this with a variety of examples (see table 1). For example, with soils mapping, time is fixed because only one date is mapped; soil types are decided in advance, so theme is controlled; and the boundaries between soil types, the spatial component, is measured. Census data also fix time because the census is taken on a single date, but here space is controlled because the country is subdivided into arbitrary administrative units, such as counties, tracts, or enumeration districts, and the number of responses to particular questions, the attribute or theme, is measured. In actual fact, even theme is far from perfectly measured, as there are issues of nonresponse from certain groups, particularly the homeless, and the attributes published are only the answers to certain predefined questions for which the responses are collated into set categories that may or may not capture the issues in which the researcher is interested.

It has been argued for many years that both historical research and geographical research have to be able to explore change over both space and time (Baker 2003). John Langton (1972) argued that using systems theory would allow researchers to move from synchronic analysis, in which change over time is measured simply by comparing two snapshots assumed to be in equilibrium, to diachronic analysis, in which change could be explored over multiple times as a continuous and ongoing phenomenon. More recently, Doreen Massey (1999, 274) argued for incorporating both space and time into an analysis. She argued that time is important because it allows us to tell the story of how a place develops. Without including space, however, one can tell only one story and imply that the story was inevitable and will inevitably happen elsewhere. By incorporating time and space together, an analysis can tell multiple stories, and thus "space could be imagined as the sphere of the existence of multiplicity, of the possibility of the existence of difference. Such a space is the sphere in which distinct stories coexist, meet up, affect each other, come into conflict or cooperate. This space is not static, not a cross-section through time; it is disrupted, active and generative." Following such an approach, an analysis using the GIS to research change over time should attempt to identify the different journeys that places follow through time. That approach is arguably particularly important in quantitative work, where the GIS has been criticized for being, among other things, a return to "the very worst sort of positivism" (Taylor 1990, 211).

GISs and the Census

As previously discussed, national historical GISs are databases that hold spatially referenced data covering a long time period. To be used effectively, they must be able to provide information about spatiotemporal change. At the core of most of these systems are census statistics, quantitative data published at the aggregate level using discrete and arbitrary spatial units that are only available for limited snapshots, usually once every 10 years. As we noted, they are based on responses to a limited number of questions, whose results are often aggregated into categories that are also frequently arbitrary. In spite of these limitations, the census is unquestionably the best and most comprehensive source detailing how many countries' populations have developed over time. To date, however, the census has not been used particularly effectively in either historical or geographical research. The problem of intercensal boundary changes has meant that analyses using census data have usually either focused on the spatial detail of a single snapshot or have had to massively aggregate data to relatively stable units such as American states or British counties to

TABLE 1. Representation of Geographic Data in Various Formats

Format	Fixed	Controlled	Measured
Soils data	Time	Theme	Location
Topographic map	Time	Theme	Location
U.S. census data	Time	Location	Theme
Raster data	Time	Location	Theme
Weather reports	Location	Time	Theme
Flood tables	Location	Time	Theme
Tide tables	Theme	Location	Time
Airline schedules	Location	Theme	Time

Source. Langran and Chrisman (1988, 2).

look at long-term change. For reasons that will be subsequently described, aggregation is highly unsatisfactory and should be avoided. Instead, the use of GIS and spatial statistics opens up new possibilities for exploring change over time while preserving the spatial detail in the data.

The problem with using highly aggregate data is that, as data become increasingly aggregate, it becomes increasingly difficult to determine whether the patterns within them are a genuine representation of patterns that appear in the underlying population surface or are largely driven by the results of the boundaries used to subdivide the data. A simple example of this issue, known as the modifiable areal unit problem (MAUP; Openshaw 1984), is illustrated in figure 1. Figure 1a shows infant mortality rates in England and Wales in 1928, as taken from the *Registrar General's Annual Report*. The report used local government districts, which are represented on the map by use of 1,827 polygons. The average infant mortality rate is 57.3 deaths before the age of 1 year per 1,000 live births, and the standard deviation is 31.4. The lowest rate was zero and the maximum 400, although both of these may be a result of small numbers of births. The spatial pattern is very complex, with some evidence of high rates in urban areas and in coalfields and heavy industrial regions. High rates are also found in other areas, such as to the south of The Wash and on the Devon-Somerset border, both of which are very rural. The data aggregated to county level are shown in figure 1b. There are only 63 counties, with a mean death rate of 59.0, which is not noticeably different from the raw data. The standard deviation, however, is only 12.1, down from 31.4 in figure 1a, and the minimum and maximum are now 38.1 and 92.5, respectively, giving a range of 54.4, down from 400. The spatial pattern is now a very simple one that shows a clear and unambiguous north-south divide. This result is the impact of the MAUP: the pattern has become highly simplified and much harder to interpret in a meaningful way.

The MAUP not only simplifies patterns in a single dataset but also affects multivariate approaches to analysis. Stan Openshaw and Peter Taylor (1979) demonstrated this vividly with a correlation analysis, in which they showed that by aggregating data from 99 counties in Iowa to six regions in different ways they could find correlations of between –0.99 and 0.99 between the population over 65 and the population voting Republican in the 1968 congressional elections. Stewart Fotheringham and David Wong (1991) showed that similar problems exist with regression analysis. The key lesson is that the use of highly aggregate data should be avoid-

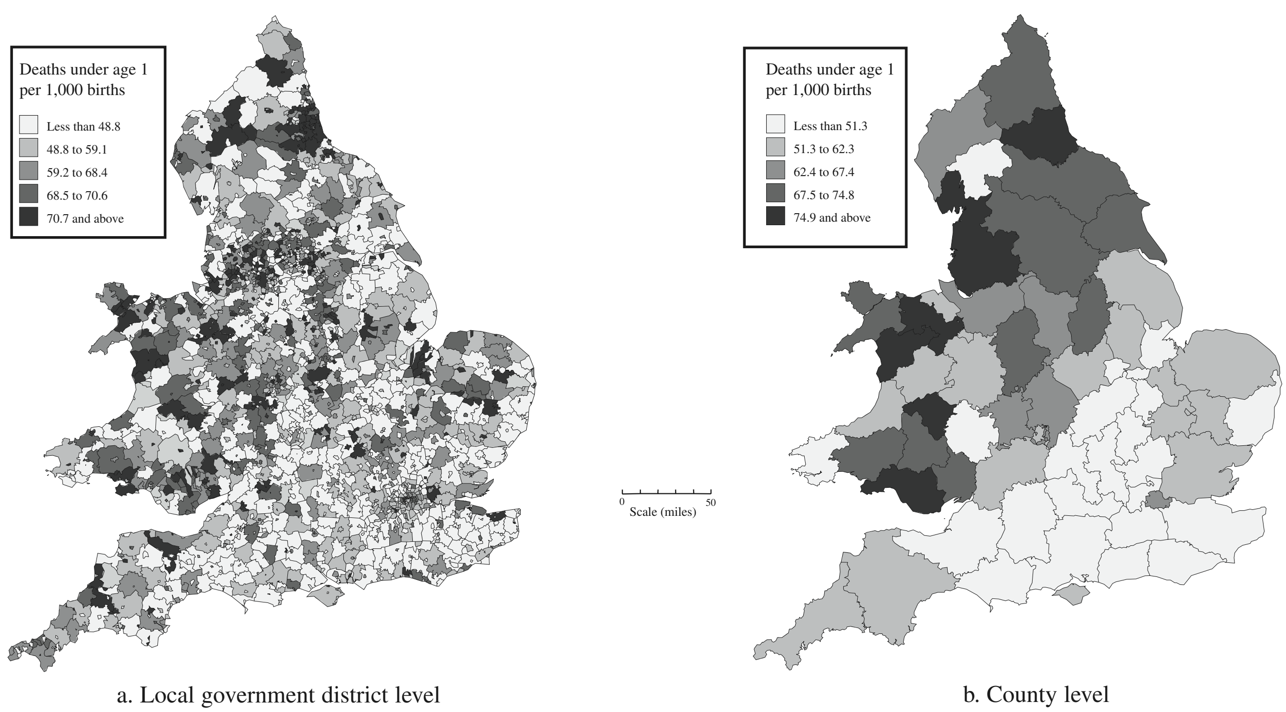


FIGURE 1. Infant mortality in 1928 at (a) local government district level and (b) county level. Each class interval contains approximately 20 percent of total population.

ed, as any patterns found are very difficult to interpret. Thus, if we are interested in analyzing change over space and time, we should attempt to do so with as much spatial detail as possible on the basis of the source material's limitations.

This has traditionally left researchers using census data with a dilemma: if the spatial detail is to be preserved, intercensal boundary changes mean that it is impossible to compare one census with the next; thus, it is only possible to look at one snapshot in time. To compare data over time, the data must be massively aggregated to a consistent form of administrative unit such as Irish or British counties or U.S. states, but this seriously degrades the data. What is required is the ability to explore change over time and space at as close as possible to the spatial scale at which the data were originally published. National historical GISs contain data in this form; however, there is still the problem of finding a way to compare them over time without the impact of the boundary changes' affecting the pattern and, hence, the analysis.

The GIS's ability to manipulate spatial data allows a solution to this problem to be implemented. Key to this is the ability to overlay two layers of data and merge both their spatial and attribute components (Flowerdew 1991). By overlaying two sets of administrative units, the GIS can use *areal interpolation* to estimate the population of a set of target zones, using the populations of a set of source zones. Doing this involves estimating how the population is distributed within the source zones (Gregory and Ell 2005).

The basic problem is illustrated in figure 2, where the values of a variable such as total population or infant mortality for three target units—Bristol, Chipping Sodbury, and Thornbury—are required. The difficulty is that these data were published for four source units. In addition to Bristol, Chipping Sodbury, and Thornbury, there is also Barton Regis, which was subsequently abolished as a result of boundary changes. Using a GIS overlay operation, we can measure the degree of overlap between Barton Regis and the three target units. The simplest method of then allocating data from Barton Regis is to use a process called *areal weighting*, in which it is assumed that the source data are evenly distributed within Barton Regis and data are allocated proportionally to area (Goodchild and Lam 1980). Thus, as approximately 50 percent of Barton Regis's area lies within the target zone of Bristol, so a corresponding proportion of its population will be allocated to Bristol; approximately 30 percent will be allocated to Thornbury and approximately 20 percent to Chipping Sodbury.

This is clearly an unsatisfactory approach, as Bristol is a major city whose center lies just north of the word "Bristol" in figure 2, whereas Thornbury and Chipping Sodbury are quite rural. It seems likely that the total population of Bristol will be underestimated and the values for the other two target districts will be overestimated. Other variables might pose alternative difficulties; for example, early-twentieth-century infant mortality has an urban bias, so this problem will be made worse, whereas variables with a strong rural bias (e.g., agricultural workers) will lead to the problem's being reversed, and too many such workers will be allocated to Bristol. The problem is caused by the underlying assumption of the technique, namely, that the population is evenly distributed across the source zones. A variety of strategies have been developed to relax this assumption, and it has been shown that these can significantly improve the accuracy of interpolated data (Gregory 2002b). Thus, with areal interpolation we can now create time series of data using consistent spatial units and thus explore change over time without losing spatial detail as a result of aggregation. Areal interpolation can also be used to compare data published for two different administrative geographies from the same dates.

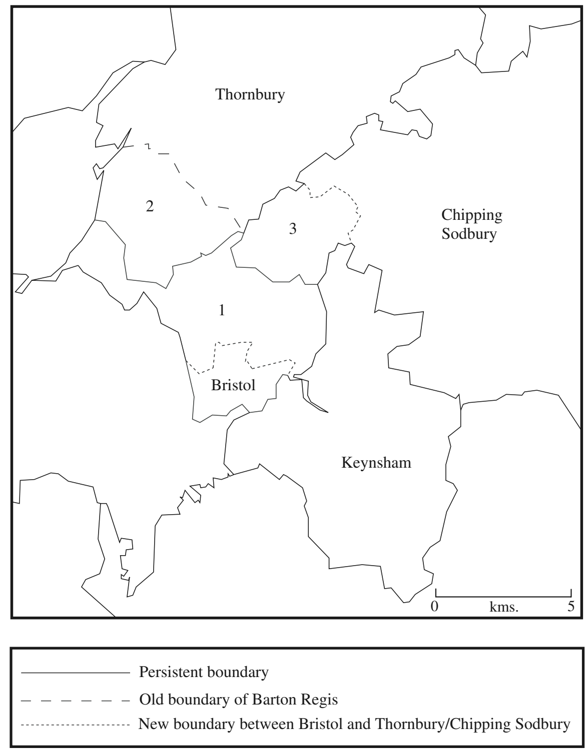


FIGURE 2. Boundary changes following abolition of Barton Regis registration district in 1905. Area 1 was allocated from Barton Regis to Bristol, area 2 to Thornbury, and area 3 to Chipping Sodbury.

Change over Space and Time: Population Loss and the Great Irish Famine

Using areal interpolation allows us to create long-run time series of spatially detailed data from statistics published with different administrative geographies. The potential to reexplore issues associated with spatiotemporal change using significantly enhanced data is enormous. However, making best use of these data requires new analysis techniques. The aim must be to stress spatiotemporal changes and to explore differences rather than to attempt to find generalizations that are meant to be equally applicable to all areas. To demonstrate this, we use the example of population loss in Ireland during and after the Great Famine. This issue covers only a relatively short time period for a relatively small country, but it is nevertheless sufficient to demonstrate the techniques and issues concerned.

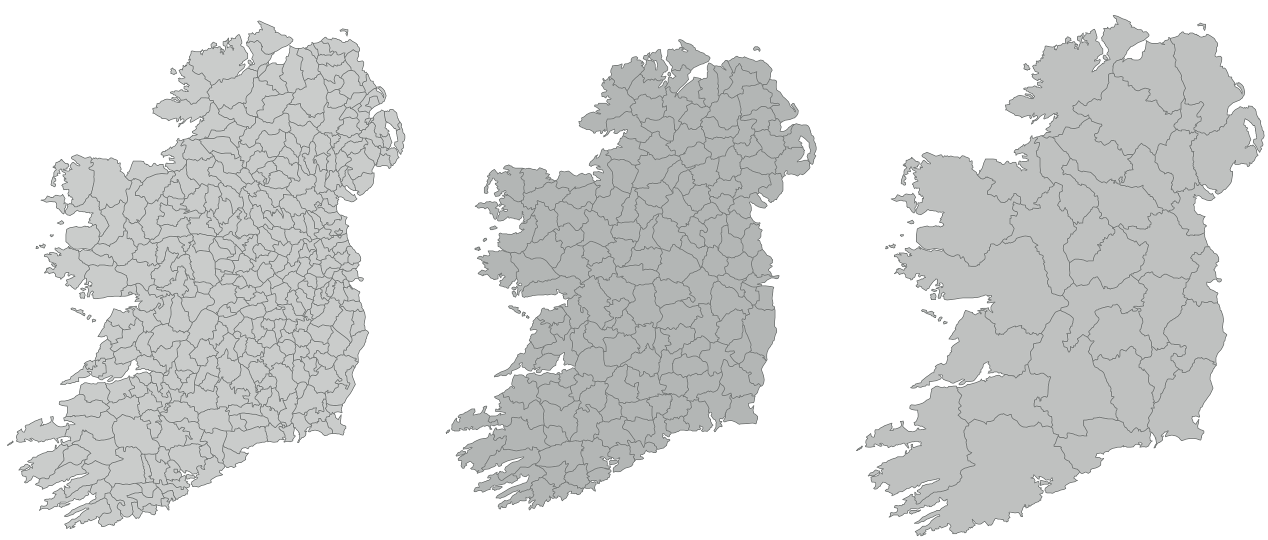
In the early nineteenth century, Ireland was heavily dependent on the potato crop. Between 1845 and 1849, this crop's repeated failure resulted in a catastrophic breakdown in Irish society, with mass starvation, disease, and emigration. The subsequent population loss is summarized in table 2. Over the decade from 1841 to 1851, Ireland lost 1.6 million people out of a population of 8.2 million, a decline of 20 percent. The decline continued for decades after the immediate impact of the Famine; by 1881, the population had fallen by 3 million. This was a drop of 37 percent over 40 years, at a time of rapid population growth in many western European countries (Kennedy et al. 1999).

Data on nineteenth-century Ireland are available from three levels, shown in figure 3. The most detailed of these is the barony, of which there were 334 in 1851, and for which the majority of demographic census data are available. Above these is the incompatible level of the Poor Law union. A variety of agricultural data are available at this level, as well as data on the amount of money spent on relief of the

TABLE 2. Population Decline in Ireland, 1841–81

		Loss in previous decade		Loss since 1841	
Date	Total pop.	Total	%	Total	%
1841	8,175,124	—	—	—	—
1851	6,552,385	1,622,739	19.8	1,622,739	19.8
1861	5,796,645	755,740	11.5	2,378,479	29.1
1871	5,412,337	384,308	6.6	2,762,787	33.8
1881	5,174,835	237,502	4.4	3,000,289	36.7

Note. — = data not applicable.
Source. Database of Irish Historical Statistics.



a. Baronies b. Poor Law union c. Counties

FIGURE 3. Major Irish administrative geographies in the mid-nineteenth century.

poor in each union. There were 163 units at this level in 1851; however, most of the data published at this scale are only available from 1851 onward, as the Poor Law system was being established in the early 1840s after the 1841 census. The final level consists of the 32 counties for which a variety of other datasets (including data on income and emigration) are available for certain, somewhat erratic dates.

Measuring the impact of the Famine is difficult. The most obvious effect is, of course, death, but mortality data associated with the Famine are notoriously poor. This has led researchers such as Joel Mokyr (1983) and S. H. Cousens (1960) to create their own county-level estimates of mortality during the Famine. Any analysis using these must be highly suspect, because they suffer from being highly aggregated and from being estimates. A second impact of the Famine was emigration. County-level data on emigration rates are available, but distinguishing between death and emigration is not as simple as it may seem. Vessels carrying migrants from Ireland had such notoriously high death rates, especially during the Famine itself, that they became known as "coffin ships," showing that emigration and death were by no means separate outcomes. A third impact was a marked decline in fertility. Mokyr acknowledged this when he used what he called his "upper bound" estimates, which include those who would have been born had it not been for the Famine. These he terms, rather prosaically, "the unborn dead."

Areal interpolation allows the calculation of intercensal population change at spatially detailed levels. This includes all three major demographic effects of the Famine—death, emigration, and fertility decline—although it is not possible to distinguish between them. One problem with this measure is that it will be affected by internal migration, so an area's loss of population through death or emigration may be offset by gains from people's migrating from other parts of the island.

The first stage in our analysis was to interpolate all the data onto a standard set of target units to allow all data to be directly compared. In this case, we used 1851 Poor Law unions. Once we interpolated the data onto these target units, we calculated population loss as follows:

$$\Delta y = [0 - (y_{t=2} - y_{t=1})/(y_{t=2} + y_{t=1})] \times 100, \quad (1)$$

where Δy is the loss in population y between time 1 (t = 1), the start of the decade, and time 2 (t = 2), the end of the decade. We subtract the value from zero to give a measure of population loss rather than the more conventional population gain, as population change in this period was almost exclusively negative. Using the sum of the start and the end populations as a denominator gives a number between 100, which is the complete depopulation of a previously populated area, and –100, which is a population moving into a previously unpopulated area, with 0.0 showing no population change over the decade. This normalized measure has the advantage over a more conventional rate, in which only the start population $y_{t=1}$ is used as the denominator, as it gives a symmetrical measure; thus a value of +20, for example, is the exact inverse of –20, which is not the case for conventional rates (Bracken and Martin 1995).

Figure 4 maps decennial population decline at the Poor Law union level for the decades between 1841 and 1861. The class intervals used, equal intervals of 10, remain constant to show how the pattern has changed. Clearly, the patterns change between the two decades, partly in response to the fall in population decline over time, but there also appears to be a geographical shift in the location of the decline—or gain, in rare cases. Choropleth maps such as these stress spatial patterns. Time, however, is not handled particularly well, as maps are required for each decade and the reader has to compare them. For only two decades, this is not too problematic, but for longer time periods, such as a century, it would impose serious difficulties. Subdividing attributes into classes simplifies them but may distort the pattern (Evans 1977). A further problem is that these maps are able to show only one variable at a time, in this case population loss, and are thus unable to suggest what factors may be linked to population loss.

Another problem with these maps is that it is difficult to draw generalizations, especially about where the largest population gain and loss are occurring and the changes over time. To illustrate these more clearly, one can use spatial statistical techniques to smooth the data. One approach is to use G_i statistics. A value of G_i is calculated for each polygon. It measures the degree to which a variable, in this case population decline, clusters around the polygon. High positive values of G_i indicate that the polygon has mainly high values of decline around it, whereas high negative values indicate that the surrounding area is bounded by low population decline or population gain. A distance decay model is used so that polygons close to the polygon being calculated have a higher influence on the result than those farther away (Getis and Ord 1996). Figure 5 shows that, in the 1840s, decline was concentrated in the west, especially around County Mayo, and also in the southwest. By the 1850s, this axis had shifted significantly with areas in the center of the south of the country; the area around Limerick and Tipperary was particularly affected. In all cases, the province of Ulster—in particular, east Ulster—was the least badly affected area, especially in the 1850s, when population change appears to be at its most polarized. Although these patterns can be distinguished from the raw data in figure 4, this use of exploratory statistics makes it easier to discern the pattern.

G_i statistics are a good example of a local analysis technique (Fotheringham 1997). Many statistical techniques can be criticized as being inappropriate for geographical data because they provide only a single summary statistic or model, with the implication that this applies equally across the study area. All well-known statistical methodologies, including correlation and regression analysis, do this. They can thus be termed *global* or *whole map techniques*, as they do not explicitly allow for the possibility

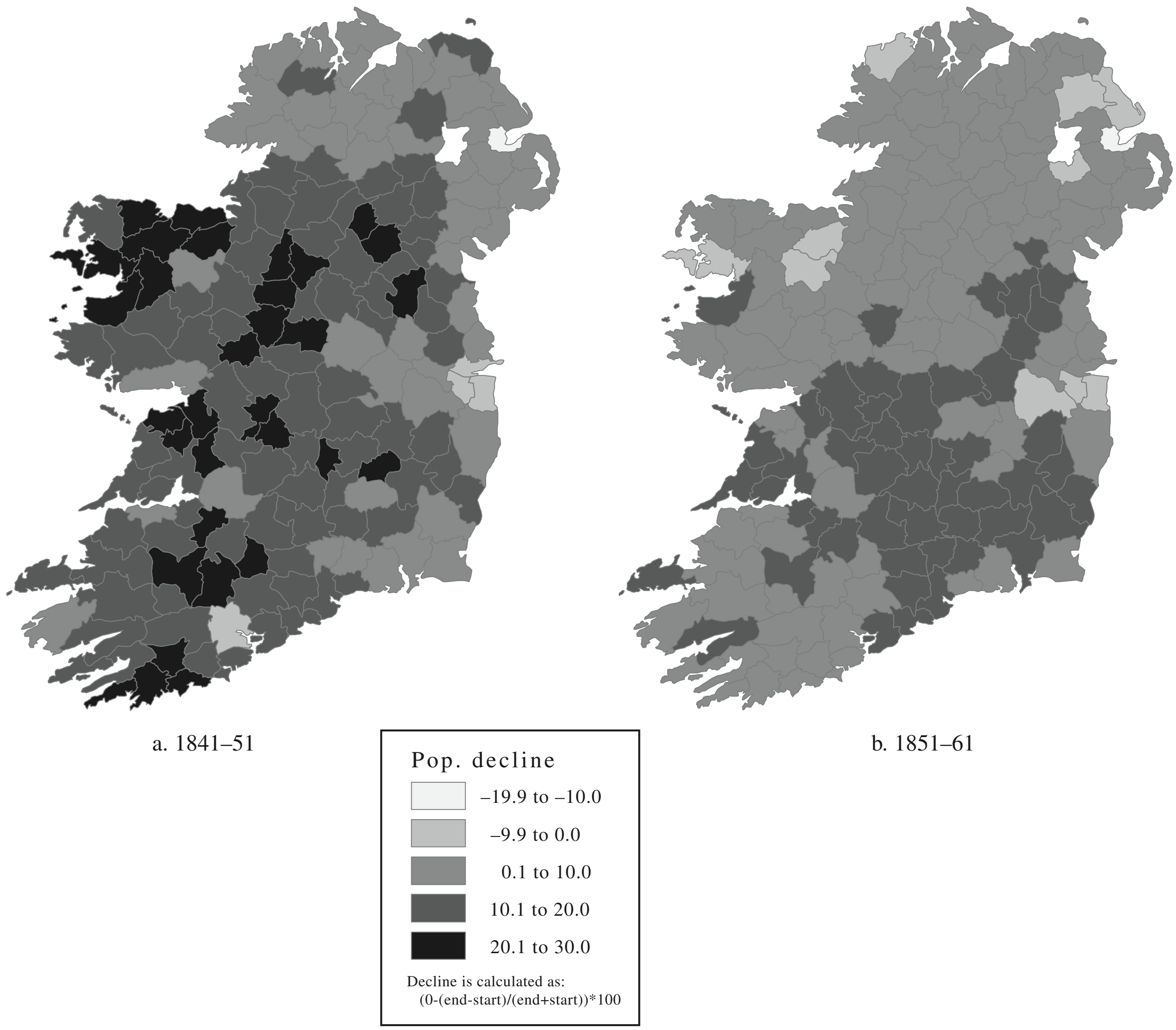


FIGURE 4. Population change in Ireland at Poor Law union level: (a) 1841–51 and (b) 1851–61. The legend uses equal intervals of 10 percentage points to allow comparisons to be made between the two maps.

that spatial patterns or relationships may change over space. Random variations from the overall pattern by individual observations can be explored, for example, by mapping regression residuals as shown later in this article, but more general statements about spatial variation such as "this is only found in the west" or "there is a strong positive relationship in the north and south but a negative relationship in the central area" cannot easily be made by global techniques. As a consequence, these techniques fail to allow space to be what Massey (1999, 274) called the "sphere of the existence of multiplicity."

Local analysis techniques aim to allow spatial variation by calculating separate values or models for every location on the study area, so in this example we calculate a separate value of G_i for each Poor Law union, allowing us to see where high values of population decline are concentrated, where (relatively) low values of population decline are concentrated, and where the pattern is less pronounced. The results of local analysis techniques are well suited to the GIS, where the spatial component of data is readily available and maps are frequently the most effective way of presenting the results.

The choropleth maps and G_i statistics used so far stress the spatial aspects of the data being investigated. However, they do not deal well with change over time and are only concerned with a single variable. An obvious next question is: Which sectors of society were most affected by this dra-

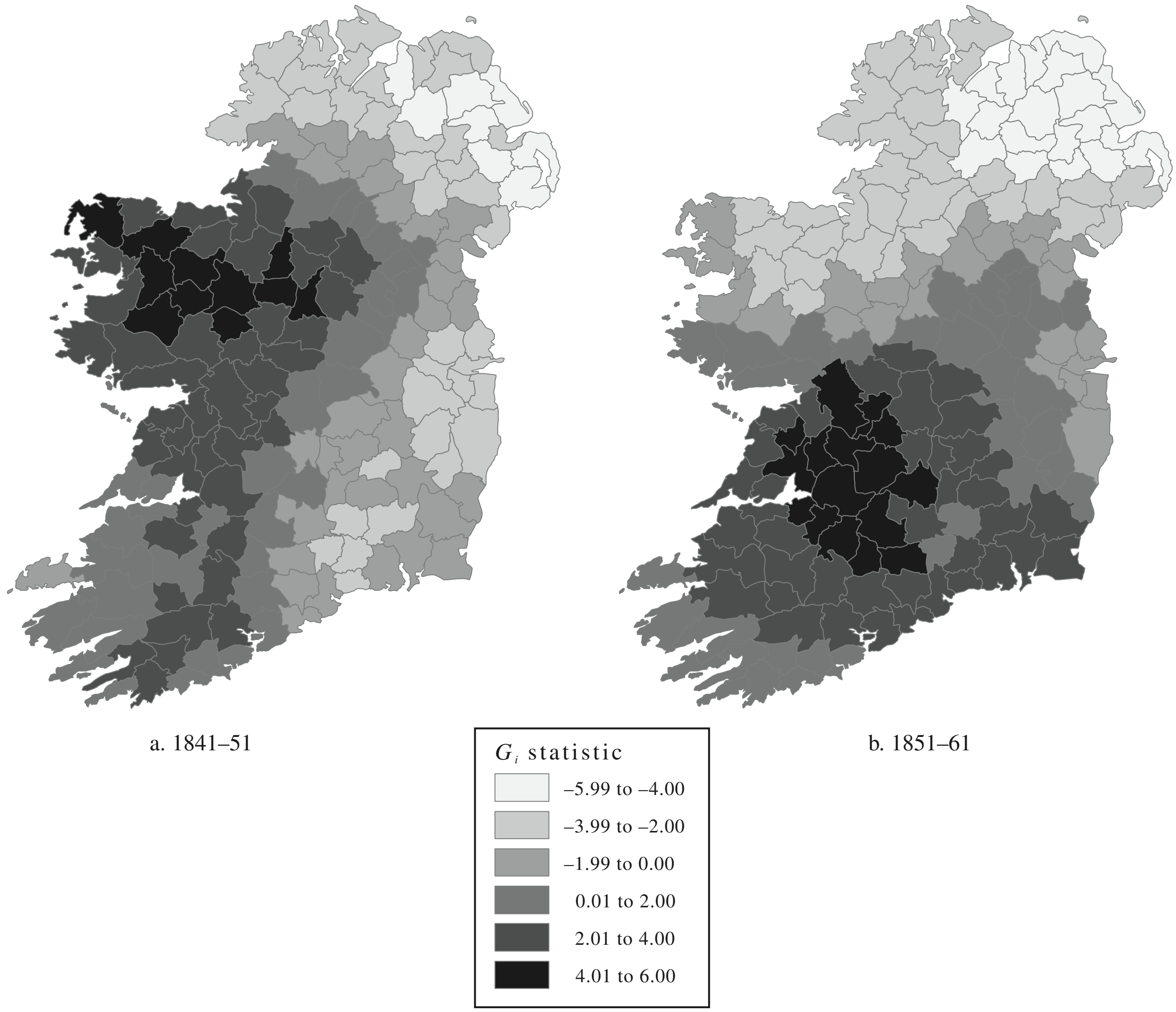


FIGURE 5. G_i statistics of population change in Ireland at Poor Law union level: (a) 1841–51 and (b) 1851–61. The legend uses equal intervals of 2 percentage points to allow comparisons to be made between the two maps.

matic decline in population, and did this change over time? The Irish censuses of this period subdivided the population into those who could read and write, those who could only read, and those who could do neither. Figure 6 shows that, through the 1840s and 1850s, a major reduction in the number of illiterate people took place that was not compensated for by an increase in the number of literates. Table 3 expands this graph. It shows that, in 1841, 3.77 million people, or 46 percent of the population, were classified as illiterate. By 1851, this had declined to 2.77 million, a loss of over 25 percent; but as a result of the loss of population, the illiterate had only declined to 42 percent of the total population. Over the same period, the population who could read but not write declined from 1.41 million to 1.20 million, but these individuals remained at around 18 percent of the population. The number of fully literate people remained roughly constant at around 2 million; however, as a result of decline in the other sectors, literates as a proportion of the total population rose from 24 percent to 30 percent. This trend of loss of the illiterate and semiliterate population without any significant gain in literates continues in the 1850s. Unfortunately, the data were not published in this form in 1871, so we cannot make longer term comparisons.

Kennedy et al. (1999) largely attributed these changes to a spread of literacy rather than a loss of the illiterate population through death or emigration. They wrote: “As the national

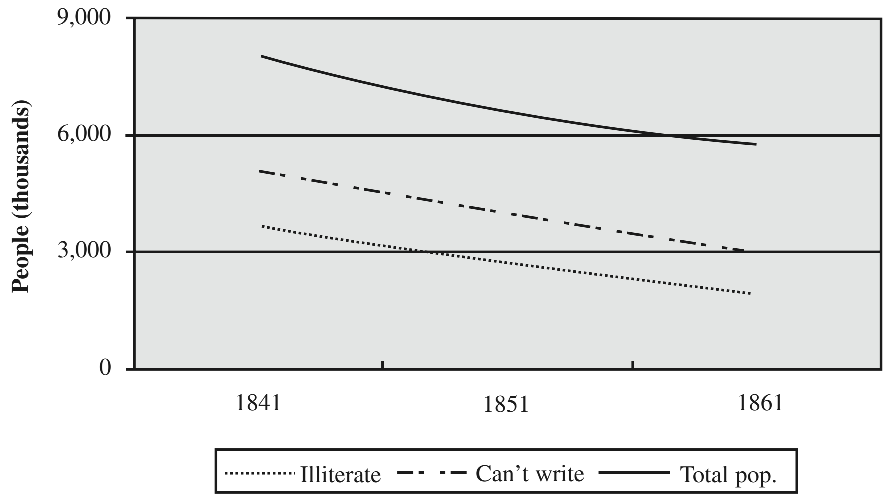


FIGURE 6. Changing patterns of literacy in Ireland, 1841–61. The people below the "illiterate" line are illiterate, those below the "illiterate" and "can't write" line can read but not write, and those between the "can't write" line and "total population" can both read and write. Note that children under age 5 are not included in diagram. (*Source*: Database of Irish Historical Statistics.)

TABLE 3. Changing Patterns of Literacy in Ireland, 1841–61 (in thousands)

		Illiterate		Read only		Literate	
Date	Total pop.	Total	%	Total	%	Total	%
1841	8,175	3,766	46.1	1,413	17.3	1,966	24.0
1851	6,552	2,766	42.2	1,203	18.4	1,939	29.6
1861	5,796	1,973	34.0	1,023	17.7	2,106	36.3

Note. Only people over age 5 are included in literacy figures, so these do not add up to the total population.
Source. Database of Irish Historical Statistics.

school system diffused literacy to the younger generations, illiteracy became a minority phenomenon" (ibid., 94), and they went on to say that "the direction of change in literacy in the period 1841–71 is not in doubt. Explaining the course of change is more difficult. A short-term effect was the Famine, one of whose effects was to reduce the numbers and proportions of illiterate people in the population. Of longer-term significance, there was the wider exposure of Irish economy and society to market forces. Increasing monetisation [sic] and a deeper involvement in production for the market made rudimentary numeracy and writing skills in the post-Famine decades" (ibid., 96). Figure 6 and table 3 contradict their conclusions. The literate population remains roughly constant at around 2 million, whereas the population who could read but not write declined by 400,000 and the illiterate population declined by 1.8 million. Therefore, the reason literacy rates rose in Ireland over this period was the decline in illiteracy or, more precisely, a decline in the number of illiterate people without a subsequent rise in the number of literate people. This lack of a rise in the literate population strongly suggests that improvements in schooling were largely irrelevant. The increase in the proportion of the population who were literate instead seems to have been caused by a loss of illiterate people, and the most probable cause for this appears to be emigration or death. Although this pattern appears at the national level, our article will show that the regional pattern is far more complicated.

Figure 7 shows changes in housing quality over the same period. The Irish census subdivided households by using a four-way classification. Fourth-class houses were mud cabins consisting of only one room; third-class houses were still made of mud but had two to four rooms and windows: second-class houses were good farmhouses or town houses having five to nine rooms and windows; and first-class houses were still better than the others (Kennedy et al. 1999). The pattern of change shown is similar to that shown for illiteracy in figure 6. The number of fourth-class houses declines dramatically from just over half a million in 1841 to 136,000 in 1851 to 89,000 in 1861, a loss of over 400,000 dwellings, around 80 percent of the total. Third-class housing shows a more modest decline from 542,000 in 1841 to 512,000 in 1851 and then 489,000 in 1861, a loss of 53,000 houses, or around 10 percent. First- and second-class housing both show modest increases from a combined total of 306,000 houses in 1841, through 368,000 to 416,000 in 1861, a rise of just over 100,000 houses or 30 percent, which nowhere near compensates for the loss of fourth-class housing in particular.

Kennedy et al. (1999, 79) offered a very different explanation for this decline than for the decline in illiteracy. They said that "the Famine swept away more than one million souls and more than a million others fled the country during the course of the calamity. Cottiers and labourers bore the brunt of the crisis. With the destruction of these social elements went the destruction of their housing, which was generally the poorest available. . . . Thus change in housing during the Famine was essentially a by-product of the class-specific nature of the Famine."

Superficially, there appears to be a consistent pattern shown by these two variables. The dramatic loss of illiterate individuals and of fourth-class housing suggests that population loss during the Famine was concentrated among the poor. Little evidence exists that these people were becoming better educated or better housed; it seems far likelier that they were either emigrating or dying. This decline continues through the 1850s, suggesting that the population loss among the poor lasted longer than the immediate impact of the Famine. Although these statements are quite convincing when one looks at figures 6 and 7 and table 4, they are nevertheless broad generalizations; any conclusions based on them are likely to be tentative. In addition, they present only one story and suggest that it was happening throughout the country, with no variations between east and west, north and south, and country and town.

There is no way of accurately measuring or even precisely defining poverty in mid-nineteenth-century Ireland; however, illiteracy and poor housing provide good indicators. As such, do they provide a consistent indicator of poverty across the whole country? By showing a scatter plot of the proportions of fourth-class housing and illiteracy in 1841, figure 8 indicates that they broadly do so before the Famine. It shows that there is a clear relationship between the two, which, using simple linear regression, we can quantify as

$$\text{Illiteracy} = 20.0 + 0.75\text{frth\_class}, \qquad (2)$$

with an r^2 of 61.3 percent showing a very strong relationship between the two that corresponds to a positive correlation of 0.78, an unusually strong relationship with social science data.

Although this is a strong relationship, conventional statistics do not tell us where this relationship is strongest and, perhaps more important, which places do not fit the pattern. Figure 9 is a map of the regression residuals from figure 8. For each Poor Law union, it maps how far the actual values of illiteracy deviate from the trend line shown in equation (2). The map is interesting in that it shows that in Ulster,

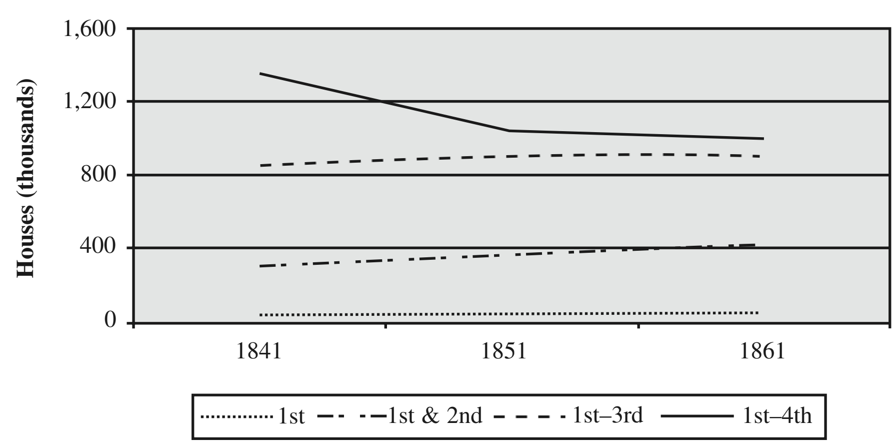


FIGURE 7. Changing housing types in Ireland, 1841–61. The area below the lowest line represents first-class housing; between this and the next line is second-class housing, and so on. (*Source*: Database of Irish Historical Statistics.)

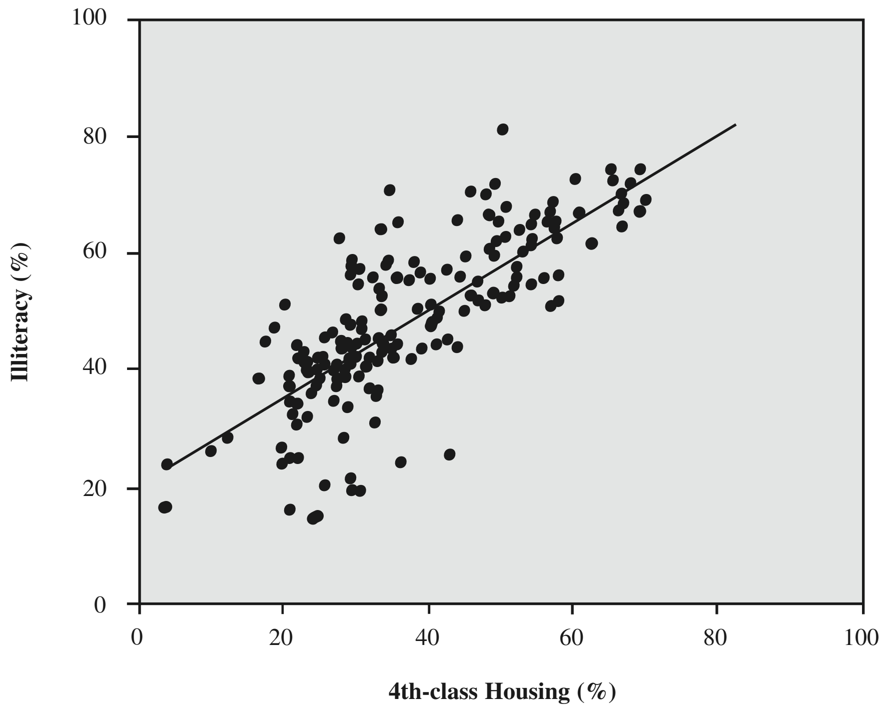


FIGURE 8. Illiteracy and fourth-class housing in 1841. The graph includes a trend line with the equation shown in equation (2). Illiteracy is shown as a percentage of the population over age 5; fourth-class housing is shown as a percentage of total housing.

particularly in Counties Antrim and Londonderry, the link between fourth-class housing and illiteracy is different from that for much of the country, with illiteracy being far lower than would be expected, given the amount of fourth-class housing. This may simply be because of better education in these areas, but it could just as easily be a sign of a more complex social structure in Ulster, which had a large Protestant population and was more industrialized than the rest of the country. Illiteracy is higher than would be expected in Galway and Cork. It is not our intention to attempt to explain this here; we simply note that, although there is a very strong relationship between fourth-class housing and illiteracy for most of the country, there are exceptions, and these exceptions have a pronounced geographical pattern.

So was there a relationship between poor housing and illiteracy and population loss during the Famine and, if there was, did population loss after the Famine sustain this relationship? An additional question, suggested by the maps of population change shown in figure 4, is: Do urban areas also affect the pattern? We could again use a regression analysis to answer this question, but there are two major problems. The first is that a conventional regression analysis does not allow relationships to vary over space because it is a global technique. Mapping residuals, as shown in figure 9, is one option, but this merely illustrates where the model is accurate and where it overestimates or underestimates the value of the dependent variable, which in figure 9 is illiteracy. An alternative approach is to use geographically weighted regression (GWR), which allows each of the coefficients to vary over space. GWR is a local technique in which a model is calibrated for each location in the study area. These are allowed to vary because a distance decay model is used such that observations near the location currently being calculated have more weight than those farther away. In this way, it becomes possible to investigate how each variable in the model varies over space (Brunsdon, Fotheringham, and Charlton 1996; Fotheringham, Brunsdon, and Charlton 2002).

The second problem is that regression analysis requires the independent variables to be independent of each other, or the parameter estimates produced will be unreliable (this is termed *multicollinearity*). We have already demonstrated that there is a close relationship between fourth-class housing and illiteracy in 1841, so using these variables in the same model is problematic. The only way to handle this is not to include both variables in the same model. Similarly, if we are interested in all four classes of housing, these also cannot be included in the same model.

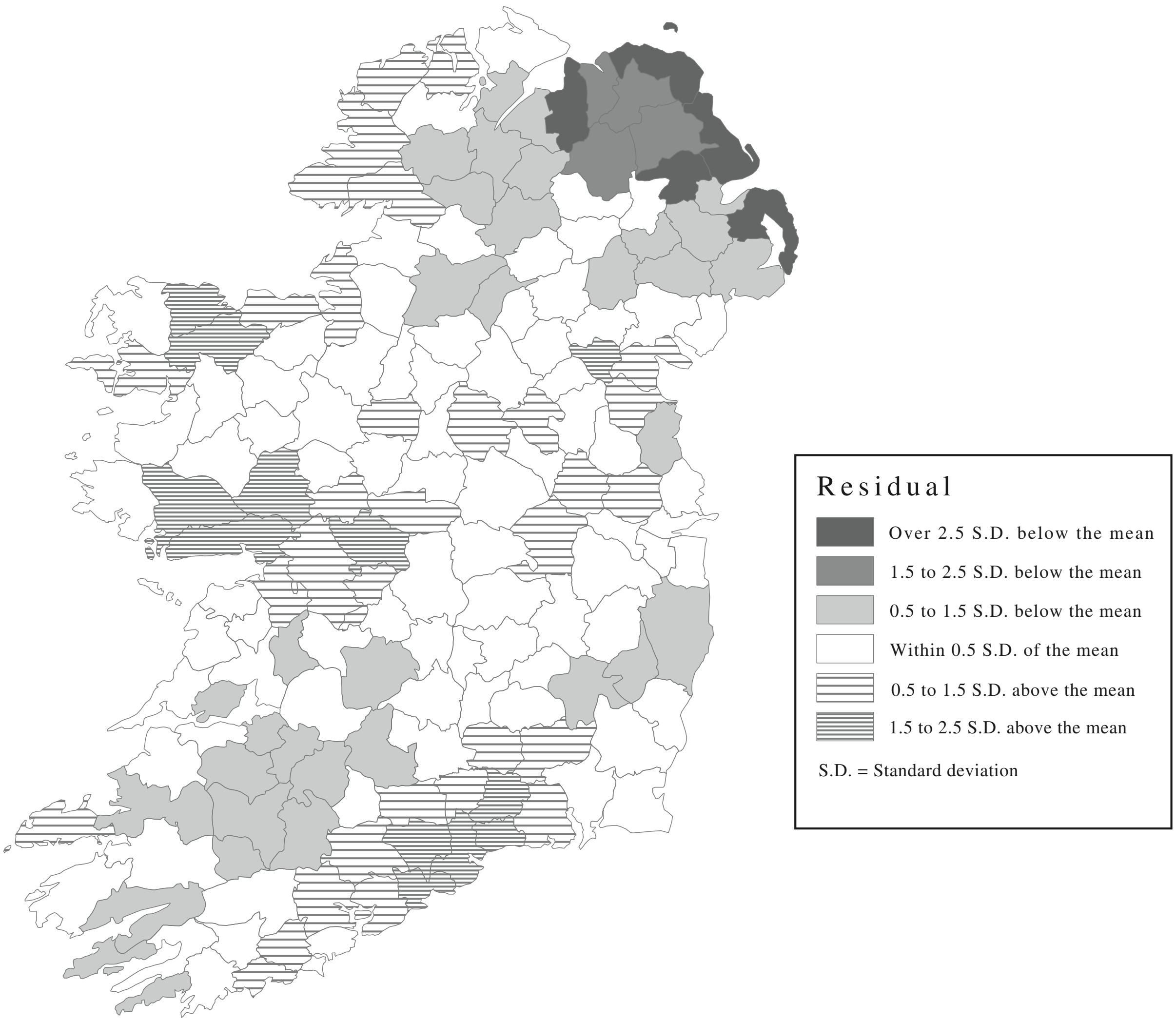


FIGURE 9. Regression residuals predicting illiteracy from fourth-class housing, 1841 (see fig. 8). The regression equation is given in equation (2). The mean residual is 0.00, and the standard deviation is 8.99.

A number of possible independent variables are available to us at the subcounty level that may help to predict population loss over the 1840s. Illiteracy and the four classes of housing provide five. The degree of urbanization is another. One way of representing urbanization is to use population density, but that is problematic because population density is highly sensitive to boundary effects. At the county level, it is all but meaningless. At the Poor Law union level, a small town may be surrounded by a large rural area, thus lowering its population density considerably; it may entirely fill the union; or, in the case of the largest cities such as Dublin or Belfast, it may be spread between two or more unions. Indeed, Poor Law union boundaries were specifically intended whenever practical to include both poorer rural areas and potentially wealthy towns, so that the union had the resources to support its poor. Therefore, in this analysis, we use binary variables rather than population density to represent whether or not a union contains a town. The 1841 census lists Irish towns and their populations. As large towns such as Dublin and Belfast may have a different impact than do small towns, we handle them as two separate variables: we class small towns as those with populations from the 1841 census of 5,000 to 50,000, whereas large towns have populations of over 50,000. There are only six unions that contain large towns: two for Dublin, and one each for Belfast, Limerick, Waterford, and Cork. This gives two more independent variables; both of them are binary variables that have a value of one if they contain a small or large town, respectively, and a value of zero otherwise.

To avoid multicollinearity, we have to generate five models to help predict population decline. In each model, we include the presence or absence of small and large towns, plus one of the four classes of housing or illiteracy. The parameters we generated from conventional regression by using these five models are shown in table 4. The pattern demonstrated is consistent: population loss over the 1840s was positively related to fourth-class housing and illiteracy and was negatively related to first- and second-class housing. Third-class housing has no significant relationship with population loss. In all cases, the presence of small towns reduces population loss, whereas the presence of a large town reduces it even more. In performing the analysis, we subtracted the mean from the numeric independent variables. Thus, the intercept term now provides the predicted population loss from a union with the average value of housing type or illiteracy and without a town, rather than values of zero for housing type or illiteracy. These values are high, with a normalized value—as calculated in equation (1)—of around 13.5 percent. This suggests that all over Ireland there was a large amount of population loss. In areas with high amounts of fourth-class housing and illiteracy, the loss was more substantial, whereas high rates of first- or second-class housing reduced the loss. The presence of towns, particularly large towns, reduces even population loss once we take the impact of housing or illiteracy into account. This supports the suggestions that the majority of population loss was highest among areas with high levels of fourth-class housing and illiteracy, although loss is occurring at high rates everywhere (see figs. 6 and 7). The main additional piece of knowledge gained is that urban areas are experiencing lower levels of population decline, particularly in the largest towns. This may happen because the impact of the Famine was lower in urban areas, or because population loss from urban areas is being offset by migration into towns from the country.

The use of GWR shows that the patterns are far more complicated than these simple results suggest (see fig. 10). To keep the patterns relatively simple, the maps show only whether the parameters are either positively or negatively statistically significant at the 5 percent level. There is good evidence for the pattern identified in the global analysis down the eastern side of the country between Belfast and Cork, where population decline increases with illiteracy and fourth-class housing and decreases with first-class and, to a lesser extent, second-class housing. There is also weaker evidence that third-class housing also had a positive relationship in these areas. However, the western side of the country, where the population loss from the Famine was at its highest, shows a very different and in some cases directly contradictory pattern. Here, the relationships are far less well defined, but some areas (particularly in Mayo, Kerry, and western Cork, and, to a lesser extent, around Londonderry) with relatively large amounts of first- and second-class housing and relatively low amounts of fourth-class housing actually experience higher population loss than those that would appear to be poorer. This suggests that, in the east of the country, there was a reasonably straightforward relationship between areas with low standards of living in the pre-Famine period and population loss during the Famine. To the west, however, the pattern is more complicated. It may be that the explanation for this lies in large estates' driving the poor from the land, but investigating this would require further research. It may also be that, in the west, the simplicity of the model is having a particularly serious impact and, for example, agricultural variables unavailable at this scale for this date may be having a particularly strong impact in these areas. This is backed up by the fact that, in all cases, the local r^2 variables are lower in these western areas (where they are typically around 50 percent) than they are in the east (where they are often over 80 percent).

Repeating this analysis by using independent variables from 1851 and 1850s population decline as the dependent variable provides both similarities and differences. Table 5 shows the global regression parameters we estimated by using these data. There are some interesting differences.

TABLE 4. Global Regression Coefficients for Population Decline in the 1840s

			Housing (class)				
Intercept	Small	Large	First	Second	Third	Fourth	Illiteracy
13.61	**–2.97**	**–11.54**	**–0.57**				
13.34	–1.87	**–10.10**		**–0.24**			
13.89	**–3.54**	**–16.22**			–0.02		
13.54	**–2.39**	**–12.85**				**0.12**	
13.52	**–2.26**	**–12.87**					**0.17**

Note. The independent variables are small and large towns, and each of the five variables are separate. These variables have been offset by their mean; thus, the intercept shows that in all cases a place with an average of that type of housing or illiteracy and no towns will have a population decline of around 13.5 percent. Estimates that are statistically significant at the 5 percent level are shown in boldface type.

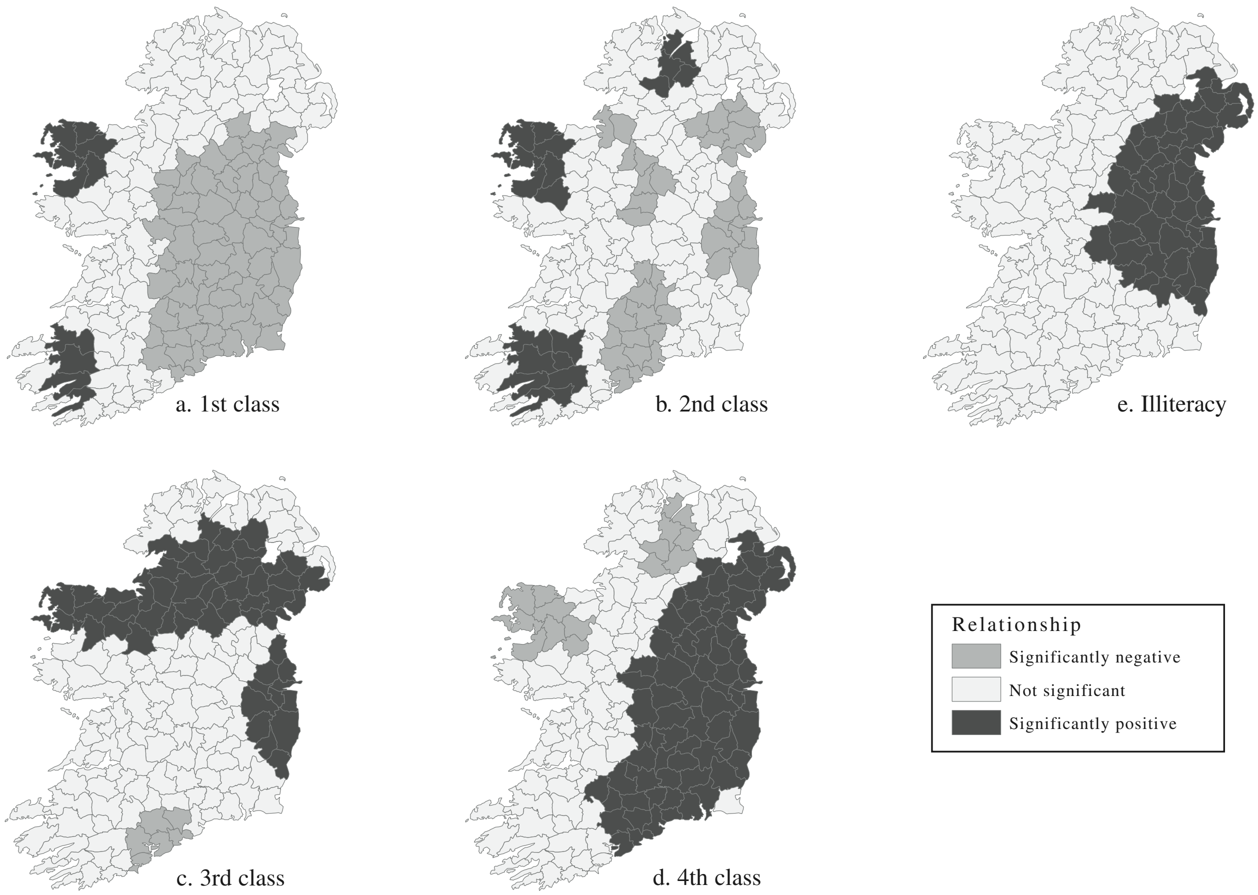


FIGURE 10. Geographically weighted regression results for population decline of 1840s. The independent variables are small and large towns, and (a) first-class housing, (b) second-class housing, (c) third-class housing, (d) fourth-class housing, and (e) illiteracy. The 5 percent level is used to determine statistical significance.

TABLE 5. Global Regression Coefficients for Population Decline in the 1850s

			Housing (class)				
Intercept	Small	Large	First	Second	Third	Fourth	Illiteracy
7.12	1.10	–4.15	–0.03				
6.98	1.61	–2.81		–0.07			
7.36	0.46	**–7.10**			–0.09		
6.98	1.58	–2.93				**0.13**	
6.94	1.77	–2.78					**0.09**

Note. The independent variables are small and large towns, and each of the five variables are separate. These variables have been offset by their mean; thus, the intercept shows that in all cases a place with an average of that type of housing or illiteracy and no towns will have a population decline of around 7 percent. Estimates that are statistically significant at the 5 percent level are shown in boldface type.

Small and large towns were very important in reducing population loss in the 1840s, but in the 1850s this impact has all but gone. Indeed, small towns are always positive and never statistically significant, and large towns have a negative impact that is only significant when we use third-class housing and that we can thus consider to be far less important than that in the 1840s. Fourth-class housing and illiteracy still have a positive relationship with population loss, and again there is no relationship with third-class housing; however, unlike the situation in the 1840s, there is no evidence that population loss was reduced by first- or second-class housing.

The GWR results suggest that the divisions shown in the 1840s have become more polarized (see fig. 11). On the east side of the country, a positive relationship exists between third- and fourth-class housing and illiteracy and a negative relationship with first- and second-class housing, as might be expected. In the west of the country, however, the reverse is true, with positive relationships with first- and second-class housing and negative relationships with the other three. It is also noticeable that there are few significant relationships in either the north or the south of the country.

The global regression suggests that, in the 1840s, population loss was generally widespread but was higher in areas with relatively high rates of fourth-class housing and illiteracy and lower in urban areas and areas with relatively high levels of first- and second-class housing. In the 1850s, the rates of population loss had declined; although they continued to be associated with fourth-class housing and illiteracy, first- and second-class housing and urban areas no longer seem to reduce these rates, suggesting perhaps that although population loss was still prevalent among the rural poor, it had indeed spread throughout society. The GWR

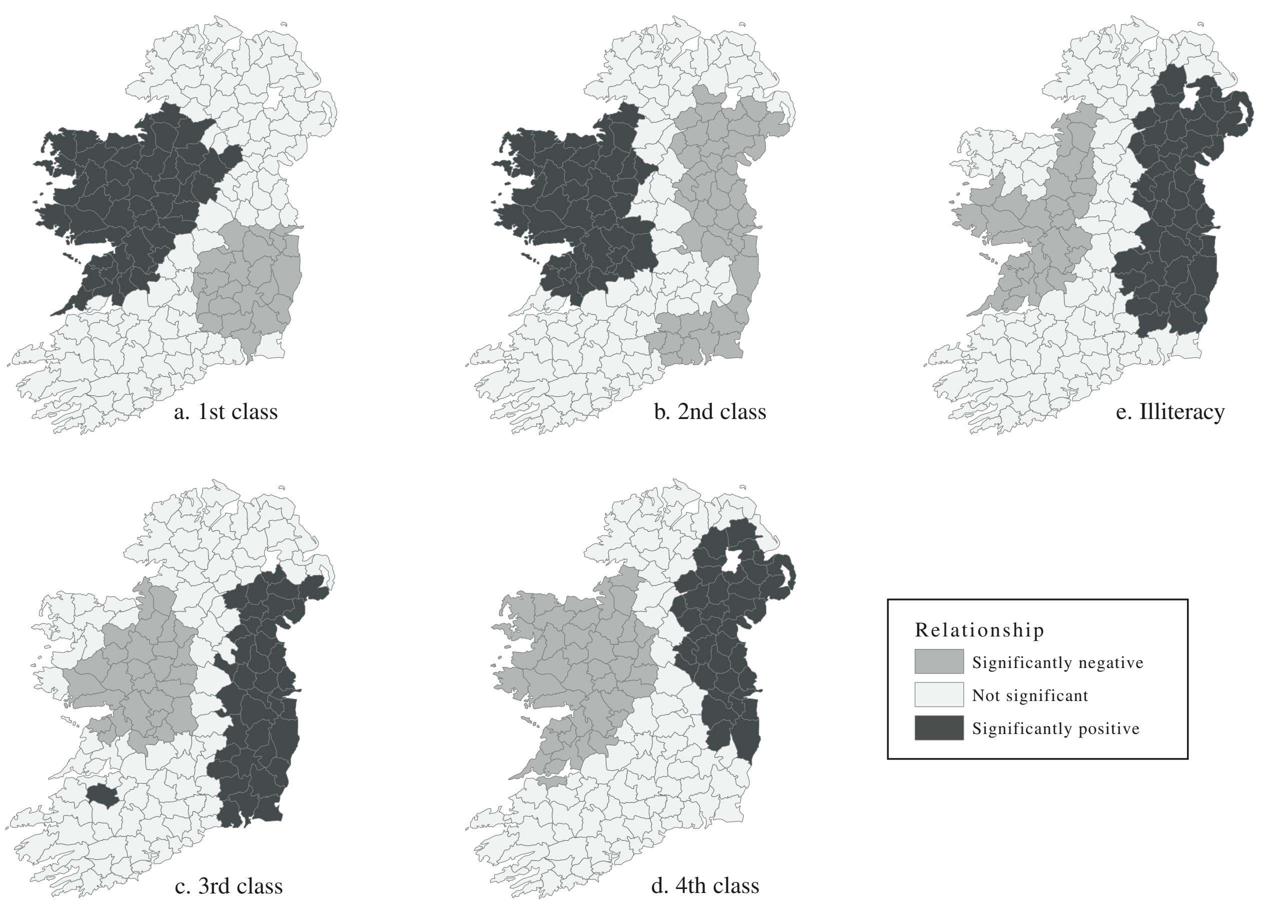


FIGURE 11. Geographically weighted regression results for population decline of 1850s. The independent variables are small and large towns, and (a) first-class housing, (b) second-class housing, (c) third-class housing, (d) fourth-class housing, and (e) illiteracy. The 5 percent level is used to determine statistical significance.

shows that this is highly simplistic and mainly prevalent in the east of the country. In the west, it appears almost as if areas with relatively large amounts of good housing are losing population, whereas there does not appear to be a relationship in the north, and there are some interesting variations in the south.

Figure 12 summarizes expectation scores for the 1840s and 1850s. For each union, we have calculated the number of "expected" and "unexpected" results. Expected results are positive relationships with fourth-class housing or illiteracy and negative relationships with first- or second-class housing; thus, an area can have a maximum count of four expected results. Unexpected results are the inverse: positive relationships with first- or second-class housing and negative relationships with fourth-class housing or illiteracy. Again, an area can have a maximum of four unexpected results. When an area has both an expected and an unexpected result, the results cancel each other out. Figure 12 is limited by having to use gray-scale cartography, and thus it only divides the country into three types of areas: (1) those with two or more expected results when there does appear to be a clear link between poverty, as measured by poor housing and literacy, and high rates of population loss; (2) those with two or more unexpected results when population loss seems higher in "wealthier" unions than in poorer ones; and (3) those with no clear relationship between these indicators and population loss.

The area of expected results is confined to the east side of the island. The unexpected results are found in a relatively small area in Mayo in the 1840s. Interestingly, this is the area where population loss is at its highest (see figs. 4 and 5). By the 1850s, this area has spread both eastward and southward to cover a far larger area. Areas of the highest population loss also have a more southerly trend in the 1850s, but the relationship between the two is far from clear. Perhaps most interesting of all is the extent of the

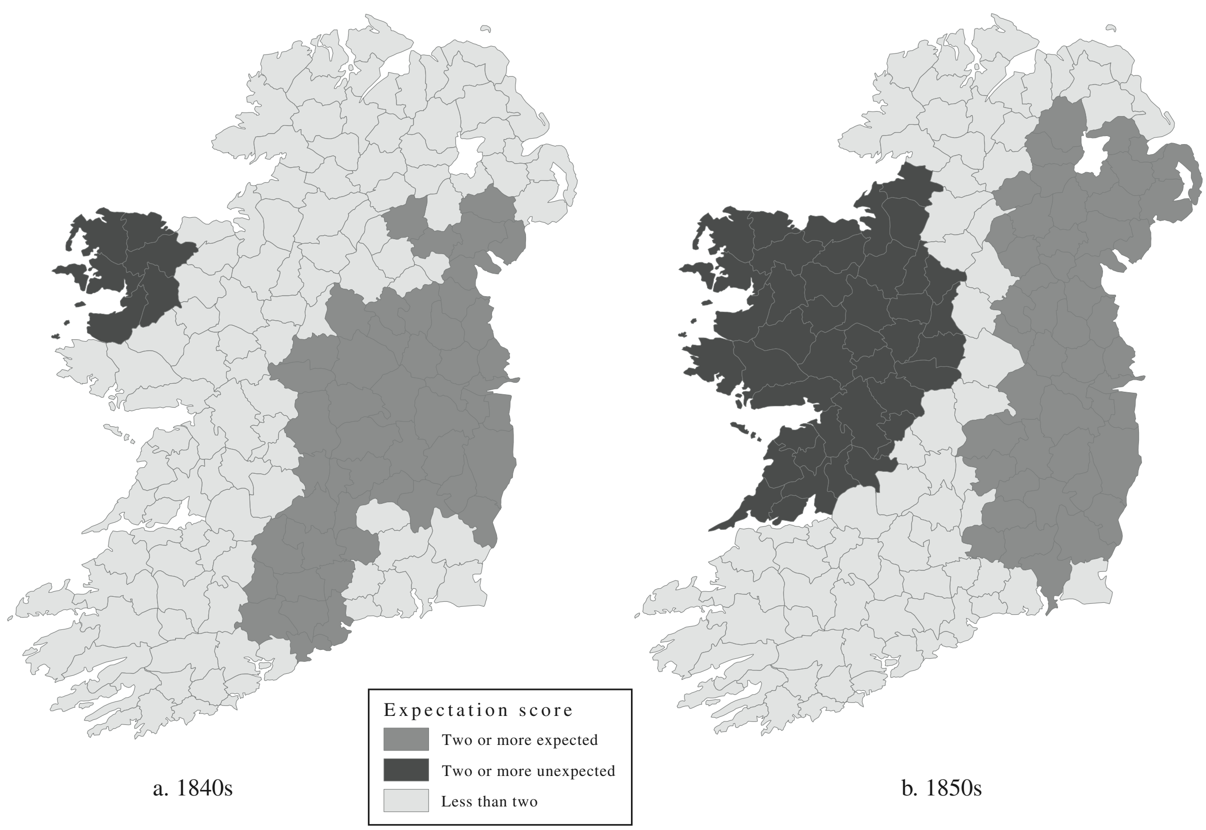


FIGURE 12. Expectation score summaries from geographically weighted regression results for 1840s and 1850s. An "expected" result is a positive relationship between population loss and fourth-class housing or illiteracy, or a negative relationship with first- or second-class housing. An "unexpected" result is the inverse: a negative relationship between population loss and fourth-class housing or illiteracy, or a positive relationship with first- or second-class housing. Thus, each union can have a maximum of four expected or unexpected results. If it has one of each, then these results cancel each other out.

country where no clear relationship exists between population loss and wealth. These areas are concentrated in the north and the south. In the north, this may be partially explained by the lack of a clear relationship between fourth-class housing and illiteracy (see fig. 9). In the south, however, we find no obvious explanation.

There clearly is a link between poverty, as defined by poor housing and illiteracy, and population loss over both the Famine decade and the one that followed. The advantage of GWR is that it allows the relationship to vary over space, and thus we can explore how this relationship varies; it is clear that it varies significantly over both space and time (see figs. 10, 11, and 12), which has two advantages. First, it may give us clues as to other variables required if more complex quantitative models are to be developed; second, and perhaps more interesting, it allows us to explore the quantitative indicators in more complex ways and perhaps explain them on the basis of nonquantitative information. We must stress, however, that these are simple relationships based on crude data. Explaining them requires additional research that may well be neither GIS based nor quantitative. What is important, however, is that it demonstrates that lessons learned either from national-level research or local case studies of places such as Mayo in the west or Skibbereen in the south may not be broadly applicable.

Conclusions

Quantitative methods have had a controversial history in both geography and history (see, e.g., Bennett 1981; Billinge, D. Gregory, and Martin 1984; S. Gregory 1976; Johnston 1983; Keat 1981; Taylor 1976). A particular criticism of quantitative approaches is that they imply, and sometimes directly lead to, logical positivism, whereby a relationship is found and it is suggested that this relationship is in some way an inevitability from which there will be few exceptions. This criticism has led to a strong suspicion of statistical approaches from many factions of both disciplines. In spite of this suspicion, quantitative approaches have maintained a strong influence in geography in particular (Flowerdew 1998).

In this article, we have demonstrated the use of a variety of techniques that can be used within GIS analysis to demonstrate how quantitative techniques may be used in a more flexible manner than was traditionally possible. Our particular aim has been to make use of all three components of the data: the attribute, the spatial, and the temporal. To handle space, we have used mapping extensively. To handle time, we have used time-series graphs. Statistical operations allow all three components to be manipulated in more complex ways than simple mapping. Much of the article relies on areal interpolation to put all the data into a single set of administrative units—Poor Law unions from 1851—to allow them to be compared; this interpolation is essential for all the calculations of rates of change. We have used G_i statistics to simplify the spatial patterns in a single variable to make it easier to identify broad trends. Finally, we have used GWR to explore the changing relationship between population loss and housing, illiteracy, and urban areas. We have done this in an exploratory manner to see where and when there appears to be a relationship, and our conclusions have concentrated on spotting trends rather than trying to explain them. Explanation is a further step requiring additional information that does not necessarily have to be quantitative.

Our intention has been to illustrate that a strength of using the GIS and spatial analysis with quantitative sources is that, unlike logical positivism, spatial analysis techniques not only find relationships among variables but also discern where these relationships vary from the norm. Thus, rather than trying to discover inflexible laws of human behavior, our aim is to explore the diversity of the patterns found within the data.

Sources such as the census are indisputably crude, but they are also among the most comprehensive sources available for a study of the past. Using GIS and spatial statistical techniques with these data is unlikely to provide an explanation of the causes of past spatial patterns, but it does enable us to explore the spatial patterns within single variables and the relationships between variables. By allowing the patterns and relationships to vary over space, it is likely to either confirm or challenge conventional explanations without directly providing explanations. Thus, this approach is "fundamentally descriptive and provocative, rather than necessarily being interpretative and productive" (Baker 2003, 44). Although this quotation describes approaches to historical GIS based on mapping, the additional statistical approaches make little difference. The GIS and spatial analysis do not replace or reduce the importance of conventional historical approaches; in fact, they should complement, challenge, and enrich them. Without more conventional textual and narrative approaches to history, there is no way of explaining the patterns presented by quantitative, statistical GIS approaches. However, these approaches provoke textual and narrative histories to explain the patterns that the GIS finds and may challenge the assumption that an explanation found for one particular place is valid for other places.

NOTES

This research was funded by Research Methods Programme Grant H333250016 from the Economic and Social Research Council (ESRC). Thanks to Ciaran Higgins (Centre for Data Digitisation and Analysis, Queen's University Belfast) and Dr. E. M. Crawford (School of Sociology and Social Policy, Queen's University Belfast) for their help with this article. Thanks also to Leonidas Housos for creating the VisStats program that helped in some of the spatial analysis routines used here. This software was written as part of his Master's of Science in GIS at the University of Portsmouth. The GWR analysis was performed with GWR 2.0 software provided by Martin Charlton, Stewart Fotheringham, and Chris Brunsdon (Department of Geography, University of Newcastle).

1. This conference was held at the Newberry Library in Chicago in March

2004; see http://www.newberry.org/hgis.
2. See http://www.ssha.org.
3. See http://www.aag.org.
4. See http://www.rgs.org.
5. See http://www.nhgis.org.
6. See http://www.fas.harvard.edu/~chgis.
7. *Polygon* is the technical term for the structure that a GIS uses to represent an area or zone.

REFERENCES

Bailey, T. C., and A. C. Gatrell. 1995. *Interactive spatial data analysis*. Harlow: Longman.
Baker, A. R. H. 2003. *Geography and history: Bridging the divide*. Cambridge: Cambridge University Press.
Bennett, R. J. 1981. Quantitative geography and public policy. In *Quantitative geography*, edited by N. Wrigley and R. J. Bennett, 387–96. London: Edward Arnold.
Billinge, M., D. Gregory, and R. Martin, eds. 1984. *Recollections of a revolution: Geography as spatial science*. London: Macmillan.
Bol, P., and J. Ge. 2005. The China Historical GIS. *Historical Geography* 33:146–48.
Bracken, I., and D. Martin. 1995. Linkage of 1981 and 1991 UK censuses using surface modelling concepts. *Environment and Planning A* 27:379–90.
Brunsdon, C., A. S. Fotheringham, and M. E. Charlton. 1996. Geographically Weighted Regression: A method for exploring spatial nonstationarity. *Geographical Analysis* 28:281–98.
Chrisman, N. R. 1999. What does "GIS" mean? *Transactions in GIS* 3:175–86.
Cousens, S. H. 1960. Regional death rates in Ireland during the Great Famine from 1846 to 1851. *Population Studies* 14:55–74.
Darby, H. C., R. E. Glasscock, J. Sheail, and G. R. Versey. 1979. The changing geographical distribution of wealth in England: 1086–1334–1525. *Journal of Historical Geography* 5:247–62.
De Moor, M., and T. Wiedemann. 2001. Reconstructing Belgian territorial units and hierarchies: An example from Belgium. *History and Computing* 13:71–97.
Department of the Environment. 1987. *Handling geographic information: Report of the Committee of Enquiry Chaired by Lord Chorley*. London: Her Majesty's Stationery Office.
Ell, P. S., and I. N. Gregory. 2001. Adding a new dimension to historical research with GIS. *History and Computing* 13:1–6.
Evans, I. S. 1977. The selection of class intervals. *Transactions of the Institute of British Geographers, New Series* 2:98–124.
Fischer, M. M., H. J. Scholten, and D. Unwin. 1996. Geographic Information Systems, spatial data analysis and spatial modelling. In *Spatial analytical perspectives on GIS*, edited by M. M. Fischer, H. J. Scholten, and D. Unwin, 3–20. London: Taylor & Francis.
Flowerdew, R. 1991. Spatial data integration. In *Geographical Information Systems: Principles and applications. Vol. 1: Principles*, edited by D. J. Maguire, M. F. Goodchild, and D. W. Rhind, 375–87. Harlow: Longman.
———. 1998. Reacting to *Ground Truth*. *Environment and Planning A* 30:289–301.
Fotheringham, A. S. 1997. Trends in quantitative methods I: Stressing the local. *Progress in Human Geography* 21:88–96.
Fotheringham, A. S., C. Brunsdon, and M. Charlton. 2000. *Quantitative geography: Perspectives on spatial data analysis*. London: Sage.
———. 2002. *Geographically Weighted Regression: The analysis of spatially varying relationships*. Chichester: Wiley.
Fotheringham, A. S., and D. Wong. 1991. The modifiable areal unit problem in multi-variant statistical analysis. *Environment and Planning A* 23:1025–44.
Gatrell, A. C. 1985. Any space for spatial analysis? In *The future of geography*, edited by R. J. Johnson, 190–208. London: Methuen.
Getis, A., and J. K. Ord. 1996. Local spatial statistics: An overview. In *Spatial analysis: Modelling in a GIS environment*, edited by P. Longley and M. Batty, 261–77. Cambridge: GeoInformation International.
Goodchild, M. F., and N. S-N. Lam. 1980. Areal interpolation: A variant of the traditional spatial problem. *Geo-Processing* 1:297–312.
Gregory, I. N. 2002a. Time variant databases of changing historical administrative boundaries: A European comparison. *Transactions in GIS* 6:161–78.
———. 2002b. The accuracy of areal interpolation techniques: Standardising 19th and 20th century census data to allow long-term comparisons. *Computers, Environment and Urban Systems* 26:293–314.
———. 2003. *A place in history: A guide to using GIS in historical research*. Oxford: Oxbow.
———. 2005. The Great Britain Historical GIS. *Historical Geography* 33:132–34.
Gregory, I. N., C. Bennett, V. L. Gilham, and H. R. Southall. 2002. The Great Britain Historical GIS Project: From maps to changing human geography. *Cartographic Journal* 39:37–49.
Gregory, I. N., and P. S. Ell. 2005. Breaking the boundaries: Integrating 200 years of the census using GIS. *Journal of the Royal Statistical Society, Series A* 168:419–37.
Gregory, I. N., K. K. Kemp, and R. Mostern. 2001. Geographic information and historical research: Current progress and future directions. *History and Computing* 13:7–24.
Gregory, S. 1976. On geographical myths and statistical fables. *Transactions of the Institute of British Geographers: New Series* 1:385–400.
Holdsworth, D. 2002. Historical geography: The ancients and the moderns—generational vitality. *Progress in Human Geography* 26:671–78.
———. 2003. Historical geography: New ways of imaging and seeing the past. *Progress in Human Geography* 27:486–93.
Johnston, R. J. 1983. *Philosophy and human geography: An introduction to contemporary approaches*. London: Edward Arnold.
Keat, R. 1981. *The politics of social theory*. Oxford: Blackwell.
Kennedy, L., P. S. Ell, E. M. Crawford, and L. A. Clarkson. 1999. *Mapping the Great Irish Famine, an atlas of the famine years*. Dublin: Four Courts Press.
Knowles, A. K. 2000. Introduction: Historical GIS: The spatial turn in social science history. *Social Science History* 24:451–70.
———, ed. 2002. *Past time, past place: GIS for history*. Redlands, CA: ESRI Press.
———, ed. 2005. Emerging trends in historical GIS. Theme issue of *Historical Geography* 33.
Langran, G., and N. Chrisman. 1988. A framework for temporal geographic information. *Cartographica* 25:1–14.
Langton, J. 1972. Potentialities and problems of adopting a systems approach to the study of change in human geography. *Progress in Geography* 4:125–79.
Lawton, R. 1968. Population changes in England and Wales in the later nineteenth century: An analysis of trends by registration district. *Transactions of the Institute of British Geographers* 44:55–74.
Lee, C. H. 1991. Regional inequalities in infant mortality in Britain, 1861–1971: Patterns and hypotheses. *Population Studies* 45:55–65.
Massey, D. 1999. Space-time, "science" and the relationship between physical geography and human geography. *Transactions of the Institute of British Geographers: New Series* 24:261–76.
McMaster, R. B., and P. Noble. 2005. The U.S. National Historical GIS. *Historical Geography* 33:130–32.
Mokyr, J. 1983. *Why Ireland starved: A quantitative and analytical history of the Irish economy, 1800–1850*. London: HarperCollins.
Openshaw, S. 1984. *The modifiable areal unit problem*. Concepts and Techniques in Modern Geography, 38. Norwich: Geobooks.
Openshaw, S., and P. Taylor. 1979. A million or so correlation coefficients: Three experiments on the modifiable areal unit problem. In *Statistical applications in the spatial sciences*, edited by N. Wrigley, 127–44. London: Pion.
O'Sullivan, D., and D. J. Unwin. 2003. *Geographical information analysis*. Chichester: Wiley.
Peuquet, D. J. 1999. Time in GIS and geographical databases. In *Geographical Information Systems: Principles, techniques, management and applications*. 2nd ed., edited by P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind, 91–103. Chichester: Wiley.
Sack, R. D. 1980. *Conceptions of space in social thought*. London: Macmillan.
Taylor, P. J. 1976. An interpretation of the quantification debate in British geography. *Transactions of the Institute of British Geographers: New Series* 1:129–44.
———. 1990. Editorial comment: GKS. *Political Geography Quarterly* 9:211–12.
Vanhaute, E. 2005. The Belgium Historical GIS. *Historical Geography* 33:136–39.