USING PROBATE DATA TO DETERMINE HISTORICAL MALE OCCUPATIONAL STRUCTURES

Sebastian A.J. Keibek

Cambridge Group for the History of Population and

Social Structure & Queens' College

sk571@cam.ac.uk

ABSTRACT

Leigh Shaw-Taylor and Sir E.A. Wrigley recently published new estimates for the male occupational structure of England and Wales. Their pre-census figures are based on parish register data, but before Rose's Act of 1812, parish registers offer occupational information only in a sample of parishes, and are virtually silent about employments before 1690. This paper examines how the gaps in the parish register data can be filled using a data source which offer more universal coverage and goes back much further in time: probate records. It demonstrates how an, at first sight, critical deficiency of probate data, namely their severe bias towards capital-intensive and/or relatively well-paying occupations, can be overcome by using parish register data for calibration.

INTRODUCTION

The 'Occupational Structure of Britain 1379-1911' research programme has been running since 2003, led by Leigh Shaw-Taylor and Sir E.A. Wrigley. During those years, they and many other members of the Cambridge Group for the History of Population and Social Structure have collected, transcribed, standardised, and analysed large quantities of historical information. Current estimates of historical *male* occupational structures resulting from these data and analyses have recently been published in the fourth edition of the Cambridge Economic History of Modern Britain.[1] They are essentially based on two data sources: national censuses and, before 1841, parish registers, more specifically, the occupations of fathers as recorded in Anglican baptism records. The parish register data, on which the pre-census estimates are based, are an extremely detailed and valuable source of male occupational information.

This is true in particular after 1812, when Rose's Act made registering the occupations of fathers compulsory for Anglican baptisms, generating the data for what has rightly been called an occupational quasi census for c.1817.[2] The coverage of this quasi census is essentially universal and so geographically fine-grained that it allows for occupational analyses not only on the national level, but for individual counties, hundreds, census registration districts, or even individual parishes and chapelries. Much of the discussion on long-run economic developments and the industrial revolution has been focused on the national level, since that is the level at which Nick Crafts and Knick Harley's national account approach delivers its quantitative insights, as does the new work by Broadberry et al.[3] Quantitative occupational data at sub-national levels offer the opportunity to put spatial relationships back in the consideration of long-run economic development and the industrial revolution. Crafts and Harley rightly noted that 'regional development varied considerably and that exploring this diversity offers the potential of a set of quite different and valuable insights into the experience of the industrial revolution.'[4] Using occupational information to generate a quantitative understanding of regional developments is therefore one of the great promises of the Occupational Structure project.

Before Rose's Act too, reliable occupational data can be found in baptism registers for a significant number of parishes, all of which have been collected by the Cambridge Group. It is these data on which the eighteenth-century estimates are based. But parish registers also suffer from a number of important weaknesses as sources of historical occupational information. Firstly, an obvious deficiency regarding all pre-1851 data is that female occupations were not or not reliably recorded. Several other members of the Cambridge Group are addressing this – given the paucity of reliable historical sources – highly challenging issue; it is not addressed in this paper, which is solely concerned with male occupational structures. Secondly, many of the fathers captured in baptism registers were given the

---

[1] Shaw-Taylor and Wrigley, 'Occupational structure and population change', in Floud, Humphries, and Johnson (eds), *The Cambridge Economic History of Modern Britain. Volume 1. Industrialisation, 1700-1870* (Cambridge: Cambridge University Press, 2014), pp. 53-88.

[2] 1812, 52 Geo. III, c.146, *An Act for the better regulating and preserving Parish and other Registers of Births, Baptisms, Marriages, and Burials, in England*; for an extensive discussion of the act, see Basten, 'From Rose's *Act* to Rose's *Bill*: a reappraisal of the 1812 Parish Register Act', *Local Population Studies,* 76 (2006), p. 43. For the c.1817 quasi census, see Kitson *et al*, 'The creation of a "census" of adult male employment for England and Wales for 1817' (Cambridge, 2012), http://www.econsoc.hist.cam.ac.uk/docs/CWPESH%20number%204%20March%202012.pdf.

[3] Crafts, *British economic growth during the Industrial Revolution* (Oxford: Clarendon, 1985); Crafts and Harley, 'Output growth and the British Industrial Revolution: a restatement of the Crafts-Harley view', *The Economic History Review,* 45:4 (1992); Broadberry *et al*, *British economic growth, 1270-1870* (Cambridge: Cambridge University Press, 2015).

[4] Crafts and Harley, 'Restatement', p.721.

occupational descriptor 'labourer'; this tells us that they worked for an employer and were probably mainly remunerated by money wages, but it does not tell us what kind of work they did and to what occupational sector they should be attributed. This issue is addressed in another working paper.[5] Thirdly, parish registers – and, for that matter, most other contemporary sources – typically describe men with a single occupational denominator, whereas early modern historians generally presume that most pre-industrial men engaged in so-called by-employments in addition to their stated occupation.[6] If these were really as ubiquitous and important as has generally been assumed, an early modern occupational structure based on principal employments alone is unlikely to adequately represent the economic activities of contemporary men.[7] This issue is addressed in a separate working paper.[8] The current paper is aimed at exploring a solution to the fourth and final problem with parish registers as a data source: their incomplete coverage.

---

[5] Keibek, 'Allocating labourers to occupational (sub-)sectors using regression techniques' (working paper, Cambridge, 2017), http://www.econsoc.hist.cam.ac.uk/docs/CWPESH_number_27_March_2017.pdf.

[6] See, for example, Thirsk, 'Seventeenth-century agriculture and social change', in her Thirsk, 'The rural economy of England: collected essays', (London: Hambledon, 1984), pp. xiv,420p (p. 211); Everitt, 'Farm labourers', in Finberg (ed.), *The agrarian history of England and Wales. Vol. IV, 1500-1640* (Cambridge: Cambridge University Press, 1967), pp. 428-29; Overton *et al*, *Production and consumption in English households, 1600-1750* (London: Routledge, 2004), pp. 76-77; Woodward, 'Wage rates and living standards in pre-industrial England', *Past & Present*:91 (1981), pp. 39-42; Hey, 'A dual economy in south Yorkshire', *The Agricultural History Review,* 17:2 (1969), p. 113; Frost, 'Yeomen and metalsmiths: livestock in the dual economy in south Staffordshire 1560–1720', *The Agricultural History Review,* 29:1 (1981), p. 40; Holderness, 'Rural tradesmen 1660-1850: a regional study in Lindsey', *Lincolnshire History and Archaeology,* 7 (1972), pp. 77-83; Rowlands, *Masters and men in the West Midland metalware trades before the Industrial Revolution* (Manchester: Manchester University Press, 1975), pp. 42-43.

[7] For an expression of this sentiment, see: Swain, *Industry before the Industrial Revolution: north-east Lancashire c.1500-1640* (Manchester: Manchester University Press for the Chetham Society, 1986), p. 207. In his review of the book, James Rosenheim commended Swain for exposing 'the futility of reliance on occupational information to assess the structure of the early-modern labor force.', see: Rosenheim, 'Review of Swain's "Industry before the Industrial Revolution"', *Albion,* 20:1 (1988), pp. 99-100. See also: Clarkson, *The pre-industrial economy in England, 1500-1750* (London: Batsford, 1971), p. 77; Pahl, *Divisions of labour* (Oxford: Blackwell, 1984), p.47.

[8] Keibek, 'By-employments in early modern England and their significance for estimating historical male occupational structures' (working paper, Cambridge, 2017), http://www.econsoc.hist.cam.ac.uk/docs/CWPESH_number_29_March_2017.pdf

THE PROBLEM OF INCOMPLETE COVERAGE

Baptism registers only provide occupational information for a sample of parishes before Rose's Act. Coverage was particularly low for the mid and late eighteenth century, as shown in Map 1, and close to non-existent outside London before 1700.
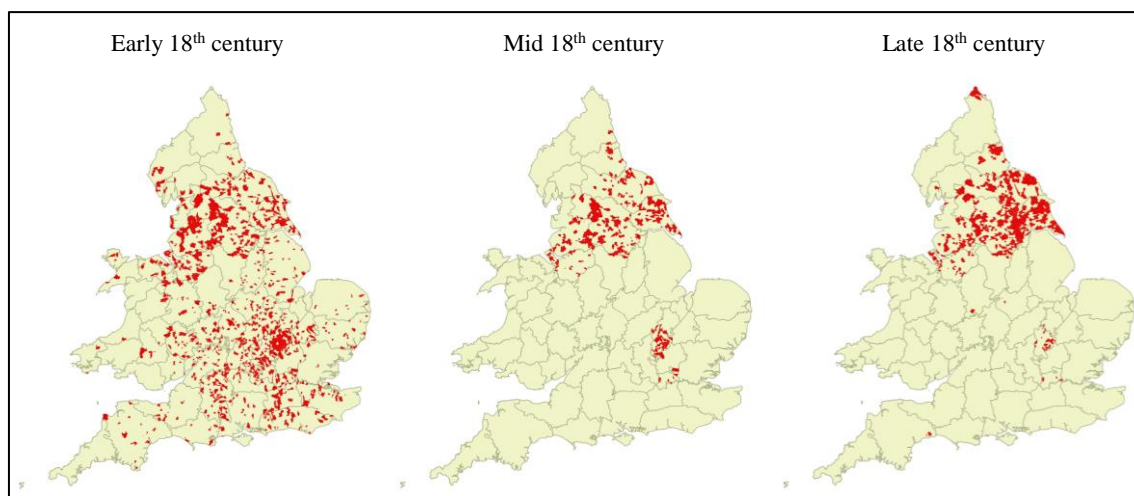
Map 1. Registration of male occupations in English and Welsh baptism registers over the course of the eighteenth century

Note: Parishes and chapelries in which male occupations were reliably recorded are indicated in red.
Data source: Parish register occupational database, created by the Cambridge Group.

In the early eighteenth century, only eleven per cent of all parishes recorded occupational information for a period of one or more years. For the mid and late eighteenth century, the figures are even lower, at only three and four per cent respectively. For occupational estimates at larger geographic scales, such as for England and Wales as a whole, small samples are not necessarily a problem, as long as the sample of parishes is representative of the population from which it is taken. However, as Map 1 shows, the samples are geographically non-random. Even in the relatively well-covered early-eighteenth century, urban areas are overrepresented, as are certain regions like industrialising Lancashire and the West Riding of Yorkshire, with other areas such as Wales, the South-West and East of England, the North, and several counties in the West Midlands covered hardly at all. Average coverage across the totality of England and Wales may have been eleven per cent in the early eighteenth century, but in two out of every three counties, coverage was below ten per cent. In the mid and late eighteenth century, lack of coverage was even more pronounced, with eight out of every nine counties below the ten per cent mark, whilst not a single parish in which occupations were registered could be found in four out of every five counties.

Even within counties, occupational structures were very far from uniform, so at this lower geographical level the sampled parishes cannot simply be presumed representative either. A case example can make that clear. Cheshire is amongst the counties with the highest parish register coverage, with occupations recorded reliably in thirty-three per cent of parishes in the early, nineteen per cent in the mid, and twenty-one per cent in the late eighteenth century. But the occupational structures derived directly from parish register data for these periods exhibit peculiar developments over time, as Figure 1 demonstrates: a sudden and sharp increase in the size of the tertiary sector between c.1725 and c.1755; an unexpected, gradual decline of that sector in the years thereafter; relatively rapid decline of the agricultural sector before the industrial revolution, followed by slight

growth after c.1785; sharp decline of the non-textiles portion of the secondary sector after c.1785. One reason for such unexpected developments is that the composition of the sample changes over time and the occupational bias is therefore not constant. Only in ten out of the thirty-one Cheshire parishes covered in c.1725, occupational information was also recorded in c.1755, and only seven parishes were covered in c.1725, c.1755, and c.1785. The rapid growth and subsequent decline of the tertiary sector over the 1715-1817 period in figure 2 may well be simply the result of 'tertiary-sector-heavy' parishes making up a greater share of the c.1755 than of the c.1725 and c.1785 datasets.
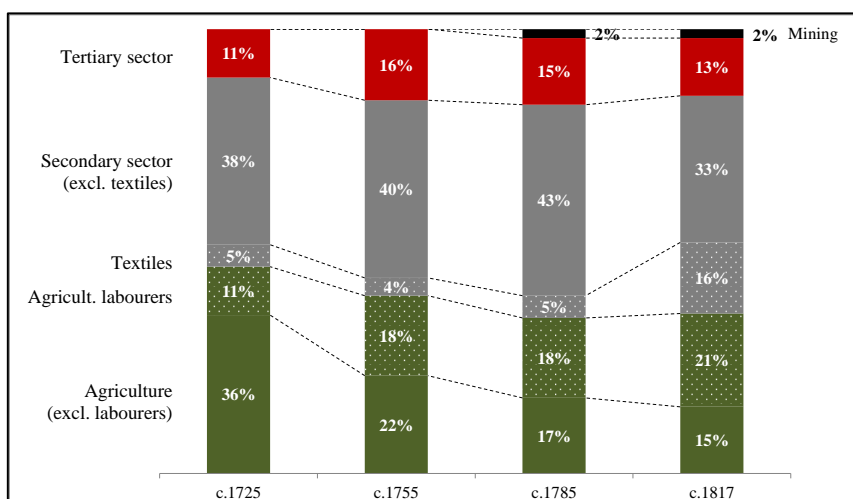

Figure 1. Cheshire's male occupational structure according to parish registers, c.1725-c.1817

Source: parish register data collected by the Cambridge Group.

Aware of these issues, Shaw-Taylor and Wrigley did not simply base their c.1710 national estimate directly on the parish register sample. Instead, they divided the pre-Rose's Act parishes along two axes: urban versus rural, and north-west England versus the rest. The underlying assumption is that the covered parishes, known not to be representative of all parishes, are much more likely to be representative of parishes on the same side of these divides. More reliable occupational structures can therefore be calculated for the rural and urban subsets of parishes, and for parishes inside and outside north-west England. These partial occupational structures are subsequently recombined to create a national one.

On a national level, and provided that, as in c.1710, the parish data have a reasonable geographic spread across the England and Wales, the reweighting method is likely to generate good results.[9] Potential regional biases within the urban and rural subsets will be much diluted on a national scale. The approach is less likely to generate reliable national results for the mid and late eighteenth century, as the required spread of parish register data across England and Wales is not available. For this reason, Shaw-Taylor and Wrigley refrained from including national estimates for these periods in the CEHMB chapter. The approach is also unreliable for smaller geographical levels, and Shaw-Taylor and Wrigley have therefore only used it nationally and on two large geographic regions. Had it been applied to a smaller region, such as a single county, the results might have been seriously misleading. For example, the Lancashire urban sample in c.1725 contains transport-dominated Liverpool but not manufacturing-dominated Manchester. This bias cannot be remedied by urban-rural reweighting. The

---

[9] Although the method is not without issues, as discussed in Keibek, The male occupational structure of England and Wales, 1600-1850 (PhD dissertation, University of Cambridge, 2017), pp. 175-7.

reweighting methodology cannot therefore fulfil one of the key goals of the Occupational Structure project: to provide quantitative insight into regional and local developments.

These limitations of the urban-rural reweighting method are the direct consequences of the limitations of the parish register data. No methodology based on these data alone can hope to generate reliable national estimates for the mid or late eighteenth century because parish registers coverage was simply too low. For the same reason, occupational information in pre-Rose's Act parish registers was too sparse to generate county or sub-county level occupational estimates. Furthermore, the parish-register-based national estimates imply that the shift from the primary to the secondary sector took place before 1700 rather than, as in the older national accounts literature, during the eighteenth century. Any proper analysis of industrialisation in Britain must therefore include the seventeenth and, potentially, earlier centuries. The lack of pre-1700 coverage disqualifies the parish register data for this purpose.

One or more new sources of data are required to fill the gaps. Shaw-Taylor, Jacob Field, and other members of the Cambridge Group therefore started collecting data from alternative sources several years ago, with the intention to use those data to overcome the problem of limited coverage. Prominent amongst these alternative sources where indexes to probate documents, of which Shaw-Taylor realised that their geographical comprehensiveness offered opportunities for complementing the parish register data. In this paper, an approach will be introduced which does exactly that. This new approach is capable of providing accurate and detailed male occupational estimates at the national, regional, and local level for the seventeenth, eighteenth, and early nineteenth century.

PROBATE DOCUMENTS AS OCCUPATIONAL DATA SOURCE

Many historical sources provide occupational information, but few do so reliably, for large areas, and for long time periods, as Paul Glennie has shown in his systematic analysis of historical data on men's trades.[10] There are really only two historical sources of sufficient detail, quality, and scope suitable for complementing the existing parish register data, namely records of court proceedings and probate documents. The former of these sources is very promising, in particular the recordings of Quarter Sessions and Assizes, which contain 'an abundance of occupational information about the people involved in the legal process in a variety of capacities'.[11] They go back to Elizabethan times and are available in large numbers; in recent research for a master's thesis, Tim Rudnicki was able to create a database of nearly 90,000 individual observations for Cheshire and Lancashire covering the early seventeenth to the early nineteenth century.[12] T.S. Cockburn has argued that they form a 'factually worthless' source of occupational information, but as Glennie has demonstrated, that judgement is far too harsh; the at first sight plausible criticism that low-status social groups are bound to be overrepresented in the court data is unjustified, as occupational information provided is not limited to

---

[10] Glennie, *'Distinguishing men's trades': occupational sources and debates for pre-census England* (Bristol: Historical Geography Research Group, 1990).

[11] Cockburn, 'The use of assize records as historical evidence', *Journal of the Society of Archivists,* 5 (1975), p. 224; Glennie, *Distinguishing men's trades*, pp. 41-3, in particular tables 3.4 and 3.5; David Cressy has demonstrated that, for an admittedly small sample, occupational information from court records is in agreement with that from parish registers; see Cressy, 'Occupations, migration and literacy in East London, 1580–1640', *Local Population Studies,* 5 (1970), p. 55.

[12] Rudnicki, *The male occupational structure of northwest England, circa 1600 to 1851* (Master thesis, University of Cambridge, 2015). For another encouraging example of the use of such data for occupational analyses, see Sugden, *An occupational analysis of the worsted industry, circa 1700-1851. A study of de-industrialization in Norfolk and the rise of the West Riding of Yorkshire* (PhD thesis, University of Cambridge, 2015), pp. 191-211.

suspected criminals but includes victims and witnesses.[13] The main problem with this data source is a practical one: relatively little of the available information has been digitised or can easily be digitised from printed documents. Where digital or printed information is available, it is often in the form of plain transcripts of the original calendars, rolls, and books. These transcripts are difficult to use because they are highly variable in form, which makes creating structured databases of occupational information from them a very time consuming process.

Probate documents have a great advantage here: their occupational information is relatively easily accessible because it is often contained in the indexes which have been created to provide access to the original wills and testaments, inventories, letters of administration, bonds, and codicils. Many of these indexes have been digitised, often in a highly structured way, splitting off occupational from other items of information such as the decedent's name and domicile and the date at which the document was created. Even when the latter has not yet been done, the information is often presented in text strings of a fairly consistent nature, making it relatively easy to extract the required information.

Shaw-Taylor, understanding the need for complementary data and recognising the promise of the probate source, started collecting testamentary indexes on a large scale in 2007. As Glennie had before him, he understood that the strengths and weaknesses of probate and parish register data were, in many ways, complementary, with the former being socially selective but providing excellent geographic coverage, whilst the latter were representative of society but geographically sparse.[14] Therefore, Shaw-Taylor saw much potential in trying to find a way to use parish registers to, somehow, reduce or even fully remove the social bias in the probate data. He found many county record offices prepared to supply such data in an electronic form. These were converted into a database, coded into the PST system, and linked to the GIS by Jacob Field, Gil Newton, and Ros Davies. These data were made available to me when I joined the Cambridge Group, providing me with a flying start for my PhD research. It proved possible to approximately double the number of usable observations by complementing the existing database with additional or superior indexes.[15] As described in more detail elsewhere, they cover the vast majority of English counties and the whole of Wales, with most of the indexes going back well into the sixteenth century.[16] Although many men were not probated and the indexes do not provide an occupational descriptor for all men that were, the total number of usable occupational observations provided by the probate database created for this research ran to over 800,000. Nesta Evans proclaimed probate documents to be 'a more fertile source of [occupational] information than any other class of document', because of their 'sheer number'.[17] It is not surprising therefore that they have been used extensively by historians to analyse occupational developments.

---

[13] Glennie, *Distinguishing men's trades*, pp. 41-3; Rudnicki, *Northwest England*, pp. 19-32.

[14] Glennie's remarks on this are discussed below.

[15] I am tremendously grateful to Leigh Shaw-Taylor for collecting these data, to Jacob Fields for 'matching' them to the PST system of occupations and to GIS information, and to the Leverhulme Trust for providing the necessary financial support for this endeavour. I am also indebted to the many county record offices and the National Archives who provided me with additional or improved indexes.

[16] Keibek, *Male occupational structure*, ch. 5.

[17] Evans, 'The occupations and status of male testators in Cambridgeshire, 1551-1800' in Arkell, Evans, and Goose (eds), *When death do us part: understanding and interpreting the probate records of early modern England* (Oxford: Leopard's Head Press, 2000), p. 176.

EXAMPLES OF THE USE OF PROBATE DOCUMENTS AS AN OCCUPATIONAL DATA SOURCE – AND THE PROBLEMS WITH THEM

Most of the many examples of the use of occupational information in wills, inventories, and other testamentary documents in the historiography have a regional or local focus.[18] In many of them, probate documents are just one of several local data sources used, and few historians have aimed to create comprehensive occupational structures from probate data or indeed from other sources. But in two very recent studies, historians have been more ambitious and attempted to use probate data for establishing country-wide occupational estimates. In the first of these, Clark et al have used the index to probate records at the National Archives to calculate the share of men working in agriculture and fishing in England for the period of the Interregnum, between 1652 and 1660.[19] Since responsibility for the probate process was temporarily transferred to the newly-established Court for the Proving of Wills and the Granting of Administrations in London during the Interregnum, the National Archive's index to those testamentary documents should, in principle, cover the whole country. Additionally, Clark et al used probate indexes for seventeen (parts of) counties, overwhelmingly situated in the southern half of the country, to generate a national estimate for the 1560-79 period. For both this and the 1652-60 period, Clark et al calculate that fifty-nine per cent of English men were involved in agriculture or fishing. Since modern-day data from developing countries show a linear relationship between the share of the population employed in farming and average income per person, Clark et al conclude that average income per person in early-modern England must also have remained stable between 1560 and 1660, in line with Gregory Clark's earlier work on the absence of any real improvement in living standards during the entire medieval and early-modern periods.[20]

A very recent working paper by Wallis et al is based on a database of probate records not dissimilar to the one on which my own research is built, originally constructed to analyse the development of medical professions during the early-modern period.[21] Wallis et al take issue with Clark et al's conclusions, arriving at figures close to Broadberry et al's social-tables-based occupational estimates,

---

[18] See, for example, Patten, 'Changing occupational structures in the East-Anglian countryside' in Fox and Butlin (eds), *Change in the countryside: essays on rural England, 1500-1900* (London: Institute of British Geographers, 1979), pp. 103-21; Patten, 'Urban occupations in pre-industrial England', *Transactions of the Institute of British Geographers,* 2:3 (1977), pp. 296-313; Ripley, 'Village and town: occupations and wealth in the hinterland of Gloucester, 1660-1700', *The Agricultural History Review,* 32 (1984), pp. 170-7; Hudson, 'Landholding and the organization of textile manufacture in Yorkshire rural townships c. 1660-1810' in Berg (ed.), *Markets and manufacture in early industrial Europe* (London: Routledge, 1991), pp. 267-71; Zell, *Industry in the countryside: Wealden society in the sixteenth century* (Cambridge: Cambridge University Press, 1994), pp. 116-9, particularly tables 5.1 and 5.2; Stobart, 'Geography and industrialization: the space economy of northwest England, 1701-1760', *Transactions of the Institute of British Geographers,* 21:4 (1996), pp. 681-96; Stobart, 'The economic and social worlds of rural craftsmen-retailers in eighteenth-century Cheshire', *The Agricultural History Review,* 52:2 (2004), pp. 141-60; Evans, 'Occupations and status', pp.176-88; Stobart, *The first industrial region: North-west England, c.1700-60* (Manchester: Manchester University Press, 2004), see in particular pp. 229-33; West, *Town records* (Chichester: Phillimore, 1983), pp. 126-7; Chalklin, *Seventeenth-century Kent; a social and economic history* ([London]: Longmans, 1965), appendix 2; Glennie, *A commercialising agrarian economy: late-medieval and early-modern Hertfordshire* (Master thesis, University of Cambridge, 1983); Glennie, *Distinguishing men's trades*, pp. 32-40; Sugden, *Worsted industry*, pp. 71-2, 249, 291; Sneath, *Consumption, wealth, indebtedness and social structure in early modern England* (PhD thesis, University of Cambridge, 2009).

[19] Clark, Cummins, and Smith, 'Malthus, wages, and preindustrial growth', *The Journal of Economic History,* 72:02 (2012), pp. 364-92.

[20] Clark, *A farewell to alms: a brief economic history of the world* (Princeton, N.J. ;Woodstock: Princeton University Press, 2007).

[21] Wallis, Colson, and Chilosi, 'Puncturing the Malthus delusion: structural change in the British economy before the industrial revolution, 1500-1800' (working paper, London, 2016), http://www.lse.ac.uk/economicHistory/home.aspx.

showing a clear decline of the primary sector share of the male labour force over the seventeenth century.

Historians working with occupational information from testamentary documents generally recognize that these 'do not record a representative cross-section of the early modern population'.[22] They deal with this issue in a variety of ways. Some attempt to limit the effects of social bias by using the data in a specific way; John Patten, for example, did not attempt to derive something resembling an occupational structure from the probate data for rural East-Anglia but merely used them to identify whether particular occupations were or were not present in certain parishes, and how this changed over time.[23] Others use probate documents only as a secondary occupational source; Peter Ripley, for example, used court records as the main source of information on Gloucestershire working men, and employed probate documents only to complement that information and to provide more insight into the details and scale of their work-related activities.[24] Some historians do use probate data to establish comprehensive occupational structures but are open about the unavoidably limited accuracy of these estimates; Jon Stobart, for example, used probate documents as his prime data source on the Diocese of Chester's occupational structure and geography, it being the only 'spatially comprehensive' source available, but freely acknowledged that he could only provide a 'socially selective picture of male employment' in this way.[25] Others set out to justify the reliability of their probate-based figures. The two examples of national probate-based estimates discussed above fall into that category.

Clark et al attempt to validate their estimated agricultural labour shares by calculating the portion of men who left a trace in the probate record as a function of population density. They provide a number of local examples but since these all share the same rationale and approach, I will focus on one specific example here: Essex in 1801. Clark et al reason that since low-density rural parishes must have been overwhelmingly agricultural and high-density urban parishes must have been overwhelmingly non-agricultural, we should expect variations in the share of men being probated as a function of population density if men employed in agriculture had a lower or higher chance of being probated than men employed outside of agriculture. Since they do not find such differences in Essex in 1801, they conclude that agricultural workers were, apparently, not overrepresented and, therefore, that their probate-based estimates are reliable.[26] The problem is that the male agricultural population consisted largely of two groups: farmers who, as will be shown, had a very high probability of being probated and farm labourers for whom, as will also be shown, this probability was very low.[27] The labourer-to-farmer ratio differed greatly over the country and over time, so the average probability of being probated also varied greatly, both spatially and temporally, for the overall agricultural population. Similarly, non-agricultural men worked in a highly diverse group of occupations. As will be shown, some non-agricultural occupations, such as tanners and merchants, were quite likely to be probated whilst for others, such as weavers and domestic servants, this was highly unlikely. Since the occupational composition of the non-agricultural share of the labour force varied over time and place,

---

[22] Zell, 'The social parameters of the probate records in the sixteenth century', *Historical Research,* 57:135 (1984), p. 113.

[23] Patten, 'Changing occupational structures'. Note that Glennie rightly took issue with even this limited use of occupational information from wills, particularly when used for inter-county comparisons, because of differences in probate coverage and parish size between counties and the potential effects of the 'rural parish' definition. Glennie, *Distinguishing men's trades*, pp. 37-9.

[24] Ripley, 'Village and Town'.

[25] Stobart, *First industrial region*, p. 40-2.

[26] Clark *et al*, 'Malthus', pp. 369-71.

[27] Similarly, non-agricultural men worked in a highly diverse group of occupations. As will be shown, some non-agricultural occupations, such as tanners and merchants, were highly likely to be probated and others, such as weavers and domestic servants, highly unlikely.

so did the average probability of being probated for non-agricultural workers. In other words, what was true in Essex in 1801 was, in all likelihood, not true in other counties and/or time periods. As will be shown in Table 2, there happens to be a good match between the probate-based estimate and the actual occupational structure in Essex in the early eighteenth century, but Essex is very much the exception here. Had another county been chosen, the results would have been quite different. In Lancashire, for example, men working in agriculture were 2.5 times as likely to be probated as non-agricultural men. In contrast, in Bedfordshire, men working outside agriculture were 1.8 times more likely to be probated than those working in the local agricultural sector, which consisted predominantly of farm labourers, only very few of whom left testamentary evidence.

A second, smaller problem with the Clark et al paper is that they provide agricultural labour share estimates for two moments in time, 1560-79 and 1652-1660, but the collections of probate records on which these shares are based are quite different. Probate documents in the earlier period were proved in local church courts, whilst those in the later, Interregnum period were proved at the new, central Court for the Proving of Wills and the Granting of Administrations in London. Clark et al's implicit assumption seems to be that the dataset of centrally-proved probate documents in the Interregnum period was as socially and occupationally diverse as that based on documents proved in the local church courts outside the Interregnum. This assumption is incorrect. The annual number of probated decedents was markedly smaller during the Interregnum than just before and after, particularly in counties far away from the London court, as Figure 2 shows. It is therefore not surprising that the occupational composition of the locally-proved probate record just before and after the Interregnum also differed from that of the centrally-proved record during the Interregnum, particularly again in counties far removed from London, as Figure 3 shows. The social composition of the set of centrally proved probate documents during the Interregnum differed from those proved in local church courts outside the Interregnum. This makes comparisons between the 1560-79 and the 1652-60 datasets unreliable.
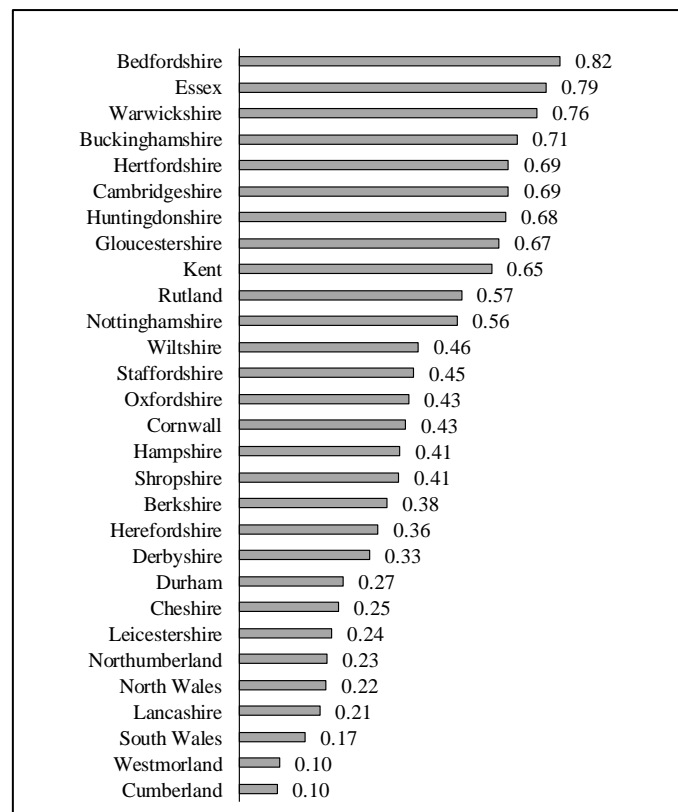
Figure 2. The ratio between the annual number of probated individuals per county in the 1652-60 period and the equivalent number in the decades preceding and following that period

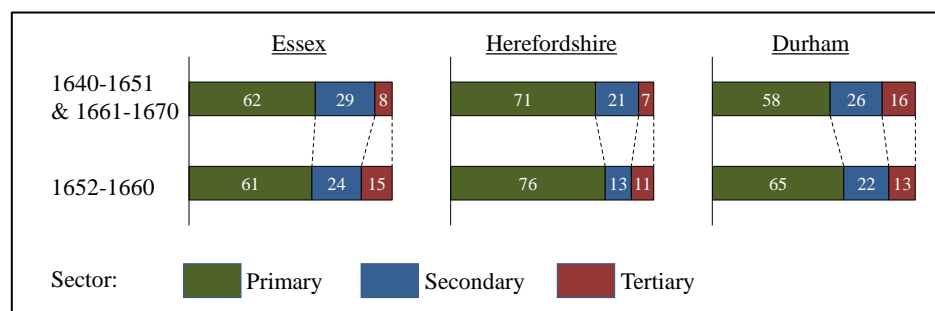*Sources*: The National Archives; county record offices.

Figure 3. Comparison of sectoral shares in the probate record during and just outside the Interregnum in selected counties

Sources: The National Archives; county record offices.

Wallis et al, try to justify the reliability of their results in three ways. Firstly, they compare the probate-derived estimate to Broadberry et al's figures derived from social tables around 1700, apparently failing to appreciate that these figures are themselves highly problematic – as I have discussed elsewhere.[28] Secondly, Wallis et al provide a comparison with occupational estimates based on apprenticeship records but note that these form a less reliable source than the probate records themselves, concluding that where the two sources do not match, it must be because the

---

[28] Keibek, *Male occupational structure*, pp. 5-8.

apprenticeship-records-suggested trends are incorrect.[29] Thirdly, they econometrically model 'counterfactual estimates' which, they claim, provide a 'strong robustness check' on their probate-based estimates. But this would seem to be a very optimistic claim, as the econometrical estimates do not provide an actual check on the effects of social bias itself but only on the potential effects of *changes* in social bias over time, and only to the degree that such changes are driven by a changing portion of deaths covered by the probate record. Furthermore, they only provide a check on *trends* over time, not on the actual *values* of occupational shares.[30]

Rather than attempting to justify probate-derived estimates via theoretical and incomplete model calculations or comparisons to other estimates which are themselves problematic, in this paper the problem of social bias is tackled head on. A new approach aimed at resolving rather than playing down the probate record's non-representativeness problem is introduced below. But before doing that, it is necessary to analyse the nature and severity of that problem in more detail.

OCCUPATIONAL BIAS IN THE PROBATE RECORD

Only a minority of adult men left a will, inventory, or another probate document. Using Wrigley's recent work on county populations and the Cambridge Group's family reconstitution research, it is possible to calculate shares of adult men who were probated, the results of which are presented in Table 1 for several counties and time periods. As the table shows, these shares exhibited significant spatial and temporal variation. A population-weighted average of the table's figures indicates that roughly four out of five English male householders who died in the 1600-1850 period left no trace in probate documents.

Table 1. Approximate share of male householders probated in selected counties/areas and decades

County/area	1630-9	1680-9	1730-9	1780-9	1830-9
Cheshire	28%	28%	24%	13%	19%
Cornwall	19%	20%	27%	15%	14%
Cumberland	15%	22%	43%	20%	24%
Durham Diocese	13%	15%	12%	10%	13%
Essex	19%	10%	12%	7%	6%
Gloucestershire	16%	22%	16%	8%	9%
Hampshire	29%	25%	38%	14%	12%
Leicestershire	43%	34%	35%	23%	15%
Wiltshire	28%	27%	28%	15%	11%

Sources: Several county record offices; online databases of probate records from the consistory courts of York (from the Borthwick Institute, via Origins.net) and Canterbury (at the National Archives); Wrigley et al, *Family reconstitution*, tables 6.19 (p. 290), A9.1 (pp.614-5) and 5.3 (p. 149); Wrigley, *Early English censuses*, table 4.1 (pp. 104-5); Wrigley and Schofield, *Population history*, pp. 493-526.
Notes: Male householders are defined here as men older than the male average age of marriage. Age specific mortality rates ($_{10}M_x$) were derived from probabilities of dying per age interval ($_{10}q_x$) as provided by Wrigley et al, employing the relationship $_nq_x = 2 \times n(_nM_x) / [2 + n(_nM_x)]$.

The men that *were* probated were not a random subset of the adult male population. The church courts could charge for grants of probate if the decedent's estate was valued at five pounds or more, and

---

[29] Wallis *et al*, 'Puncturing', p. 28.
[30] Ibid, pp. 29-32.

therefore had a financial incentive to encourage application for probate in these instances. But they could not force anyone to make a will and/or inventory.[31] It is therefore not surprising that many decided to avoid the expense and bother of applying for probate. Since the trade-off between, on the one hand, the cost and effort of the probate process and, on the other hand, its value in case of disputes over the estate, was more likely to be positive for high-value than low-value estates, the former are overrepresented in the probate record. As the church courts were not allowed to charge for estates below the five pounds' threshold, they may have actively discouraged probate in these instances – even though they could not refuse to process testamentary documents presented to them. It is unsurprising, then, that men who were wealthy and/or engaged in activities that required much capital – such as farmers, merchants, tanners, and brewers – were more likely to make probate than men who were poor or whose occupation required little or no capital – such as tailors, weavers, domestic servants, or labourers. The occupational bias of the probate record is clear from Figure 4, in which the relative share of men making probate has been calculated for a sample of occupations by comparing probate data with parish register data from the same area and time period. As the figure shows, the probability of an early-nineteenth-century farmer in Cheshire leaving a probate document was four times higher than a butcher, twelve times higher than a weaver, and twenty-five times higher than a labourer.[32]
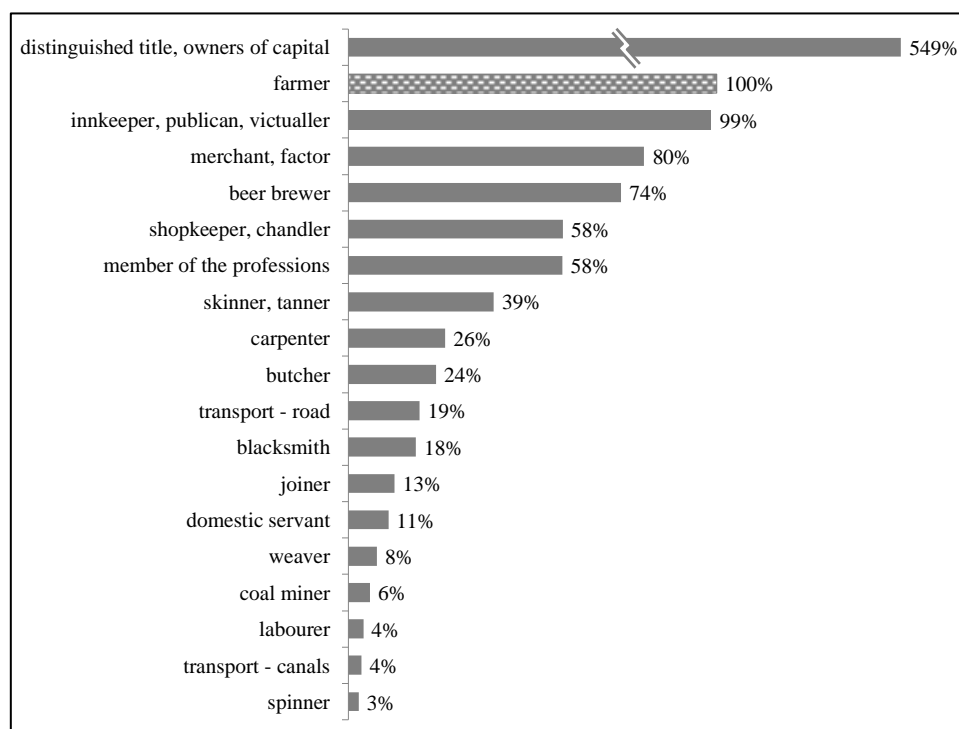

Figure 4. The relative chance of being probated for a sample of occupations in Cheshire, c.1817 (relative to farmer = 100%)

---

[31] Arkell, 'The probate process', p. 12.

[32] Other historians have come to comparable conclusions about the severity of the occupational and social bias of the probate record, based on comparisons with other unbiased sources of occupational information, For example, Lindert, 'An algorithm for probate sampling', *The Journal of Interdisciplinary History,* 11:4 (1981), pp. 662-3; Churchley, *Differing responses to an industrialising economy: occupations in rural communities in the heart of England from the restoration to the railway age (c. 1660 - c. 1840)* (DPhil thesis, University of Birmingham, 2010), p. 46, table 2.1; Oestreicher, 'The counted and the uncounted: the occupational structure of early American cities', *Journal of Social History,* 28:2 (1994), pp. 354-5.

Notes: The figures presented are the direct result of a comparison between the number of observations by occupation in parish registers and probate data for Cheshire. They represent relative rather than absolute probabilities of being probated, with farmers artificially set at 100 per cent to facilitate easy comparison. The parish register data covered the 1813-20 interval, the probate data were taken from a slightly longer time period, to create large enough samples: 1802-1832.

As a result, despite the popularity of probate documents as an occupational data source, occupational structures derived from them are highly unreliable. In Figure 5, male secondary and tertiary sector shares according to the probate record are compared to reliable figures derived from parish registers in the 1813-20 period, with each data point representing a county for which (sufficient numbers of) probate data are available. As the figure shows, the probate-derived figures bear little relationship to reality. A statistical test of correlation confirms the visual impression of Figure 5, with p-values of .70 and .14 for the secondary and tertiary sector respectively. Figure 5 also shows that the probate data lead to underestimates of the secondary sector in most counties, and to overestimates of the tertiary sector in all but three counties.
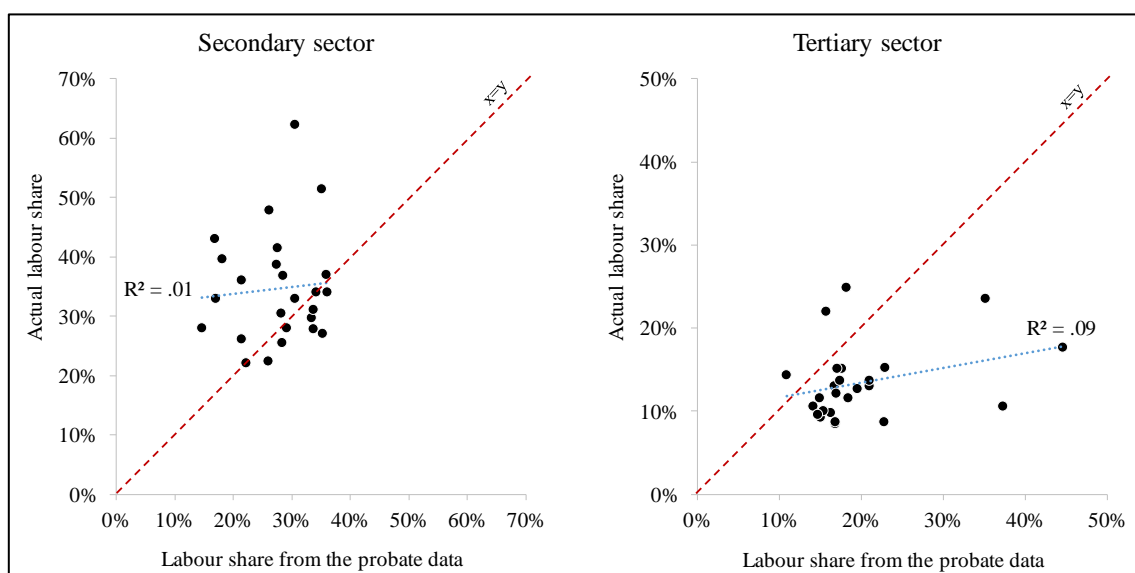

Figure 5. A county-based comparison between sectoral occupational share according to probate (x-axis) and parish register (y-axis) data (c.1817)

Notes: Each data point corresponds to a county or, in some cases, a combination of counties. Only counties for which sufficient numbers of probate data were available were included. These were: Bedfordshire, Berkshire, Buckinghamshire, Cambridgeshire, Cheshire, Cornwall, Cumberland, Devon, Dorset, Durham and Northumberland, Essex, Gloucestershire, Hampshire, Hertfordshire, Huntingdonshire, Lancashire, Leicestershire, Norfolk, North Wales, Nottinghamshire, Oxfordshire, South Wales, Suffolk, Westmorland, and Wiltshire. The parish register data cover the 1813-20 interval. The probate data were taken from a slightly longer time period, to create sufficiently large samples: 1802-1832. Labourers were allocated to sectors using the method and results presented elsewhere.[33] The red 'x=y' equilibrium line was included to visualise the number of counties for which probate data exaggerated or underestimated sectoral labour shares. The blue linear fit line was included, with its R^2 value, as an indication of the goodness of fit between probate-suggested and actual occupational shares.

---

[33] Keibek, 'Allocating labourers'.

Zooming in on smaller geographical areas and on sub-sectors confirms the problem of non-representativeness of the probate evidence, as demonstrated in Table 2. As the table shows, the degree of non-representativeness differed by county. For Essex, which Clark et al used to demonstrate the reliability of their estimates, the probate-derived estimates are fairly close to the actual occupational structure as determined from parish register data.[34] But in other counties the match is poor, for example in Lancashire and North Wales, in which the probate-derived figures underestimate the secondary sector by a factor of two, or in Durham and Northumberland, where they underestimate mining by a factor of five. In some sub-sectors, the match is particularly poor, for example in 'dealers and sellers' which is significantly overrepresented in the probate data in all counties, up to a factor of 4.6 in Durham and Northumberland. 'Services and professions' are also generally overrepresented, up to a factor of 2.5 in Lancashire. The probate data overstates the importance of the agriculture in counties dominated by family farming, but understates it in agrarian capitalist Bedfordshire, where the many farm labourers left few traces in the probate record. Indeed, according to the probate record, the share of labourers in Bedfordshire's male labour force declined from twenty-one per cent in 1621 to just three per cent in 1821 – entirely at odds with the rising importance of agrarian capitalism in this county.

Table 2. A comparison at sub-sectoral level between occupational structures according to probate and parish register data (c.1817)

	Bedfordshire			Lancashire below the Ribble			North Wales			Durham & Northumberland		
	Probate (%)	Par. reg. (%)	Factor (prob./par.)	Probate (%)	Par. reg. (%)	Factor (prob./par.)	Probate (%)	Par. reg. (%)	Factor (prob./par.)	Probate (%)	Par. reg. (%)	Factor (prob./par.)
Primary sector	**56.0**	**68.5**	**0.8**	**32.0**	**19.5**	**1.6**	**72.6**	**64.2**	**1.1**	**39.3**	**41.0**	**1.0**
Agriculture	55.9	68.1	0.8	31.3	15.6	2.0	69.7	54.5	1.3	36.0	22.2	1.6
Mining	0.0	0.0	-	0.6	3.6	0.2	2.6	9.3	0.3	2.8	17.6	0.2
Other	0.1	0.4	0.3	0.1	0.2	0.4	0.3	0.4	0.7	0.5	1.2	0.4
Secondary sector	**27.2**	**22.3**	**1.2**	**36.8**	**69.0**	**0.5**	**14.7**	**28.7**	**0.5**	**28.3**	**41.9**	**0.7**
Clothing	3.6	1.7	2.2	2.3	3.1	0.7	1.2	2.1	0.6	1.4	2.2	0.6
Footwear	2.0	3.1	0.6	1.8	2.8	0.6	1.4	3.5	0.4	2.4	4.3	0.6
Textiles	0.4	0.3	1.3	14.9	39.3	0.4	1.4	3.3	0.4	1.8	2.6	0.7
Metal trades & tools	2.4	2.2	1.1	2.9	4.7	0.6	1.9	3.5	0.6	3.7	5.3	0.7
Building	7.3	7.9	0.9	5.3	7.5	0.7	3.2	7.9	0.4	6.3	11.5	0.5
Other	11.6	7.1	1.6	9.6	11.5	0.8	5.7	8.5	0.7	12.7	16.0	0.8
Tertiary sector	**16.7**	**9.2**	**1.8**	**31.2**	**11.5**	**2.7**	**12.7**	**7.1**	**1.8**	**32.4**	**17.1**	**1.9**
Dealers and sellers	6.5	1.9	3.4	12.8	3.4	3.7	3.6	1.4	2.6	9.9	2.1	4.6
Services & professions	9.8	6.0	1.6	15.2	6.0	2.5	8.4	4.2	2.0	12.3	6.7	1.8
Transport*	0.4	1.3	0.3	3.2	2.1	1.5	0.6	1.5	0.4	10.3	8.2	1.3

	Essex			Wiltshire			Leicestershire			Cornwall		
	Probate (%)	Par. reg. (%)	Factor (prob./par.)	Probate (%)	Par. reg. (%)	Factor (prob./par.)	Probate (%)	Par. reg. (%)	Factor (prob./par.)	Probate (%)	Par. reg. (%)	Factor (prob./par.)
Primary sector	**55.1**	**63.0**	**0.9**	**50.1**	**56.8**	**0.9**	**51.6**	**36.0**	**1.4**	**70.5**	**60.2**	**1.2**
Agriculture	54.2	61.5	0.9	49.7	56.4	0.9	51.5	35.5	1.5	57.9	34.0	1.7
Mining	0.1	0.0	-	0.1	0.1	0.8	0.0	0.4	0.0	11.4	23.6	0.5
Other	0.8	1.5	0.5	0.4	0.3	1.3	0.1	0.1	0.6	1.3	2.6	0.5
Secondary sector	**28.0**	**24.0**	**1.2**	**33.5**	**33.8**	**1.0**	**32.6**	**52.8**	**0.6**	**20.8**	**30.2**	**0.7**
Clothing	1.3	1.2	1.1	1.5	1.4	1.1	9.6	27.5	0.3	0.8	1.2	0.7
Footwear	2.9	2.6	1.1	2.6	2.6	1.0	2.9	3.2	0.9	1.9	3.4	0.6
Textiles	0.8	0.5	1.5	5.5	8.4	0.7	3.1	3.3	0.9	0.6	0.3	2.1
Metal trades & tools	3.4	2.4	1.4	4.1	2.4	1.7	2.1	2.7	0.8	6.6	4.5	1.5
Building	6.9	7.7	0.9	8.9	10.5	0.9	4.8	6.3	0.8	5.3	12.2	0.4
Other	12.8	9.6	1.3	10.7	8.6	1.2	10.2	9.8	1.0	5.7	8.6	0.7
Tertiary sector	**16.9**	**13.0**	**1.3**	**16.4**	**9.4**	**1.7**	**15.8**	**11.2**	**1.4**	**8.8**	**9.7**	**0.9**
Dealers and sellers	6.0	2.4	2.5	6.9	1.8	3.8	3.9	2.2	1.8	2.5	1.3	1.9
Services & professions	9.4	8.1	1.2	8.5	5.9	1.4	11.0	6.4	1.7	5.6	6.2	0.9
Transport*	1.5	2.4	0.6	1.0	1.6	0.6	0.8	2.6	0.3	0.6	2.1	0.3

*excluding seamen

Notes: The parish register data cover the 1813-20 interval, the probate data were taken from a slightly longer time period, to create large samples: 1802-1832. Labourers were allocated to sectors using the method and results presented elsewhere.[35] The 'factor' column present a straightforward division of the two columns to its left, to provide a straightforward measure of the degree to which the probate data exaggerate or underestimate labour force shares.

---

[34] See the discussion about Clark *et al*'s choice of county, starting on page 29.
[35] Keibek, 'Allocating labourers'.

There can be only conclusion: the bias in the probate record makes occupational structures derived from it entirely unreliable. This would seem to disqualify probate documents as an occupational data source. Fortunately, this at first sight critical defect can be remedied, as discussed in the next section.

CORRECTING THE PROBATE RECORD FOR OCCUPATIONAL BIAS

The vastly different probabilities of being probated for men in different occupations, as presented in Figure 4 are more than merely a *measure* of the probate record's occupational bias; they are also at the heart of a *solution* for it. Their reciprocal values can be used as *calibration factors* with which to multiply the probate record to reconstruct the (missing) parish register record. How this works is best illustrated in a case example.

The index of Cheshire probate documents provides occupational information on nearly twenty-five-thousand male individuals who were probated between 1710 and 1830. For the purposes of this example, this 120-year period was divided into four intervals of thirty years, the mid-points of which corresponded with the mid-points of the parish register data, that is, 1725, 1755, 1785, and 1817. Thus, a temporal match was made between parish register and probate data.

These data can now be used for calculating accurate male occupational structures for the pre-Rose's Act time periods, in a three-step process, depicted schematically in Illustration 1 for the c.1725 data. In step 1, probate and parish register data from the same time period are matched geographically, to create a like-for-like basis for comparison between the two sources. Probate data are available for the whole of Cheshire, but parish registers recorded occupations in only thirty-one out of ninety-one parishes in c.1725. Only in these 'doubly-covered' parishes can a like-for-like comparison between probate and parish register data be made.

These parishes form the basis for calculating the calibration factors, in step 2. As discussed above, these are the reciprocal values of the (relative) probabilities of being probated:

$$[\text{calibration factor for occupation } i] = \frac{[\text{number of parish records with occupation } i]}{[\text{number of probated decedents with occupation } i]}$$

In step 3, the numbers of probated men per occupational group are multiplied with these calibration factors *in all parishes*. In the thirty-one 'doubly-covered' parishes, the result of this multiplication is, by definition, equal to the occupational structure according to the local parish register data. In the sixty other parishes, the result of the multiplication is a simulation of what the occupational structure derived from parish register data would have looked like *if such data had been available*. Thus, a calibrated, unbiased estimate of the contemporary male occupational structure is generated.
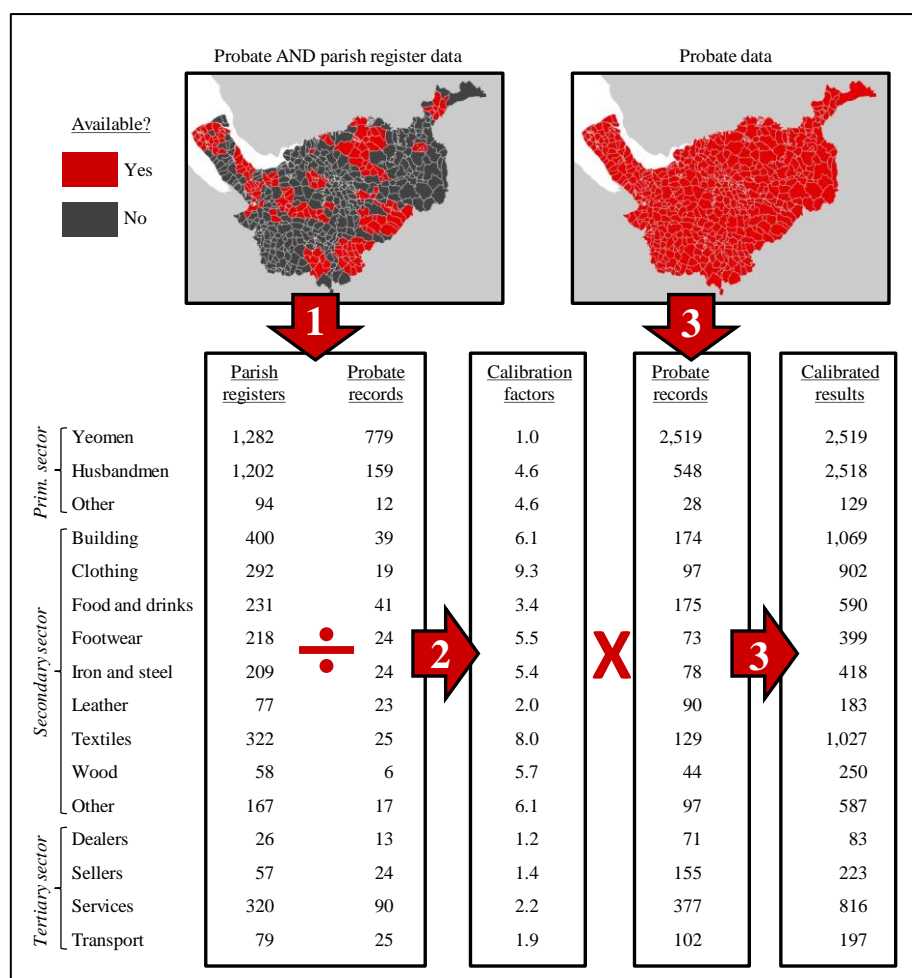
Illustration 1. A schematic depiction of the calculation of an unbiased
male occupational structure from probate records (Cheshire, c.1725)

Note: the numbers in the arrows refer to the steps of the calibration and correction process, described in the main text.

As discussed, geographical matching of probate and parish register data is required for determining reliable probate calibration factors. This is usually straightforward, but not always, and it is therefore necessary to discuss it in a bit more detail. Probate records almost always provide geographical information on the deceased, usually in the form of the name of the hamlet or township in which he lived. Generally, the information from parish registers is geographically less precise. It is available at the level of, what the Cambridge Groups has called, Anglican Registration Units (RUs). These usually coincide with an Anglican parish, but in some cases, certain chapelries within a parish recorded baptisms separately; in those cases, one parish gives rise to several RUs, one for each of the separately-recording chapelries, and one for the remainder of the parish. In rural areas, RUs generally contain several townships and hamlets.[36] Using a variety of sources of historical geographical information, the townships and hamlets mentioned in the probate data can be carefully 'mapped' onto RUs, to ensure that the two data sources are compared for the exact same geographical area when calculating the probate calibration factors. However, in large towns, the procedure is more complicated. The city of Chester was such a large town, and the only place in Cheshire for which a

---

[36] For 'The codebook of Anglican registration units', created by P. Kitson, see
http://www.geog.cam.ac.uk/research/projects/occupations/britain19c/anglicanregistration.html.

geographic match between probate and parish register data cannot be achieved in the manner described above. Like many larger towns, Chester was divided into several parishes. Not all of these Chester parishes recorded male occupations in the baptism register.[37] As Figure 6 shows, a substantial share of Chester's labour force is not covered by parish register data.
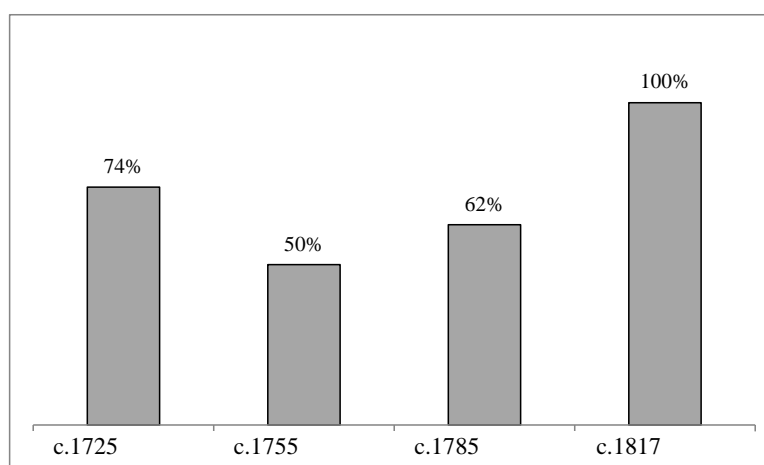

Figure 6. Estimated share of population in the Chester city parishes for which baptism register occupational data exists (c.1725-c.1817)

Sources: Wrigley, *Early Censuses*, table A.2.7; Lewis and Thacker, *A History of the County of Chester, Volume 5, Part 1, The City of Chester* (Woodbridge: Boydell and Brewer, 2003), pp. 90-7.

This would not be a problem in terms of geographically matching probate and parish data, if the former were specified at the level of parishes as well. The geographic data in probate records is usually more precise than the parish record data, because townships and hamlets were typically smaller than parishes, with several of them contained in one parish. In major towns like Chester however, the opposite is the case. Probate documents of men from Chester typically only mention that they lived in the Chester, and do not specify the decedent's parish within that town. This means that it is not possible to make a direct, like-for-like geographic match between parish and probate records for large towns like Chester.

One cannot simply presume that the occupational structure in the covered Chester parishes was similar to that in the city's non-covered parishes. It is likely that Chester, as many towns then (and now) had a non-uniform occupational topography, with certain trades concentrated in specific parts of the town, or even in one or two streets. Indeed, Figure 7 shows that the covered parishes in Chester were not representative of the whole town. In the left hand side of this figure, the probate calibration factors for Chester are compared to those of the rest of Cheshire for c.1817. Because these are post-Rose's Act parish register data, all Chester parishes provide occupational information, so a one-on-one comparison with the probate data is possible. The probate multipliers for Chester city turn out to be in line with those from rural Cheshire. For the c.1725 data, displayed on the right hand side of the figure, a one-on-one geographic match could, as discussed, be made for rural Cheshire but not for Chester city. Nevertheless, if the covered parishes in Chester city *had* been representative of the town as a whole, Chester's probate multipliers should have been in line with those for rural Cheshire. That is not the case. Ergo, the covered parishes in Chester are not representative of the whole town.

---

[37] For example, c.1725, such data were recorded in St Bridget, St Martin, St Mary on the Hill, St Michael, St Oswald, and St Peter but not in Holy Trinity, St John the Baptist, and St Olave.
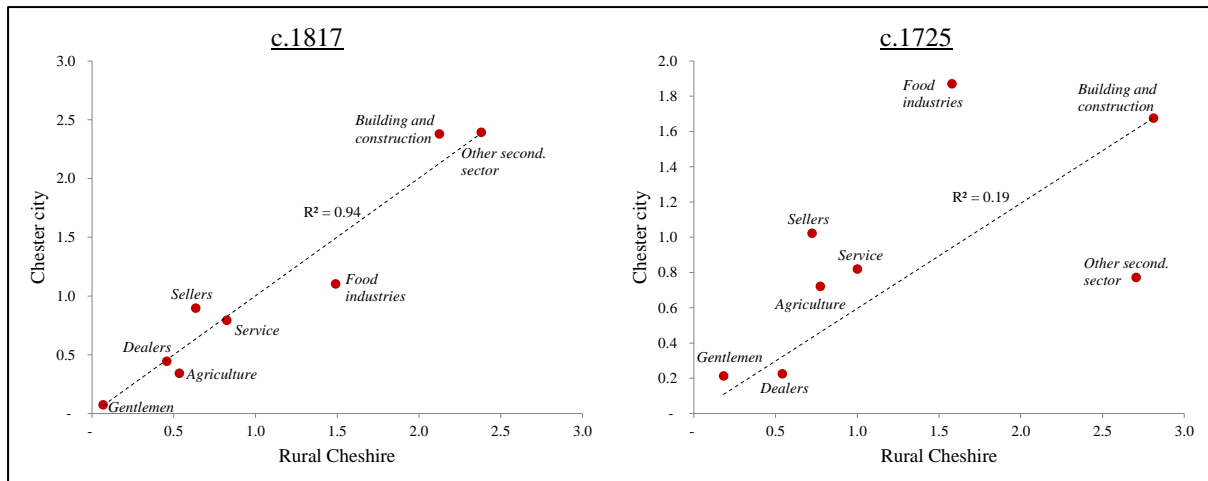
Figure 7. Probate calibration factors compared between rural Cheshire and Chester
(indexed; average male decedent = 1).

Note: only occupations encountered sufficiently frequently in the Chester city parish registers and probate documents to enable a statistically meaningful calculation of the calibration factors were included in the charts.

But, Figure 7 also provides the solution for this problem. As the chart on the left-hand side showed, the probate multipliers for c.1817 in Chester and the rest of Cheshire were comparable. That was, in all probability, also the case in c.1725. Consequently, the c.1725 probate multipliers derived from *Cheshire excluding Chester* can be applied to the probate data for *the city of Chester*, which cover the whole town. Thus, a reliable occupational structure is derived for the county as a whole.

The final result is Figure 8, providing an unbiased picture of male occupational developments in Cheshire in the eighteenth and early nineteenth centuries. The peculiar developments suggested by the parish register data in isolation, depicted in Figure 1, disappear once the probate record has been employed to reconstruct the occupational structure in the 'missing' parishes. The early, sudden fall in importance of the agricultural sector is replaced by a gradual decline throughout the period, accelerating during the industrial revolution. The rapid growth and inexplicable later decline of the tertiary sector are replaced by a more gradual and plausible development. The calibrated occupational structures for c.1755 and c.1785 differ particularly strongly from those based on parish registers alone.
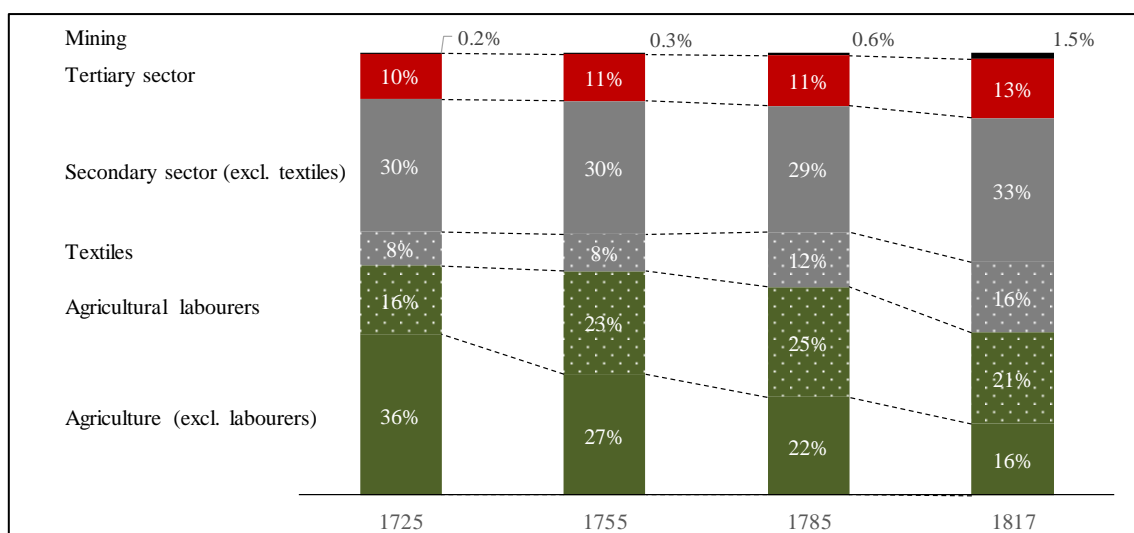
Figure 8. The occupational structure of Cheshire, from calibrated probate data (1725-1817)

Sources: probate databases; parish register database.

Two remarks regarding this calibration approach are worth making at this point. Firstly, it removes the probate data's occupational bias regardless of the *cause* of this bias. Although the overrepresentation of capital-intensive and 'wealthy' occupations is undoubtedly the principal cause of occupational bias in the probate record, other causes cannot be ruled out entirely. For example, age bias in the probate record would translate into occupational bias if a non-trivial share of men altered occupations during their lifetimes. If a significant share of, say, men who started their working life as textile workers changed to, say, farming later in life, this would create a non-wealth-related, additional bias in the probate data. But this additional source of bias would automatically be reflected in a higher value of the calibration factors for workers in the textiles industry, and would therefore equally automatically be removed when applying these factors to the probate data.

Secondly, the approach obviously relies on the source of the calibration, that is parish registers, to be an unbiased and accurate source of occupational information itself. Any potential bias in the parish register is reflected in the calibration factors and, therefore, replicated when these are applied to the probate data. When collecting parish register data, members of the Occupational Structure project have been very careful to minimise potential sources of bias. Only those baptism registers where included in the Cambridge Group's parish register dataset in which occupations were recorded for at least 95 per cent of the baptised children for which a father's occupation could be expected to be given – so excluding illegitimate children. This was done to avoid bias produced by, for example, parish clerks frequently omitting an occupational descriptor for fathers employed in the dominant, 'default' occupation.[38] By comparing them to other sources, Kitson et al have shown that the parish register data are, indeed, reliable.[39] At first sight, servants present a problem. Since servants were highly unlikely to be married and, therefore, to father children they are underrepresented in the baptism registers compared to sources which include men before the age of marriage. However, when combined with labourers, 'they comprise almost exactly the same share of the overall male workforce' since the vast majority of both servants and labourers worked in agriculture, with farm

---

[38] A real risk, as is clear from an analysis of precisely such a type of bias in the Gloucestershire 1608 muster list – as discussed on page 80.

[39] Kitson *et al*, 'Creation', pp. 10-5.

servants becoming farm labourers after marriage.[40] Domestic servants *are* somewhat underrepresented in the parish registers, however, as Kitson et al's comparison to the 1841 demonstrated, and a correction factor was therefore applied to the domestic servant numbers in the parish register data and the calibrated probate data presented in this paper.

USING THE APPROACH WHERE/WHEN NO PARISH DATA ARE AVAILABLE

The probate calibration approach employs the strengths of one data source to eliminate the weaknesses of the other. Parish records have little or no bias, and can therefore be used to remove the probate record's inherent occupational bias. Conversely, probate data provide full geographic coverage and can therefore be used to reconstruct the missing parishes from the parish record. The methodology utilises to its advantage a contrast between the two data source which Glennie identified more than twenty years ago: 'The relatively sparse spatial and temporal coverage provided by parish registers is a major weakness, but is partially compensated for by their broad social coverage. More or less the opposite is true of wills and probate inventories.'[41] Indeed, the probate and parish record complement each other beautifully. In combination, they allow one to determine a reliable and unbiased occupational structure for every cohesive geographical area for which an index to probate documents and a reasonable number of parish records with occupational information are available. But, the methodology can also be used to determine male occupational structures in time periods and geographies for which no parish register data are available at all. How and why that works is explained in this section.

As demonstrated in the Cheshire case example, the methodology works by multiplying probate data with calibration factors derived from a comparison with parish register data which is 'near' in time and place. Near in time, as the comparison is limited to a relatively short time interval of three decades. Near in space, as the comparison is limited to parishes and chapelries in one and the same county. But what if there are no parish data which are near in time and space? In eighty per cent of English and Welsh counties, occupations were not reliably recorded in a single parish between 1730 and 1813. And before 1695, virtually no parishes with reliable occupational records can be found at all, outside London. Can we assume that the occupational bias of the probate records is sufficiently stable to justify 'borrowing' calibration factors from further away in time and place? Historians have generally assumed that the answer to this question is no. Glennie, for example, thought it: 'unlikely that the "sample" of men represented by probate documents will possess precisely the same … occupational bias in different areas, or for one area at different times.'[42] And Clark has recently suggested that the whole impression of a consumer revolution in the eighteenth century derived from probate inventories may not be a reflection of historical reality at all but simply the result of increasing bias in the probate record over that century.[43] It will be shown, however, that such fears are exaggerated.

Figure 9 tests the degree to which calibration factors can be 'borrowed' from nearby areas in the same time periods. Figure 9 focuses on southern England since there, the lack of parish register data for calculating local calibration factors was particularly problematic, with for example the South-West

---

[40] Ibid, p. 11.

[41] Glennie, *Distinguishing men's trades*, p. 32.

[42] Ibid, p. 40.

[43] Clark, 'The consumer revolution: turning point in human history, or statistical artifact?' (Davis, 2010), http://www.econ.ucdavis.edu/faculty/gclark/papers/Consumer%20Revolution.pdf.

and South-East hardly having any parish register coverage at all. But similar charts were made for other parts of the country, with similar results.
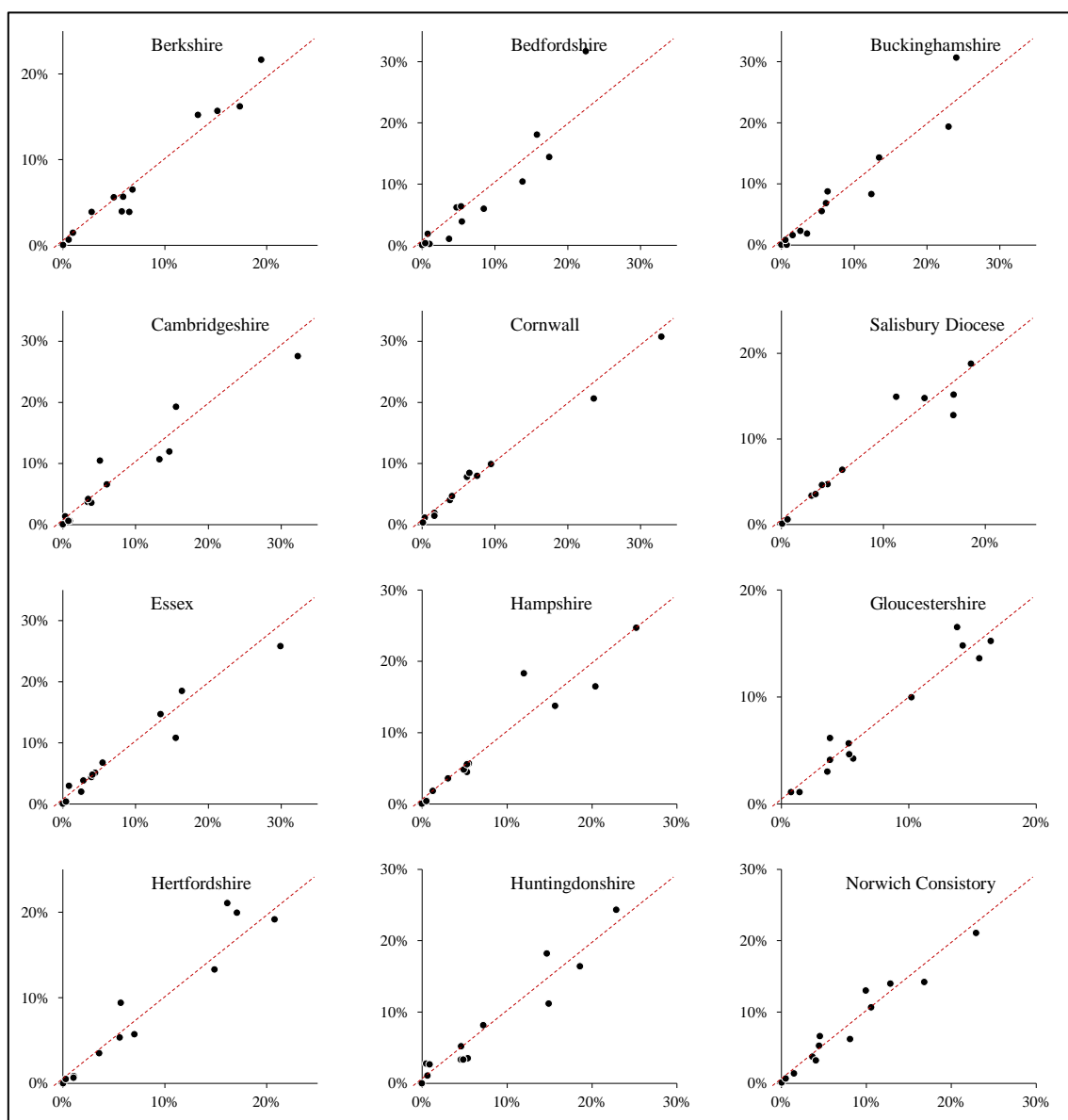

Figure 9. A comparison between parish register and probate-derived male occupational structure using calibration factors from nearby and comparable counties
(southern England, c.1817)

Notes: Each data point represents an occupational sub-sector (agriculture, mining, other primary sector, clothing, footwear, textiles, metal and tools, building, other secondary sector, dealers and sellers, services and professions, and transport). Labourers and seamen were excluded, for reasons explained elsewhere.[44] Actual male labour shares, derived from parish register data, are plotted on the horizontal axis, with the vertical axis representing those calculated from the calibrated probate data. The calibration factors used in each county/area are those derived from all *other* counties/areas. The red lines are 'x=y' lines; had the calibration been perfect, all data points would have landed on these lines. All charts represent a single county, except for 'Salisbury Diocese' (Wiltshire + those Dorset parishes in that probate jurisdiction) and 'Norwich Consistory' (Norfolk + those Suffolk parishes in that probate jurisdiction). Only parishes in which probate and parish register data are available were included, as only for those can a one-on-one comparison be made.

---

[44] Keibek, *Male occupational structure*, pp. 132-8.

In each of the scatter plots in Figure 9, the x-axis represents the male occupational share per occupational sub-sector according to parish registers in one county, whilst the y-axis represents the share according to the calibrated probate data in the same county. But the calibration factors used are not those from the county itself, but from the combination of all *other* counties in the figure. In other words, the calibration factors that were used to calibrate the Berkshire probate data were derived from a comparison of probate and parish register data in Bedfordshire, Buckinghamshire, Cambridgeshire, etcetera. Each dot represents an occupational sub-sector. A red 'x=y' line has been included in each chart; had the probate calibration been perfect, all dots in each chart would have been on this line. As the figure shows, the calibration is not perfect, but the match with the parish register data is quite good. The coefficient of determination (R^2) varied between .89 for Bedfordshire to .97 for Cornwall. In short, for areas in which, in the absence of parish register data, no local calibration factors can be determined, using those from counties in the same general geographic area generates fairly accurate results.

The value of this what one might call 'spatial extendibility' of probate calibration factors is useful for areas for which, in the absence of parish register data, no local calibration factors can be determined. For example, no occupational data from parish registers exist for Cornwall before Rose's Act, so no local probate calibration factors can be calculated for, say, the early eighteenth century. But for many southern English counties, parish register data *are* available. By using probate calibration factors derived from those, a reliable, unbiased male occupational structure can be calculated from Cornwall's probate data.

This 'spatial extendibility' is also useful for another reason. To optimally remove the probate data's occupational bias, it is beneficial to work with data at low levels of occupational abstraction, since the probability of being probated varied strongly even within sub-sectors. For example, the textiles sub-sector encompasses a variety of occupations – clothiers, fullers, dyers, weavers, male spinners, etcetera – with very different average wealth levels and capital requirements. Using one calibration factor for the entire textiles sub-sector potentially leads to sub-optimal results; the composition of the sub-sector varied over time and place, and this variation cannot accurately be captured using a single calibration factor. But there is a limit to the degree to which sub-sectors can be split into smaller occupational groups when calculating calibration factors: there have to be sufficient numbers of probate and parish register data for a reliable result. A balance needs to be struck between, on the one hand, occupational specificity and, on the other hand, statistical significance. For common occupations – yeomen, coal miners, tailors, blacksmiths, weavers, schoolmasters, etcetera – the numbers of parish register and probate observations were usually large enough to accurately determine calibration factors at the level of individual counties. But for other, rarer occupations – foresters, button makers, fullers, nailors, clock makers, naval officers, etcetera – this was only possible in some counties or for larger geographical areas. By combining counties, occupational specificity can be improved and low-level calibration factors can be calculated. This enables one to use local calibration factors for (locally) frequent occupations whilst using regional or national calibration factors for occupations that were (locally) rare. The calculations in this paper are based on a set of more than one hundred calibration factors, listed in Table 3. This high degree of occupational specificity in unbiasing the probate data was possible because not all of these calibration factors needed to be calculated from local data for every county; where local data were statistically insignificant, regional or even national figures could be used.

Table 3. Probate calibration factors and their average, national values c.1817

Sector	Sub-sector	Occupational group	Cal. factor
Primary	Agriculture	Farmer/yeoman	1.00
		Husbandman	5.63
		(Market) gardener	3.90
		Animal husbandry, sheep	6.81
		Animal husbandry, other	1.04
		Agriculture, management	7.18
		Agriculture, other	14.51
	Mining	Coal miner	24.53
		Miner (type unspecified)	10.76
		Mining, other	3.82
		Quarrying	9.11
	Rest of primary	Fisherman	9.90
		Forester	4.49
		Game/parkkeeper	8.86
		Primary sector, other	4.56
Secondary	Clothing	Tailor, clothing maker	5.28
		Button maker	1.84
		Clothing, other	5.29
		Hatter/glover	7.95
	Footwear	Shoe/bootmaker/clogger	6.72
	Textiles	Weaver	13.33
		Spinner	23.65
		Clothier/textile fabric/products maker	2.85
		Textile processor	9.40
		Dyer	6.60
		Wool comber, carder, fuller, shearer	8.40
		Textiles, other	3.09
	Metal manufacture and products	Blacksmith	5.15
		Iron/steel manufacture	12.71
		Nail/screw/pin maker	10.48
		Iron/steel products making, other	5.58
		Non-ferrous metal, raw and products	3.91
	Machines and tools making	Machine making	3.40
		Tool making	5.77
		Edge tool manufacturer	6.42
		Engineer/milwright	7.38
	Building and construction	Carpentry	6.04
		Bricklaying	6.04
		Masonry	7.50
		Painter/decorator/plasterer	7.28
		Plumber/glazier	2.14
		Roofer/slater/tiler	7.84
		Building, other, specialist trades	0.97
	Rest of secondary sector	Brewer	1.72
		Baker/confectioner	3.13
		Butcher	3.39
		Miller	3.80
		Food and drinks, other	1.15
		Skinner/tanner	2.88
		Saddle/harness maker	2.72
		Leather industries, other	2.44
		Sawyer, timber worker	21.83

Continued on next page

Continued from previous page

Sector	Sub-sector	Occupational group	Cal. factor
	Rest of second'y sector (cont'd)	Furniture maker	4.63
		Cooper, cask maker	3.40
		Wooden products, other	4.80
		Wheelwright	3.96
		Ship building	6.36
		Brick and tile making	4.93
		Chemical, salt, soap production	2.93
		Clock/watch/scientific instruments makir	3.02
		Coach/cart building	5.33
		Glass production	11.72
		Paper making	9.99
		Pottery, earthenware manufacture	7.83
		Precious metal working, jewelry making	2.62
		Printer, book producer	12.39
		Rope making	6.24
		Secondary sector, other or unspecified	4.17
Tertiary	Dealers and sellers	Shopkeeper/chandler/grocer	1.53
		Ironmonger	0.98
		Jeweller	3.01
		Peddler, hawker, itinerant trader	2.45
		Retail, other	1.91
		Merchant	1.00
		Draper/mercer	1.50
		Dealer/factor/wholesaler	1.45
	Services and professions	Domestic service	14.09
		Soldier/non-commisioned officer/militiar	15.89
		Army, other	2.97
		Naval officer	11.92
		Marine	20.64
		Navy, other	8.79
		Hospitality services	1.25
		Entertainment	5.54
		Hairdressing	3.85
		Warehousing	3.57
		Services, government	1.49
		Services, customs and excise, inland rev	4.41
		Services, other	1.87
		Professions, education	2.60
		Professions, legal	6.05
		Professions, medical	1.31
		Professions, religious	1.29
		Professions, other	2.84
		Profession, support	2.97
	Transport and communications	Road transport, drivers	7.68
		Road transport, other	13.32
		Inland navigation	17.30
		Merchant navy, officer	1.10
		Merchant navy, other	8.50
		Maritime services	0.38
		Transport, other	6.71
Labourer	Labourer	Labourer	38.44
Not included in occupational structure		Distinguished titles and owners of capita	0.29

But how to calibrate probate data in time periods when no parish register data are available at local, regional, or even national level? Figure 10 examines the 'temporal extendibility' of probate calibration factors for the c.1710-c.1817 time period, in a similar fashion to that in which Figure 9 tests 'spatial extendibility'. Again, the x-axis in each chart represents the male occupational share per sub-sector according to parish registers, whilst the y-axis represents the share according to the calibrated probate data, with each data point being an occupational sub-sector. The probate calibration factors used for each chart were those derived from a comparison of probate and parish register data for c.1817.[45] As the charts show, projecting calibration factors backwards in time yields good results, with the vast majority of sub-sector 'dots' on or close to the 'x=y' lines. Apparently, calibration factors are remarkably stable over time, even over the full century separating the c.1710 probate data in the lower three charts from the c.1817 calibration period.
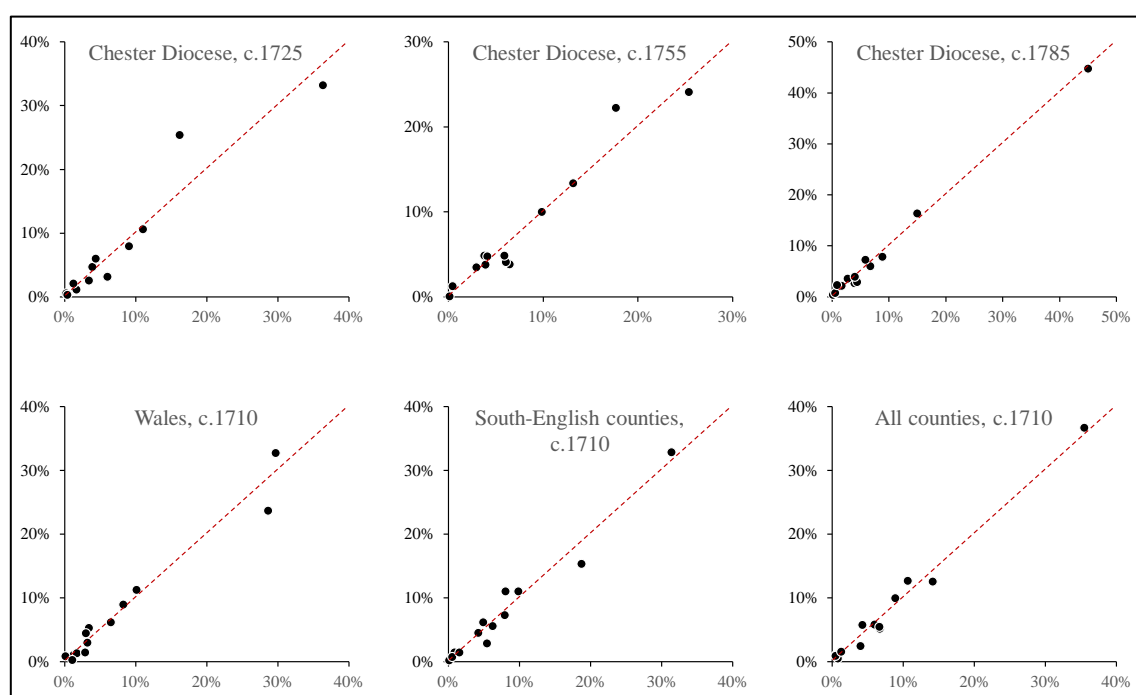

Figure 10. A comparison between parish register and probate-derived male occupational structure using calibration factors from c.1817 (England and Wales, c.1710-c.1785)

Notes: Each data point represents an occupational sub-sector (agriculture, mining, other primary sector, clothing, footwear, textiles, metal and tools, building, other secondary sector, dealers and sellers, services and professions, and transport). Labourers and seamen were excluded, for reasons explained elsewhere.[46] Actual male labour shares, derived from parish register data, are plotted on the horizontal axis, with the vertical axis representing those calculated from the calibrated probate data. The calibration factors used in each county/area are those derived for c.1817 for the same geographic area, parish register data from the 1813-20 period and probate data from the 1802-32 period. The red lines are 'x=y' lines; had the calibration been perfect, all data points would have landed on these lines. Only parishes in which probate and parish register data are available were included, as only for those can a one-on-one comparison be made.

Since calibration factors are, apparently, sufficiently 'temporally extendible' to generate accurate results, they can be used to interpolate between two points in time for which both probate and parish register data are available. As discussed above, for the overwhelming majority of English and Welsh

---

[45] That is, parish register data for 1813-20 and probate data for 1802-32.
[46] Keibek, Male occupational structure, pp. 132-8.

counties, no parish registers can be found in which occupations were reliably recorded for the mid and late eighteenth century. This is problematic, as this is a critically important time period from the perspective of economic history, since it coincides with the immediate run up to and initial phase of the Industrial Revolution. As probate data *are* available for this period in the overwhelming majority of English and Welsh counties, this problem can now be solved. Figure 11 provides an example for Wiltshire.
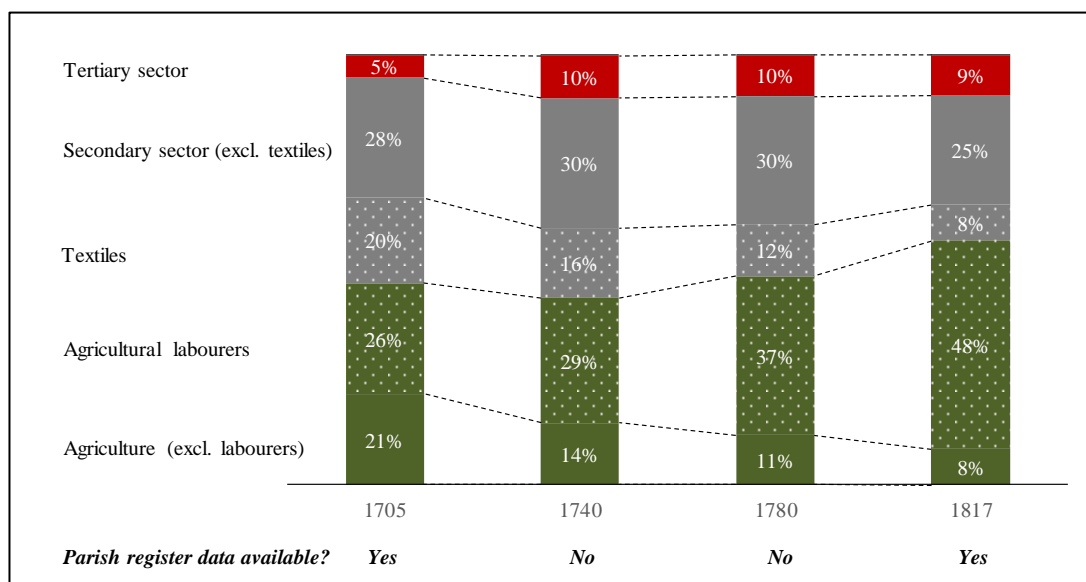

Figure 11. Development of Wiltshire's male occupational structure, according to calibrated probate data, c.1705-1817

Notes: parish register data for calibrating contemporary calibration factors are available for c.1705 and, of course, post 1812. For c.1740 and c.1780, contemporary probate data and calibration factors 'borrowed' from c.1705 and c.1817 were used.
Sources: Wiltshire probate index, obtained from the Wiltshire Record Office; parish register database created by the Cambridge Group.

The 'temporal extendibility' of the calibration factors cannot only be used to *inter*polate between two points in time for which parish data are available, as in Figure 11, but also be used to *extra*polate outside the period with parish data altogether. As discussed, before 1700, parish registers with reliable occupational information are very rare. By employing probate data and the – presumed constant – calibration factors derived from the parish register period, it is nevertheless possible to estimate male occupational developments in the seventeenth century. In Figure 12, this was done for Cheshire.
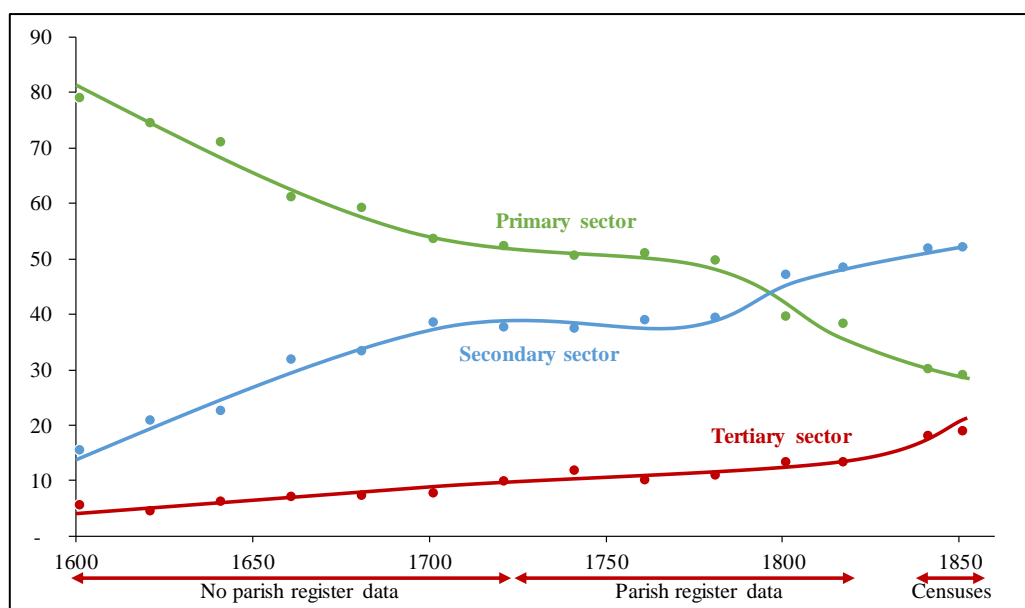
Figure 12. Male labour share percentage by occupational sector (Cheshire, 1601-1851)

Notes: before c.1725, no parish register data are available for determining contemporary calibration factors. The c.1725 factors were therefore used, and applied to pre-1725 probate data. Labourers were allocated to sectors in line with the approach set out in a separate working paper.[47]

Sources: Probate index from the Chester Record Office; parish register database created by the Cambridge Group; 1841 and 1851 census.

But Figure 12 assumes that early-eighteenth-century calibration factors are temporally extendible into the seventeenth century. Unfortunately, it is not possible to test this assumption by creating the seventeenth-century version of Figure 10, since the required parish register data are lacking. Fortunately, two other sources of occupational data with (almost) comprehensive social coverage exist for the seventeenth century, albeit only for relatively small areas, and only for a single year. The first of these is formed by the returns for the 1660 poll tax for the Cheshire hundred of Northwich, which was transcribed for the Record Society of Cheshire and Lancashire by G.O. Lawton, and which contains occupations for most men.[48] All except those on poor relief were expected to pay this tax. An occupation was recorded for 73 per cent of the three thousand men listed in the returns. Table 4 provides a comparison of the male occupational structure according to the poll tax returns with those derived from probate data; the latter were calculated using calibration factors for Cheshire in c.1725, so over sixty years later. As the table shows, the two sources provide a very good match.

---

[47] Keibek, 'Allocating labourers'.
[48] Transcribed in Lawton, *Northwich hundred: poll tax 1660 and hearth tax 1664* (Bath: Pitman Press, 1979).

Table 4. Comparison of occupational structure according to poll tax
and calibrated probate data (Northwich hundred, 1660)

	From poll tax (%)	From probate (%)
Primary sector	**64.9**	**66.9**
Agriculture	64.8	66.9
Mining	0.0	0.0
Other	0.0	0.0
Secondary sector	**29.6**	**29.6**
Clothing	5.2	4.7
Footwear	2.7	2.2
Textiles	4.6	6.5
Metal trades & tools	3.8	2.7
Building	4.1	2.8
Other	9.2	10.7
Tertiary sector	**5.5**	**3.5**
Dealers and sellers	1.4	0.6
Services & professions	3.7	2.7
Transport	0.3	0.2

Notes: The probate data used were for the period 1640-1670 (with relatively few data from the 1650s, when the probate process was centralised in London). C.1725 calibration factors were applied to the probate data. Labourers were allocated to sectors in line with the approach outlined elsewhere.[49]

A second seventeenth century source of male occupational information with comprehensive social coverage is comprised by the 1608 muster list for Gloucestershire.[50] This source lay at the basis for 'An occupational census of the seventeenth century' in a deservedly famous paper of the same name by A.J. and R.H. Tawney.[51] Deriving a reliable occupational structure from the muster list requires addressing two issues. Firstly, the share of men provided with an occupation in the muster list varies from place to place, and this variation is not random, as Figure 13 demonstrates: the more agricultural the area, the lower the share of men for whom an occupation was listed. If the muster list data would be taken at face value, the agricultural sector would be underrepresented. Why the occupations of a relatively high proportion of men were not specified in farming-dominated areas is a matter for speculation. It may simply have been caused by farming being *so* dominant in these area that it was considered the 'default' activity, with occupational specification (mostly) reserved for the few men who were *not* involved in agriculture. It is also possible that in the hamlets and small villages in which these men lived, an occupational denominator was simply not required to distinguish one man from another, as the likelihood of the same first name/surname combination occurring more than once was very low in these small settlements.

---

[49] Keibek, 'Allocating labourers'.

[50] Smyth, *The names and surnames of all the able and sufficient men in body fitt for his Majestie's service in the warrs, within the City of Gloucester and the inshire of the same*, manuscript (1608).

[51] Tawney and Tawney, 'An occupational census of the seventeenth century', *The Economic History Review*, a5:1 (1934).
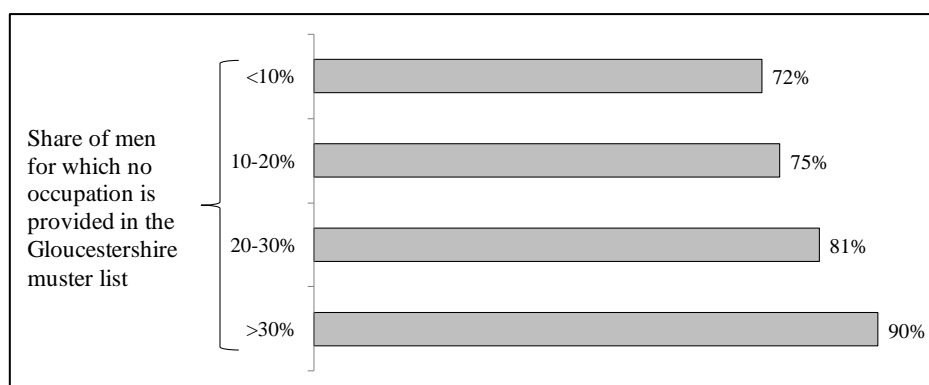
Figure 13. The agricultural share of the male labour force in the uncorrected probate record as a function of the percentage of men of unspecified occupation in the muster list (Gloucestershire, 1608)

A second problem with the muster list data is the sizeable number of servants. The Tawneys allocated these to occupational sectors based on the occupation of their masters. They considered servants of masters specified by status only – gentlemen, knights, esquires, etcetera – to have been domestic servants.[52] The Tawneys' approach leads to 39 per cent of servants being allocated to agriculture, 24 per cent to the secondary sector, and 37 per cent to the tertiary sector, almost all as domestic servants in the households of gentlemen and other men of status. This distribution seems rather light on farm servants. The Tawneys admit that the number of domestic servants is overstated as 'some of those classed as household servants … were undoubtedly farm servants'; indeed, they note that in a number of stately households, the servants are specified as domestic or agricultural, with the latter making up around forty per cent of the total.[53] But their assumption that servants of secondary sector workers were employed in their masters' trade is also debatable. It is possible that many of them, perhaps even the overwhelming majority, were farm servants. As discussed in a separate working paper, a significant proportion of households whose male 'household head' worked in a secondary sector occupation was engaged in farming as a household by-employment, with most of the farming undertaken by household members other than the 'household head' such as his wife, living-in children, and (farm) servants.[54]

Table 5 presents Gloucestershire's estimated occupational structure in 1608 as derived from the muster list, taking the above two issues into account. The agricultural sector has been upwardly corrected for under-specification in farming-dominated parishes, as depicted in Table 5. To express the uncertainties in the allocation of servants to sectors, two scenarios were calculated. One copies the Tawneys' approach, allocating servants in line with their masters' occupation. In the alternative approach, forty per cent of the servants of gentlemen and other men of status were allocated to farming rather than considered domestics, and all servants of secondary-sector workers were considered farm servants too. The result is a (manageable) range rather than a single set of values. Table 5 also includes the occupational structure derived from contemporary probate data, using calibration factors from the early eighteenth century. Despite a gap of more than a century between those calibration factors and the probate data to which they were applied, the muster list and

---

[52] Ibid, pp. 33-4.
[53] Ibid, p. 33; see also footnote 2 on that page.
[54] Keibek, 'By-employments'.

calibrated probate estimates match well – with the exception of the textiles sub-sector, which requires special treatment, as discussed elsewhere.[55]

Table 5. A comparison of male occupational structures as derived from muster list and calibrated probate data (Gloucestershire, c.1608)

	From muster list (%)	From probate (%)
Primary sector	**54-58**	**59**
Agriculture	53-58	58
Mining	1	0
Other	0	0
Secondary sector	**35-37**	**37**
Clothing	5	4
Footwear	2	2
Textiles	14-15	20
Metal trades & tools	3	2
Building	4	5
Other	6-7	6
Tertiary sector	**6-8**	**5**
Dealers and sellers	1-2	1
Services & professions	3-5	2
Transport	2	1

Notes: Excluding Bristol, which was not covered in either data source and excluding Bilbury Peculiar, which was not included in the probate index. See main text for background. Note that domestic servants are included in 'services & professions'.
Sources: Probate database; parish register database; Gloucestershire muster list.

CLOSING REMARKS

In conclusion, probate documents, in their 'raw' form, are a gravely unreliable source of male occupational information, but they can be converted into a reliable data source through calibration with parish register data from the same parishes and time periods. Furthermore, the calibration factors obtained by this procedure can be used to remove the occupational bias from probate data from time periods and geographic areas in which no parish register data are available. The approach effectively exploits the complimentary properties of the parish register and probate data, using the latter to calibrate the former, and the former to interpolate and extrapolate the latter. In this way, historical male occupational structures can be reconstructed at national, county, and sub-county levels for the 1600 to 1850 period – as presented in my recent PhD dissertation.[56]

It should be stressed at this point, however, that the above only describes the general approach to using probate data for the purpose of estimating the historical composition of the male labour force. Occupational bias is the central problem of the probate data, but it is not the only one. There are many smaller, more restricted issues which require addressing too: how to achieve accuracy for particularly underrepresented occupations such as labourers; how to deal with the absence of miners in the probate record in some counties; how to account for changes in calibration factors over time for occupations experiencing major technological change – such as those in textiles – or increases in scale – such as

---

[55] Keibek, *Male occupational structure*, pp. 131-32.
[56] Ibid, ch. 7-9 and appendix B.

farming in southern England; etcetera. These specific, technical, but not insubstantial issues, and their solutions, are addressed elsewhere.[57]

<p style="text-align:center">*   *   *</p>

BIBLIOGRAPHY

Arkell, T., 'The probate process', in Tom Arkell, Nesta Evans and Nigel Goose (eds), *When death do us part : Understanding and interpreting the probate records of early modern England* (Oxford: Leopard's Head Press, 2000), pp. 3-13.

Basten, S., 'From Rose's *Act* to Rose's *Bill*: a reappraisal of the 1812 Parish Register Act', *Local Population Studies,* 76 (2006), pp. 43-62.

Broadberry, S., B.M.S. Campbell, A. Klein, M. Overton, and B. Van Leeuwen, *British economic growth, 1270-1870* (Cambridge: Cambridge University Press, 2015).

Chalklin, C.W., *Seventeenth-century Kent; a social and economic history* ([London]: Longmans, 1965).

Churchley, R.A., *Differing responses to an industrialising economy: occupations in rural communities in the heart of England from the restoration to the railway age (c. 1660 - c. 1840)* (DPhil thesis, University of Birmingham, 2010).

Clark, G., *A farewell to alms: a brief economic history of the world* (Princeton, N.J. ;Woodstock: Princeton University Press, 2007).

———, 'The consumer revolution: turning point in human history, or statistical artifact?' (Davis, 2010), http://www.econ.ucdavis.edu/faculty/gclark/papers/Consumer%20Revolution.pdf.

Clark, G., J. Cummins, and B. Smith, 'Malthus, wages, and preindustrial growth', *The Journal of Economic History,* 72:02 (2012), pp. 364-92.

Clarkson, L.A., *The pre-industrial economy in England, 1500-1750* (London: Batsford, 1971).

Cockburn, J.S., 'The use of assize records as historical evidence', *Journal of the Society of Archivists,* 5 (1975), pp. 215-31.

Crafts, N.F.R., *British economic growth during the Industrial Revolution* (Oxford: Clarendon, 1985).

Crafts, N.F.R., and C.K. Harley, 'Output growth and the British Industrial Revolution: a restatement of the Crafts-Harley view', *The Economic History Review,* 45:4 (1992), pp. 703-30.

Cressy, D., 'Occupations, migration and literacy in East London, 1580–1640', *Local Population Studies,* 5 (1970), pp. 53-60.

Evans, N., 'The occupations and status of male testators in Cambridgeshire, 1551-1800', in Tom Arkell, Nesta Evans and Nigel Goose (eds), *When death do us part : Understanding and interpreting the probate records of early modern England* (Oxford: Leopard's Head Press, 2000), pp. 176-88.

Everitt, A., 'Farm labourers', in H. P. R. Finberg (ed.), *The agrarian history of England and Wales. Vol. IV, 1500-1640* (Cambridge: Cambridge University Press, 1967), pp. 396-465.

Frost, P., 'Yeomen and metalsmiths: livestock in the dual economy in south Staffordshire 1560–1720', *The Agricultural History Review,* 29:1 (1981), pp. 29-41.

Glennie, P., *A commercialising agrarian economy: late-medieval and early-modern Hertfordshire* (Master thesis, University of Cambridge, 1983).

———, *'Distinguishing men's trades': occupational sources and debates for pre-census England* (Bristol: Historical Geography Research Group, 1990).

Hey, D.G., 'A dual economy in south Yorkshire', *The Agricultural History Review,* 17:2 (1969), pp. 108-19.

Holderness, B.A., 'Rural tradesmen 1660-1850: a regional study in Lindsey', *Lincolnshire History and Archaeology,* 7 (1972), pp. 77-83.

Hudson, P., 'Landholding and the organization of textile manufacture in Yorkshire rural townships c. 1660-1810', in Maxine Berg (ed.), *Markets and manufacture in early industrial Europe* (London: Routledge, 1991), pp. 261-91.

---

[57] Ibid, ch. 6.

Keibek, S.A.J., 'Allocating labourers to occupational (sub-)sectors using regression techniques' (working paper, Cambridge, 2017), http://www.econsoc.hist.cam.ac.uk/docs/CWPESH_number_27_March_2017.pdf.

———, 'By-employments in early modern England and their significance for estimating historical male occupational structures' (working paper, Cambridge, 2017), http://www.econsoc.hist.cam.ac.uk/docs/CWPESH_number_29_March_2017.pdf.

———, *The male occupational structure of England and Wales, 1600-1850* (PhD dissertation, University of Cambridge, 2017).

Kitson, P.M., L. Shaw-Taylor, E.A. Wrigley, R.S. Davies, G.H. Newton, and A.E.M. Satchell, 'The creation of a "census" of adult male employment for England and Wales for 1817' (Cambridge, 2012), http://www.econsoc.hist.cam.ac.uk/docs/CWPESH%20number%204%20March%202012.pdf

Lawton, G.O., *Northwich hundred: poll tax 1660 and hearth tax 1664* (Bath: Pitman Press, 1979).

Lindert, P.H., 'An algorithm for probate sampling', *The Journal of Interdisciplinary History,* 11:4 (1981), pp. 649-68.

Oestreicher, R., 'The counted and the uncounted: the occupational structure of early American cities', *Journal of Social History,* 28:2 (1994), pp. 351-61.

Overton, M., J. Whittle, D. Dean, and A. Hahn, *Production and consumption in English households, 1600-1750* (London: Routledge, 2004).

Pahl, R.E., *Divisions of labour* (Oxford: Blackwell, 1984).

Patten, J., 'Urban occupations in pre-industrial England', *Transactions of the Institute of British Geographers,* 2:3 (1977), pp. 296-313.

———, 'Changing occupational structures in the East-Anglian countryside', in H. S. A. Fox and R. A. Butlin (eds), *Change in the countryside: essays on rural England, 1500-1900* (London: Institute of British Geographers, 1979), pp. 103-21.

Ripley, P., 'Village and town: occupations and wealth in the hinterland of Gloucester, 1660-1700', *The Agricultural History Review,* 32 (1984), pp. 170-7.

Rosenheim, J.M., 'Review of Swain's "Industry before the Industrial Revolution"', *Albion,* 20:1 (1988), pp. 99-100.

Rowlands, M.B., *Masters and men in the West Midland metalware trades before the Industrial Revolution* (Manchester: Manchester University Press, 1975).

Rudnicki, T., *The male occupational structure of northwest England, circa 1600 to 1851* (Master thesis, University of Cambridge, 2015).

Shaw-Taylor, L., and E.A. Wrigley, 'Occupational structure and population change', in Roderick Floud, Jane Humphries and Paul Johnson (eds), *The Cambridge Economic History of Modern Britain. Volume 1. Industrialisation, 1700-1870* (Cambridge: Cambridge University Press, 2014), pp. 53-88.

Smyth, J., *The names and surnames of all the able and sufficient men in body fitt for his Majestie's service in the warrs, within the City of Gloucester and the inshire of the same*, manuscript (1608).

Sneath, K.G., *Consumption, wealth, indebtedness and social structure in early modern England* (PhD thesis, University of Cambridge, 2009).

Stobart, J., 'Geography and industrialization: the space economy of northwest England, 1701-1760', *Transactions of the Institute of British Geographers,* 21:4 (1996), pp. 681-96.

———, 'The economic and social worlds of rural craftsmen-retailers in eighteenth-century Cheshire', *The Agricultural History Review,* 52:2 (2004), pp. 141-60.

———, *The first industrial region: North-west England, c.1700-60* (Manchester: Manchester University Press, 2004).

Sugden, K., *An occupational analysis of the worsted industry, circa 1700-1851. A study of de-industrialization in Norfolk and the rise of the West Riding of Yorkshire* (PhD thesis, University of Cambridge, 2015).

Swain, J.T., *Industry before the Industrial Revolution: north-east Lancashire c.1500-1640* (Manchester: Manchester University Press for the Chetham Society, 1986).

Tawney, A.J., and R.H. Tawney, 'An occupational census of the seventeenth century', *The Economic History Review,* a5:1 (1934), pp. 25-58.

Thirsk, J., 'The rural economy of England: collected essays', (London: Hambledon, 1984), pp. xiv,420p.

———, 'Seventeenth-century agriculture and social change', *The rural economy of England : Collected essays* (London: Hambledon, 1984), pp. 183-216.

Wallis, P., J. Colson, and D. Chilosi, 'Puncturing the Malthus delusion: structural change in the British economy before the industrial revolution, 1500-1800' (working paper, London, 2016), http://www.lse.ac.uk/economicHistory/home.aspx.

West, J., *Town records* (Chichester: Phillimore, 1983).

Woodward, D., 'Wage rates and living standards in pre-industrial England', *Past & Present*:91 (1981), pp. 28-46.

Zell, M., 'The social parameters of the probate records in the sixteenth century', *Historical Research,* 57:135 (1984), pp. 107-13.

———, *Industry in the countryside: Wealden society in the sixteenth century* (Cambridge: Cambridge University Press, 1994).