Using probate data for estimating historical male occupational structures

S.A.J. Keibek (sk571@cam.ac.uk)

Working paper, Cambridge, January 2016

Note

This paper is preliminary and should not be cited without written permission. An updated and extended version will be included in my PhD thesis 'The male occupational structure of England and Wales, 1650-1850'.

Abstract

The pre-census estimates of the male occupational structure of England and Wales, recently published in the CEHMB, 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 or well-paying occupations, can be overcome by using parish register data for calibration.

The 'Occupational Structure of Britain 1379-1911' research programme has been running since 2003, led by Leigh Shaw-Taylor, 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. 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 after1812, 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 '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*. 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

[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 Working Papers in Economic and Social History*, 4 (2012), http://www.econsoc.hist.cam.ac.uk/docs/CWPESH%20number%204%20March%202012.pdf.

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.'[3] 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. This paper, however, is solely concerned with male occupational structures. Secondly, many of the fathers captured in baptism registers were given the 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.[4] 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.[5] 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.[6] This issue is addressed in a separate working paper also.[7] 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.

---

[3] Crafts and Harley, 'Output Growth and the British Industrial Revolution: A Restatement of the Crafts-Harley View', *The Economic History Review*, 45:4 (1992), p.721.

[4] Keibek, 'Allocating labourers to occupational sectors using regression techniques', *Campop paper*, http://www.campop.geog.cam.ac.uk/research/projects/occupations/abstracts/paper28.pdf.

[5] See, for example, Thirsk, 'Seventeenth-Century Agriculture and Social Change', in her *The Rural Economy of England : Collected Essays* (London, 1984), 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.

[6] 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.

[7] Keibek, 'By-employments and occupational structure in pre-industrial England', *Campop paper*, http://www.campop.geog.cam.ac.uk/research/projects/occupations/abstracts/paper30.pdf.

The problem of incomplete coverage

Baptism registers only provide occupational information for a small minority of parishes before Rose's Act. Coverage is particularly low in the mid and late eighteenth century, as figure 1 shows, and close to non-existent outside London before 1700. This significantly impedes realisation of the promise of regional and local insights into economic developments, discussed above.

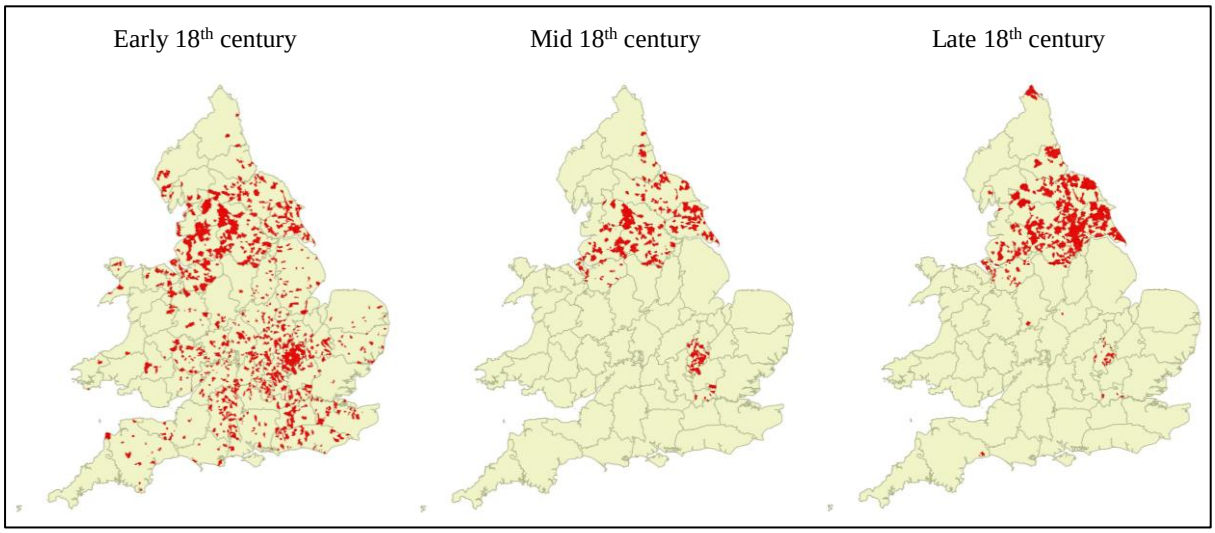

Figure 1. Registration of male occupations in English and Welsh baptism registers over the course of the eighteenth century.

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. Small samples are not necessarily a problem, of course, as long as these are representative of the population from which they were taken. However, as figure 1 shows, the samples were 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, and the North 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 even at this lower geographical level the sampled parishes cannot simply be presumed representative. 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 2 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.

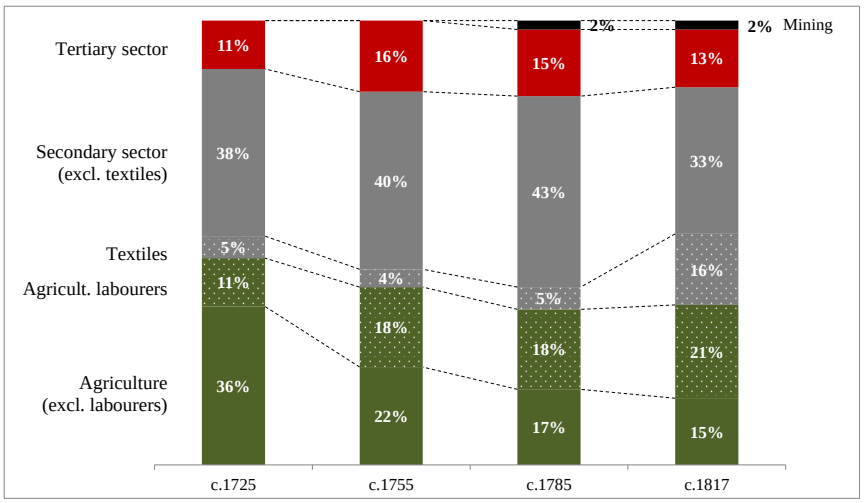

Figure 2. Cheshire's male occupational structure according to parish registers, c.1725-c.1817

Source: parish register data collected by the Cambridge Group.

Well aware of these issues, two different remedial approaches have been used by members of the Group. The first of these, developed to generate earlier, preliminary occupational estimates published online in 2010, was merely a stop-gap measure, as it was based on an assumption which was known to be questionable. It assumed that the bias resulting from a non-random early eighteenth century sample was equal to the bias of that same set of parishes within the c.1817 dataset.[8] That this is a dangerous assumption can be illustrated by testing it for the pre-industrial occupational structure of the future heartland of the industrial revolution, what Jon Stobart has called 'the world's first industrial region': Lancashire and Cheshire.[9] These counties are amongst those with the highest parish register coverage for the early eighteenth century. Nevertheless, when using the c.1817 dataset to reweight the early eighteenth century parish sample one ignores the fact that the most rapidly industrialising and therefore fastest growing regions are entirely missing from that sample in both counties. Neither Manchester in Lancashire nor the neighbouring hundred of Macclesfield in Cheshire are covered in the early sample, meaning that the c.1817 weights are unlikely to be representative of the situation in the eighteenth century.

[8] Shaw-Taylor *et al*, 'The Occupational Structure of England C.1710 to C.1871, Work in Progress' (Cambridge, 2010), http://www.geog.cam.ac.uk/research/projects/occupations/britain19c/papers/paper3.pdf, p.7.
[9] Stobart, *The First Industrial Region : North-West England, C.1700-60* (Manchester: Manchester University Press, 2004).

For the new CEHMB, a second and much better approach was therefore developed by Shaw-Taylor and Wrigley. It simultaneously divides the pre-Rose's Act parishes into 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 methodology is undoubtedly a great improvement over the 'raw' data and likely to generate fairly good results. 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 new approach is also unreliable for smaller geographical levels, and Shaw-Taylor and Wrigley have therefore only used it nationally and on two quite large regions. Had it been applied to a smaller region, such as a single county, in c.1725, the urban subset would have been found to be significantly biased. For example, in Lancashire the sample contains transport-dominated Liverpool but not manufacturing-dominated Manchester. This bias cannot be remedied by urban-rural reweighting. The reweighting methodology cannot therefore fulfil one of the key goals of the Occupational Structure Project, as discussed above: 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 parish register data alone can hope to generate reliable national estimates for the mid or late eighteenth century, let alone for the seventeenth or earlier centuries, because parish registers coverage is simply too low. Nor can the Occupational Structure project's promise of quantitative insight into regional development be realised with parish register data alone. One or more new sources of data are required to fill the gaps. Members of the Cambridge Group, eminently aware of the issue, started collecting non-parish register data some time ago, with the intention of using those data to overcome the problem of limited coverage. It will be shown in this paper how one of these data sources, probate records, can provide a solution to the coverage problem.

Probate documents as an alternative source of occupational information

The information offered by the new data sources will have to be detailed and reliable, provide wide and well-specified geographic coverage, and extend the covered timeframe to well before the eighteenth century. Despite the fact that occupational evidence is actually not scarce in historical sources, meeting these requirements is not easy, as Paul Glennie has shown in his excellent systematic analysis of historical data on men's trades.[10] Occupational descriptions of individuals are not rare, but scattered across many different types of sources such as letters, biographies, literary sources, and official documents. Accurate results can be achieved by linking these sources at the level of

[10] Glennie, *'Distinguishing Men's Trades' : Occupational Sources and Debates for Pre-Census England* (Bristol: Historical Geography Research Group, 1990).

individuals, but this is highly labour intensive and therefore feasible only for local studies or for research limited to specific trades.[11]

Information on historical occupational structures can also be gained from contemporary studies, commissioned by local authorities keen to improve their knowledge of the composition of the local population and economy. But although these may, again, provide valuable information for local historians, coverage over space and time is sparse, typically limited to a handful of parishes for one specific moment in time. The Gloucestershire 1608 muster role used by A.J. and R.H. Tawney has the widest geographically spread of any of these sources in the early-modern period, but even that only covers one county.[12] As Glennie has shown, militia lists provide a detailed and reliable source of occupational information, but they too are only available for a few counties, and only from c.1760.[13] Contemporary social tables, such as those by King, Petty, Massie, and Colquhoun are open to a range of interpretations and of difficult to gauge reliability. Also, of course, they describe the national scale only.

There are really, only two historical sources of sufficient detail, quality, and scope to be suitable for complementing the existing parish register data, namely records of court proceedings and probate documents.[14] And the latter of these two sources has one great additional advantage: its occupational data is relatively easily accessible because it is contained in many of the indexes which have been created to provide easy access to the original documents.

Shaw-Taylor and other members of the Cambridge Group, understanding the need for complementary data and recognising the promise of the probate source, started collecting testamentary indexes on a large scale in 2007. They found many county record offices prepared to supply such data in an electronic form. These were made available to me when I joined the group and, combined with additional indexes collected since, they form the basis for the research in this paper and my forthcoming doctoral thesis.[15] In total, they cover nearly ninety per cent of all English and Welsh counties. For the Interregnum, all English probate documents were proved at the Prerogative Court at Canterbury (PCC) rather than the local dioceses thus providing, in theory, full national coverage for

[11] For examples of such local and/or trade-specific studies of (amongst other things) occupational structure, see Woodward, *Men at Work: Labourers and Building Craftsmen in the Towns of Northern England, 1450-1750* (Cambridge: Cambridge University Press, 1995); Wrightson and Levine, *Poverty and Piety in an English Village: Terling, 1525-1700* (New York: Academic Press, 1979); Levine and Wrightson, *The Making of an Industrial Society: Whickham 1560-1765* (Oxford: Clarendon, 1991); Burn. 'Shiftin' Coal and Mekkin' Ships: Ordinary Newcastle Occupations, C.1600-1720', paper presented at the EHS Annual Conference, Oxford, 2012, http://www.ehs.org.uk/ehs/conference2012/Assets/EHSBooklet2012.pdf, pp. 37-42.

[12] Tawney and Tawney, 'An Occupational Census of the Seventeenth Century', *The Economic History Review,* a5:1 (1934), pp. 25-58.

[13] Glennie, *Distinguishing Men's Trades*, pp. 46-65.

[14] Regarding court record data: T.S. Cockburn has argued that it is 'factually worthless', but as Glennie has demonstrated, that judgement is too harsh, and it can be a valuable source of systematic occupational information. See 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. 42-3, particularly 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.

[15] I am tremendously grateful to Dr Shaw-Taylor for collecting these data, to Dr 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.

the 1650s; an electronic index to the PCC wills was kindly made available by the National Archives.[16] These indexes together provide information on two million men over the 1550 to 1850 period. Nesta Evans was surely right in proclaiming probate documents to be 'a more fertile source of [occupational] information than any other class of document', because of their 'sheer number'.[17]

Unfortunately, probate documents are also a very problematic source of occupational information. This is not because the individual documents are unreliable, as Mark Overton *et al* have argued. They criticised the accuracy of the occupational descriptors in probate records, claiming that the occupation given in the inventory 'often differed from that stated by the decedent in his or her will'.[18] This criticism is severely overblown, as a systematic analysis has demonstrated.[19] But probate records *are* problematic for a very different reason: they suffer from strong social bias. Wealthier estates are much more likely to be represented in the testamentary evidence than poorer ones for a simple reason: the trade-off between, on the one hand, the cost of having a will or inventory made and, on the other hand, the value of such a document in case of disputes over the estate, was more positive for high-value than low-value estates. Financial disincentives for the church courts in registering and exhibiting probate documents for low-value estates probably exacerbated the inherent wealth bias.[20]

Consequently, individuals in capital intensive and/or stock rich occupations such as farmers, brewers, tanners, and traders are severely overrepresented in the probate record. Conversely, the poor and/or those involved in occupations which did not require expensive equipment or raw materials and did not produce much stock are strongly underrepresented; examples of such occupations are weavers, transport workers, tailors, miners, and labourers. The severity of the probate record's occupational bias can be made visible by comparing probate data with parish register data from the same parishes and time period. As figure 3 shows, the probability of an early-nineteenth-century farmer in Cheshire was four times higher than a butcher, twelve times higher than a weaver, and twenty-five times higher than a labourer.[21]

---

[16] An index to these records is available via the National Archives, http://www.nationalarchives.gov.uk/records/wills.htm.

[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.

[18] Overton *et al*, *Production and Consumption*, p. 34.

[19] Keibek and Shaw-Taylor, 'Rural by-Employments', pp. 251, 253-4. Others also found little evidence of the claimed variation in occupations between documents referring to the same individual; see Muldrew, *Food, Energy and the Creation of Industriousness: Work and Material Culture in Agrarian England, 1550-1780* (Cambridge: Cambridge University Press, 2011), p. 166; Shaw-Taylor, 'The Nature and Scale of the Cottage Economy' (unpublished book chapter, Cambridge, 2002), http://www.geog.cam.ac.uk/research/projects/occupations/abstracts/paper15.pdf, pp. 11-2.

[20] Arkell, 'The Probate Process', in Arkell *et al* (eds), *When Death Do Us Part*, p. 12.

[21] 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.

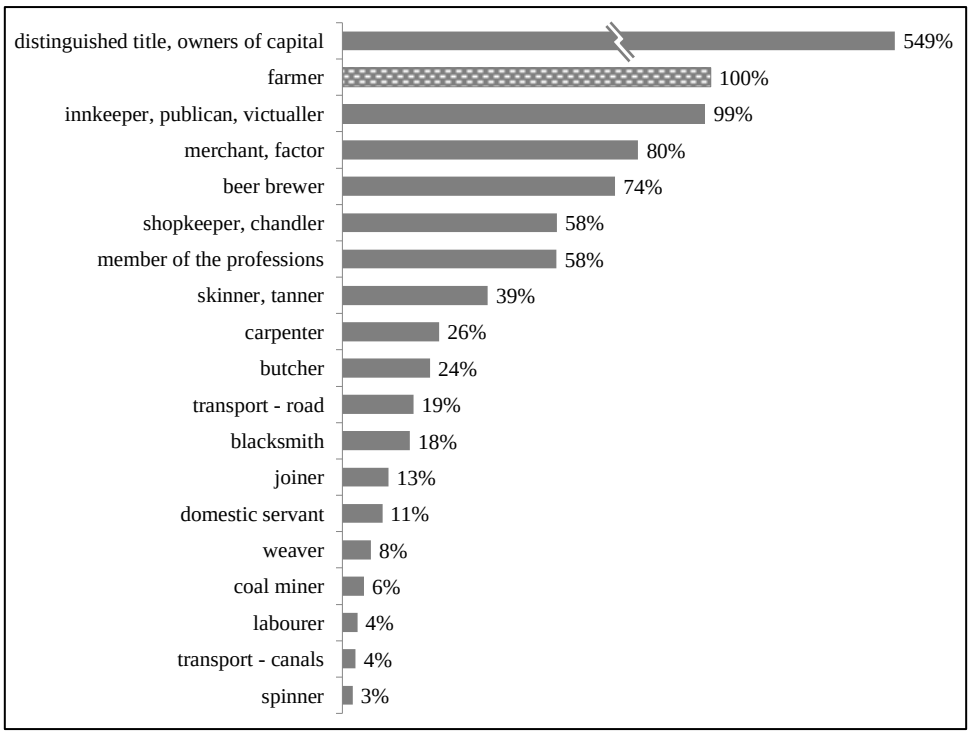

Figure 3. The relative chance of being probated per occupational group in Cheshire, c.1817
(relative to farmer = 100%)

Sources: Parish register database, constructed by the Cambridge Group; probate database; several county record offices.

As a result, occupational structures derived from probate data distort historical reality, and, as figure 4 shows for the county of Cheshire, the distortion is severe.

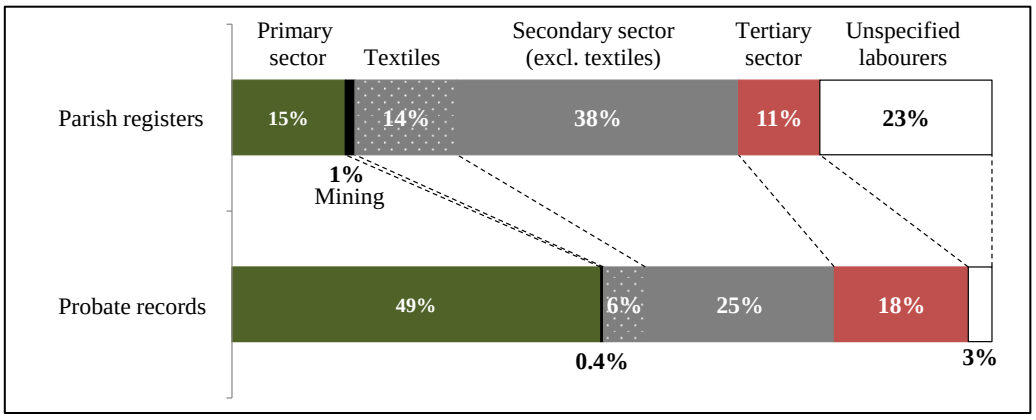

Figure 4. The male occupational structure of Cheshire according to parish registers and probate records, c.1817.

Sources: Parish register database, constructed by the Cambridge Group; probate database; Cheshire Record Office.

Historians have quite frequently underestimated or downplayed probate documents' occupational bias, and used them directly for estimating occupational structures.[22] The most recent example is a 2012 paper by Greg Clark *et al*, which presents occupational estimates for the Interregnum based on the index to PCC wills.[23] Unfortunately, these wills present an even more distorted image of contemporary society than those proved before and after the Interregnum at the local ecclesiastical courts. Comparisons between the PCC and local probate indexes show that annual numbers of wills made were much lower during the Interregnum than in the immediately preceding and following decades. In Cheshire, for example, numbers fell to sixty-three per annum during the Interregnum, a quarter of what they were before and after the period. But numbers fell much more steeply for occupations with an already low chance of being probated, further exacerbating the probate record's occupational bias. For example, numbers fell three times faster for labourers than for farmers.

Given the vast differences in probabilities of being probated between occupations, depicted in figure 3 above, it might seem unlikely that probate documents can be used to determine reliable, unbiased historical occupational structures at all. However, the probabilities in figure 3 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.

A case example: Cheshire, c.1725-1817

In total, 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 exercise, this 120-year period is divided into four intervals of thirty years, the mid-points of which correspond with the mid-points of the parish register data, that is, 1725, 1755, 1785, and 1817. Thus, a temporal match is 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 figure 4 using 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.

---

[22] For example, in 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; 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.

[23] Clark, Cummins, and Smith, 'Malthus, Wages, and Preindustrial Growth', *The Journal of Economic History,* 72:02 (2012), pp. 364-92.

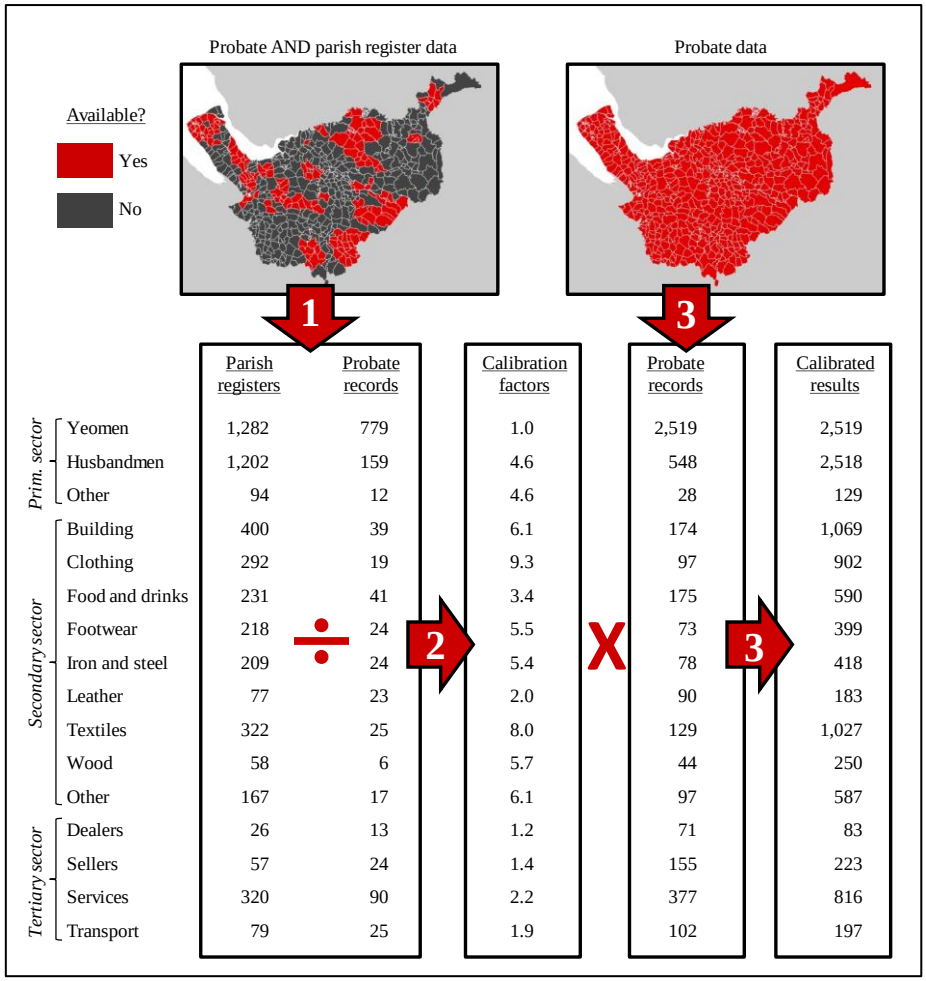

Figure 5. Schematic depiction of the calculation of an unbiased male occupational structure for Cheshire, c.1725.

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

Sources: see figure 4.

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.

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 only available at the level of parishes or, if they registered baptisms separately, chapelries. These 'Anglican Registration Units', typically contain several townships and hamlets.[24] Using a variety of sources of historical geographical information, the townships and hamlets mentioned in the probate data can be carefully 'mapped' onto Anglican Registration Units, 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 geographic match between probate and parish register data cannot be achieved in the manner described above. As other large towns, Chester was divided into several parishes. Not all of these Chester parishes recorded male occupations.[25] As figure 6 shows, a substantial share of Chester's labour force is not 'covered' by parish register data.

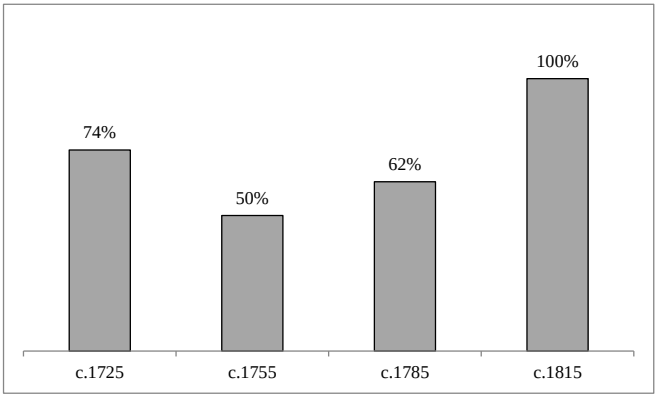

Figure 6. Estimated share of population in the 'covered' parishes; Chester city, c.1725-1815.

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 was the case. Probate documents of men from Chester men 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

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

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.

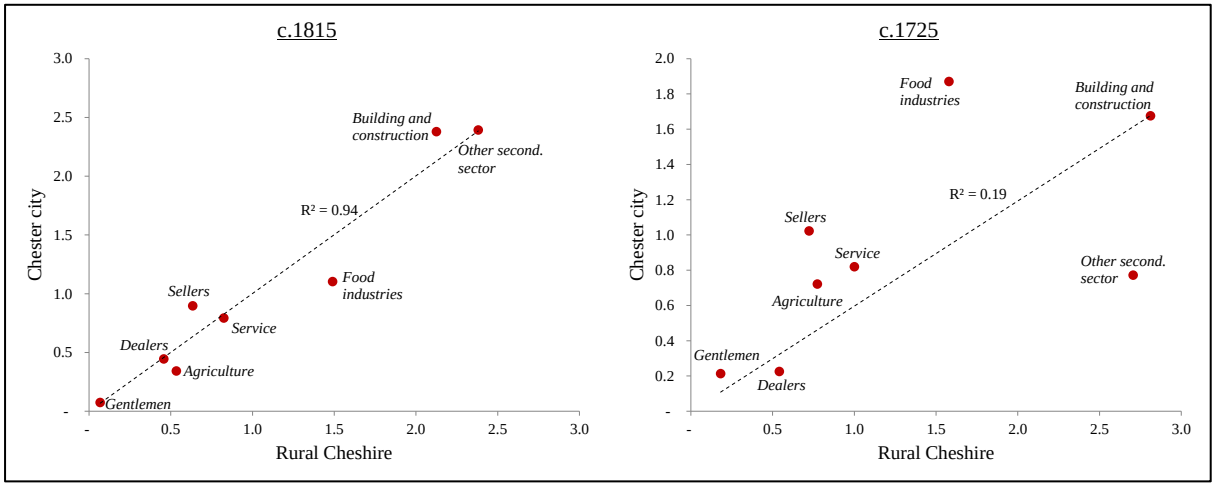

Figure 7. Probate calibration factors compared (indexed; average male decedent = 1).

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.

Sources: see figure 4.

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. To ensure that this does not affect the overall weight of Chester within the occupational structure of the county, the occupational structures for rural Cheshire and the city of Chester are calculated separately, and then combined, using contemporary population figures as weights.[26] 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 2, 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 in figure 2 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 much more gradual and logical development. The calibrated occupational structures for c.1755 and c.1785 differ particularly strongly from those based on parish registers alone.

[26] For sources, see those for figure 4.

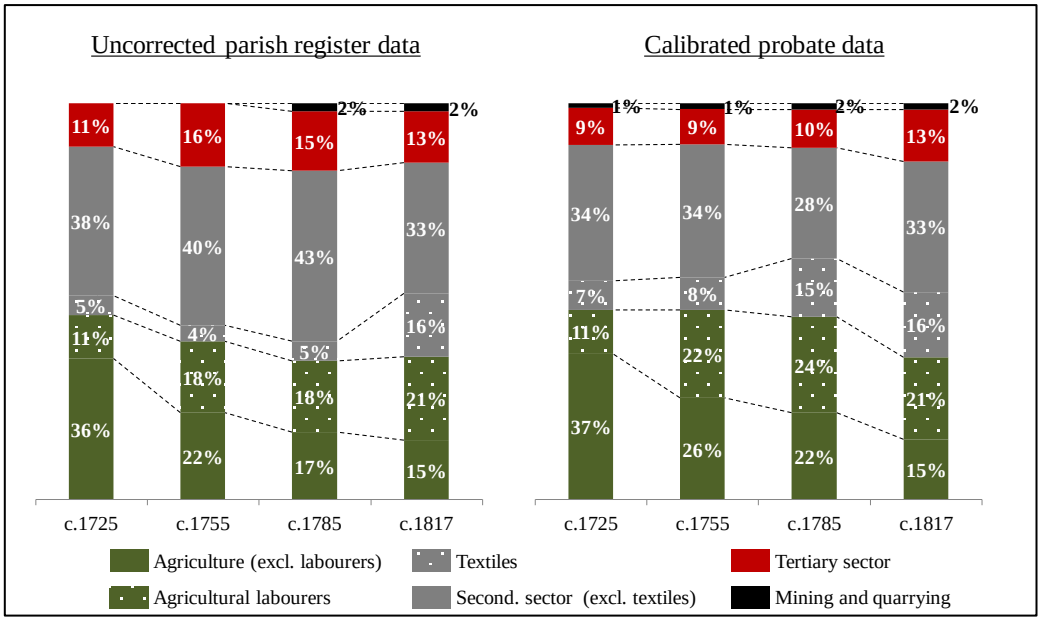

Figure 8. The occupational structure of the county Cheshire, 1725-1817, before (left) and after (right) applying the correction methodology.

Sources: see figure 4.

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.'[27] 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. The latter are used to calibrate the former, whilst the former are used to interpolate the latter. 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

[27] Glennie, *Distinguishing Men's Trades*, p. 32.

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.'[28] 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.[29] It will be shown, however, that such fears are exaggerated; probate calibration factors are, in fact, fairly constant over time and place.

Constancy over time is explored in figure 9, in which calibration factors for Cheshire c.1725 are compared with those for c.1817. Despite being nearly a century apart, the calibration factors are remarkably similar.

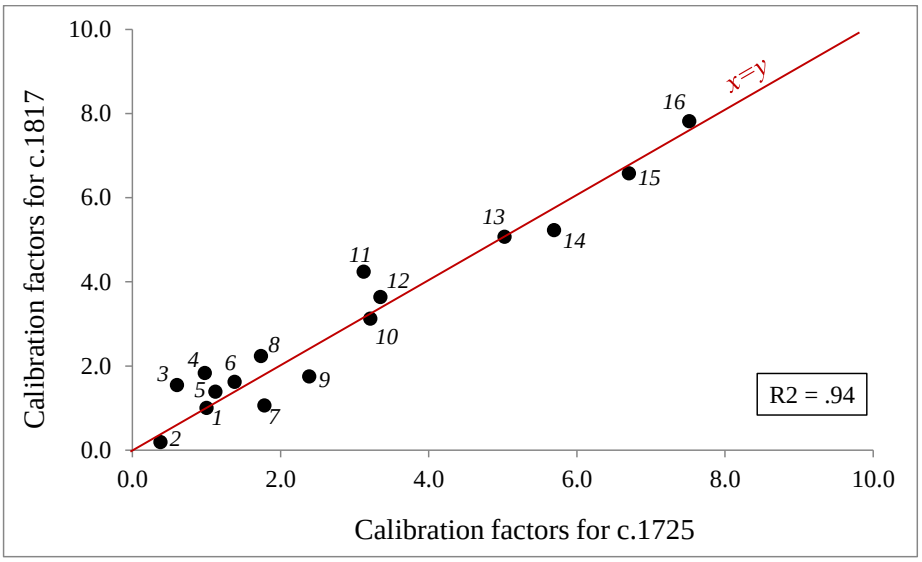

Figure 9. A comparison between probate multipliers per occupational group; Cheshire[1], c.1725 *vs.* c.1817 (indexed; yeoman farmer = 1).

Notes: every data point in the chart represents an occupation. Only occupations which occurred ten times or more in the parish register and probate records of both time periods were included, to ensure that the calibration factors were sufficiently accurate for a fair comparison. The numbers next to the data points refer to the occupations: 1=yeoman/farmer, 2=distinguished title/capital owner, 3=commercial services, 4=skinner/tanner, 5=dealers and merchants, 6=sellers and shopkeepers, 7=innkeeper/publican, 8=ship's officer, 9=professions, 10=crop producer/market gardener, 11=butcher, 12=miller, 13=(black)smith, 14=carpenter/joiner, 15=tailor/shoemaker, 16=textile worker (excl. spinner). Note that this chart was made at a much lower level of occupational detail than the occupational clusters used in, for example, figure 3. This was done to ensure a like-for-like comparison, as the composition of higher-level occupational clusters could change significantly over time. For example, the 'textiles' cluster in figure 3 contains a large number of male spinners for c.1817, but no such male spinners existed in c.1725, as it was an exclusively female occupation then.

[1] Excluding the city of Chester, as a direct, like-for-like comparison between probate and parish register data is not possible for that town in c.1725 – see the discussion on this in the main text.

Sources: electronic index to Cheshire probate documents; parish register data collected by the Cambridge Group.

[28] Ibid, p. 40.

[29] 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.

It should be noted that the calibration factors in figure 9 are indexed, with the calibration factor for yeomen/farmers set at 1. In other words, they are relative, not absolute measures of the likelihood of occupational groups being probated. In absolute terms, calibration factors *do* vary over time, as the general tendency to leave a will or inventory was not constant; a yeoman/farmer was sixteen per cent more likely to leave a probate document in c.1725 than in c.1817. However, it is irrelevant for the described methodology whether, for example, a carpenter in the early eighteenth century had a lower or higher absolute likelihood of being probated than one a century later. What *is* important is whether that likelihood changed *in comparison to other occupations*. As figure 9 shows, that was not the case for Cheshire. The overall tendency to leave a will or inventory may have changed over time, but that did not affect the social and occupational bias in the probate record. Clark's suspicion that the decreasing share of men being probated in the course of the eighteenth century might mean that the subset of probated men was increasingly socially biased – and, therefore, that the consumer revolution was merely a statistical artefact, caused by a sample of changing social composition – is not born out.[30]

The fact that probate multipliers were constant over long time intervals, at least in a relative sense, means that 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 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 an economic historical position, since it coincides with the immediate run up to and initial phase of the industrial revolution. As probate data *are* available for the period in almost all English and Welsh counties, this problem can now be solved. Figure 10 provides an example for Wiltshire.

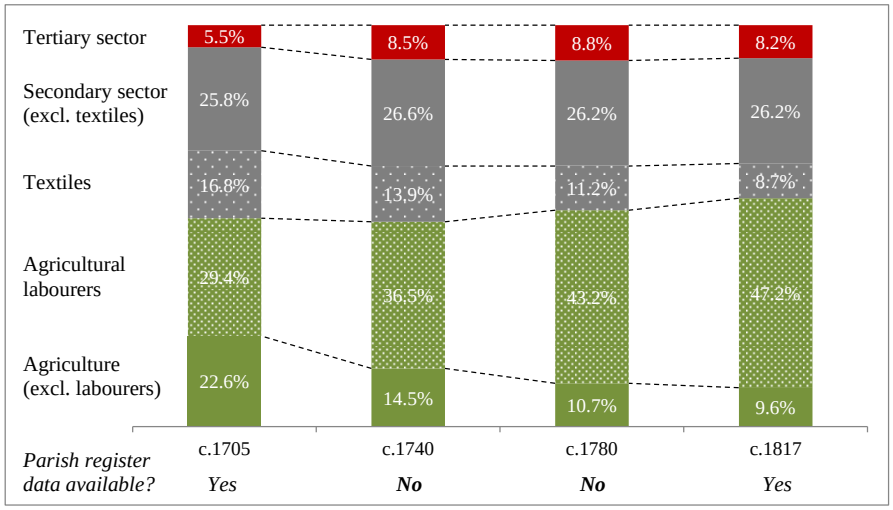

Figure 10. Development of Wiltshire's male occupational structure, according to calibrated probate data, c.1705-1817

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.

[30] Clark, 'Statistical artifact'. Analysis of the wealth bias in probate inventories also shows that it did not increase over time, even when inventories became increasingly rare, see: Keibek and Shaw-Taylor, 'Rural by-Employments', p.265.

The apparent constancy over time of the calibration factors cannot only be profited from to *inter*polate between two points in time for which parish data are available, as in figure 10, 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 11, this has been done for Cheshire.

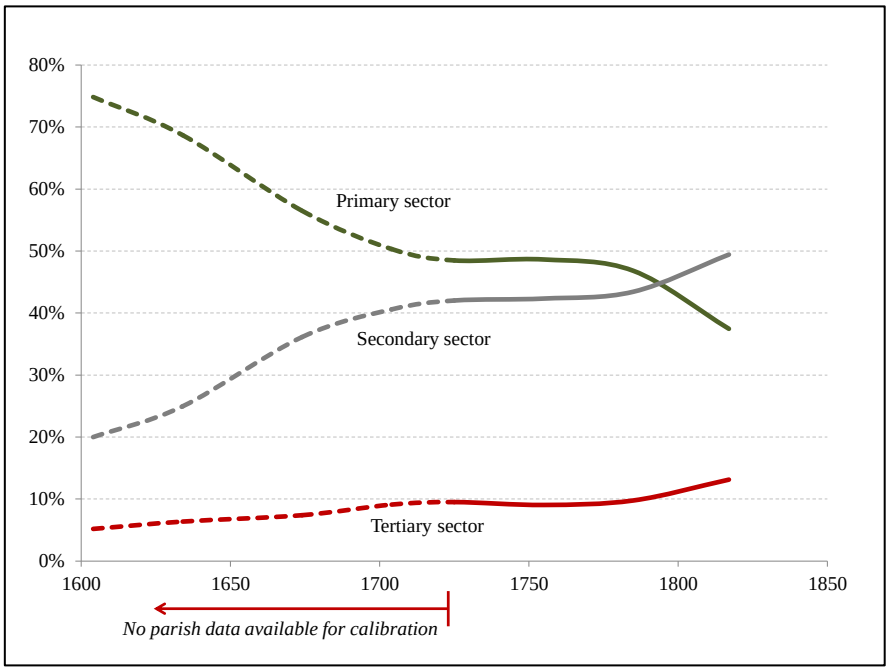

Figure 11. Cheshire male occupational structure development, c.1600-1817.

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.

Sources: see figure 4.

Admittedly, this extrapolation is more speculative than the interpolation in figure 10. For interpolation, we merely have to assume that calibration factors were constant over a certain time period, already knowing that they were the same at the beginning and end of that period. For extrapolation, we only really know the calibration factors at one end of the period. Even though the social and occupational bias of the probate record did not change over the course of the eighteenth century, they may have changed over the course of the seventeenth. It is, however, possible to compare the occupational structure according to the calibrated probate record in the early seventeenth century with an unbiased, high-quality, independent source of occupational information for one specific county: Gloucestershire. For this county, as discussed, a highly detailed and complete muster roll is available for 1608, providing the material for 'An occupational census of the seventeenth century', as the Tawneys called it.[31] As figure 12 shows, the muster roll and calibrated probate data come to very similar conclusions for the male occupational structure of the early seventeenth century, thus suggesting that eighteenth-century derived calibration factors can justifiably be applied to probate data that are a century older.

[31] Tawney and Tawney, 'An Occupational Census'.

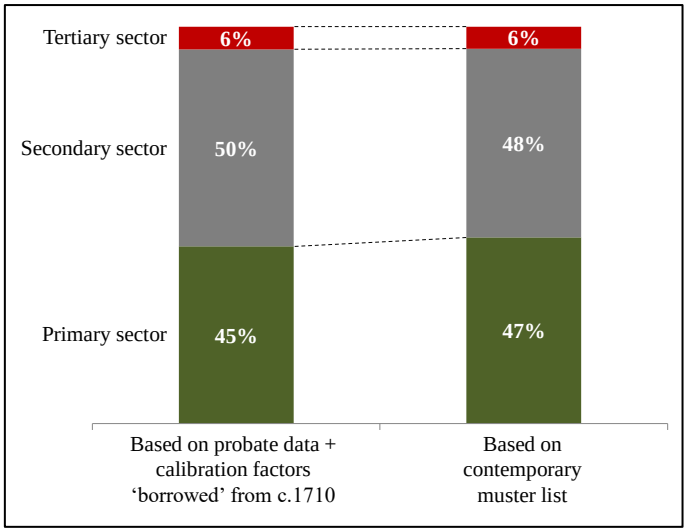

Figure 12. Comparison of estimates of the male occupational structure of Gloucestershire[1], c.1608.

Notes: [1] excluding Bristol and Bilbury Peculiar, as these were not represented in the probate index.

Sources: Gloucestershire probate index; parish register data for Gloucestershire c.1710, collected by the Cambridge Group; Gloucestershire muster list.

Rather than 'borrowing' calibration factors from other moments in time, it is also possible to use calibration factors from other areas, as they turn out to be fairly constant over medium distances in place as well. As figure 13 shows, calibration factors for neighbouring or nearby counties are quite similar. For areas of England and Wales, occupational data from eighteenth century parish registers are either completely unavailable or exist in numbers too small to allow for determining reliable local calibration factors. This is, for example, the case for large and economically important English counties like Norfolk, Essex, and Devon, as well as for the entirety of Wales. Eighteenth century male occupational structures can be calculated for areas like these by applying contemporary calibration factors determined for nearby counties.

In conclusion then, the probate-calibration methodology appears to be quite robust and widely applicable. Probate calibration factors are stable over fairly long distances in time and place, thus justifying 'borrowing' them from other areas or time periods when no local and/or contemporary parish register data are available. Given the much wider availability of probate data than pre-Rose's Act parish register occupational data, both temporally and geographically, this provides us with a powerful tool for deriving local and national male occupational structures for the eighteenth and even, at a lower degree of precision, for the seventeenth century.

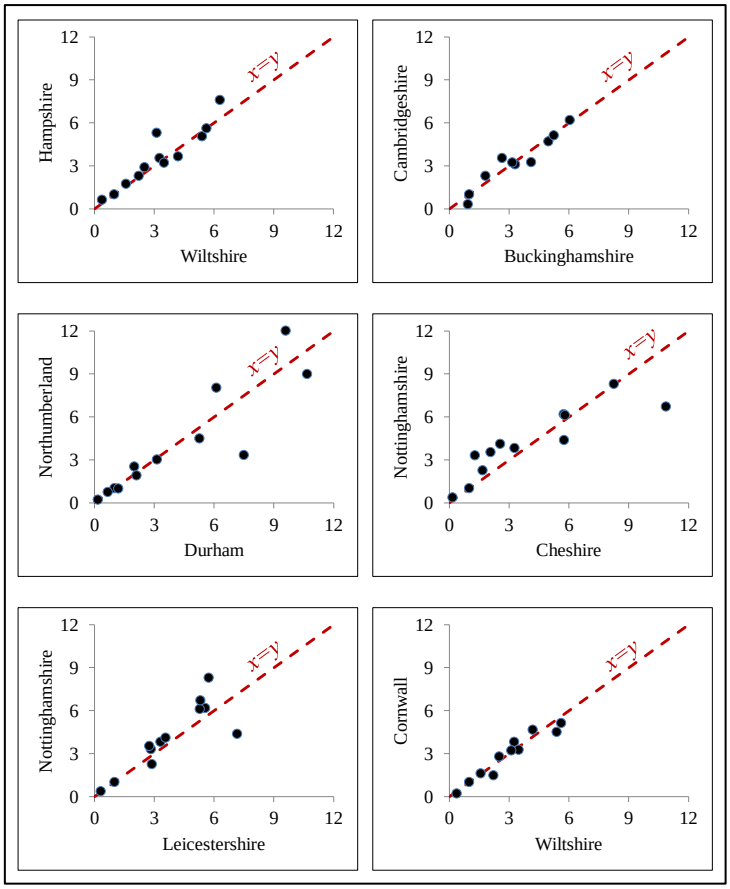

Figure 13. Comparison of probate calibration factors between counties which are geographically close, c.1817 (indexed; yeoman/farmer=1)

Notes: every data point in the charts represents an occupational group. Only groups with sufficiently many observations in both counties, in both parish register and probate data were included, to ensure an accurate comparison.

Sources: local probate indexes, obtained from several record offices; parish register data collected by the Cambridge Group.