

From probate inventories to households: correcting the probate record for wealth bias

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 version will be included in my PhD thesis 'The male occupational structure of England and Wales, 1650-1850'.

Abstract

Wealthy, capital-intensive estates are severely overrepresented in historical collections of probate inventories. This paper discusses a new methodology to neutralise the wealth bias in this important historical data source, enabling one to use probate inventories as a source of information on all households, rather than merely on the biased, probated subsection. The methodology is based on establishing the probability of being inventoried – in a certain time period and geographical area – as a function of the value of the decedent's estate. The probability function is established by means of an iterative fitting process, using occupational information from contemporary parish registers to establish the target values. The inverse of the resulting probability function is subsequently used to reweight the probate inventory dataset. Thereby the historical process which created the inventoried household subsection in the first place is reversed, thus providing an unbiased view of the full population of households. A number of example applications demonstrate the value of the approach.

Probate inventories are a powerful source of early modern historical information, as has long been recognised by historians of the period. They have been used to improve our knowledge of numerous aspects of the lives (and deaths) of early modern women and men, from the type of goods they owned, the ways in which they earned a living, the crops they grew and the livestock they kept, to the ways in which they arranged their financial affairs, the positions they occupied in the social structure, the wealth they enjoyed relative to others, *etcetera*. Probate inventories allow us to 'sketch individual lives' of these early modern men and women but arguably demonstrate their strengths most convincingly when used to 'build aggregate estimates'.¹ As sources on individual people and households they suffer from, as Overton *et al* have phrased it, 'a depressingly long list of possible reasons why any single inventory may be misleading', but these shortcomings are generally much reduced when probate inventories are used in large numbers, as a statistical source.² However, it is precisely then when probate inventories also display one of their most serious and tenacious weaknesses: wealth bias.

¹ Lindert, 'An Algorithm for Probate Sampling', *The Journal of Interdisciplinary History*, 11:4 (1981), p. 649.

² Overton *et al*, *Production and Consumption in English Households, 1600-1750* (London: Routledge, 2004), p. 31.

Not every death led to a probate inventory being drawn up. Early modern children and married women were often not inventoried because in many countries it was the male 'household head' who was the legal owner of all property in the household. But men were often not inventoried either, particularly when they were not well off. The trade-off between, on the one hand, the cost of having an inventory made and, on the other hand, the value of such an inventory in case of disputes over the estate, was naturally more positive for wealthy than poor estates.³ Also, even when all decedents were in principle under legal obligation to be probated, the authorities might in practice be quite lenient, particularly when the decedents were poor. As a result, 'like other seemingly broad, sources in social history, probate records represent the experience of an atypically prosperous segment of the population'.⁴

Historians have differed in their approach to this social skew within the probate record. Some have simply accepted it as an unfortunate but unsalvageable drawback of this otherwise so treasured historical source, conceding that their conclusions were not valid for historically society in its entirety but limited to its inventoried share. Others have attempted to fill in the gap of the non-inventoried households by a 'best guess' approach. And a few have attempted to develop more sophisticated adjustments, to apply to the collection of probate inventories itself or, *a posteriori*, to the analytical results derived from it. This paper falls squarely in the latter tradition. Its aim (and claim) is to presents a new and superior method to correct the probate record for wealth bias, enabling one to draw accurate conclusions from probate evidence about contemporary society as a whole. Or, to express it in more technical terms, to allow one to analyse the full population of households from which the probate sample was drawn, despite the historical sampling method having been non-random.

The paper is divided into four sections. First, the wealth-bias problem is investigated in more detail and across geographies. Second, existing attempts at dealing with the problem are discussed (and found wanting). Third, the new approach is introduced and explained in detail. Fourth, the approach is demonstrated for a number of example applications.

³ Arkell, 'The Probate Process', 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. 12.

⁴ Smith, 'Underregistration and Bias in Probate Records: An Analysis of Data from Eighteenth-Century Hingham, Massachusetts', *The William and Mary Quarterly*, 32:1 (1975), p. 106.

I

Most historians who use probate inventories will have seen themselves confronted with the problem of wealth bias at some moment in their research. The severity of the problem will have depended on a number of things. If an inventory is used solely as an illustration, or to study an individual or a single household, the problem of wealth bias is, of course, very slight; it merely limits one in the choice of cases, but does not affect the findings about the cases which are actually chosen. But probate inventories are, as discussed, more commonly used to determine aggregate historical measures by analysing significant groups of inventories. Even then, wealth bias might be considered only a minor problem if all or almost all members of the social group that one is studying left an inventory, for example, if that group formed a small, wealthy social sub-stratum within the overall contemporary population. Or, if one is exceptionally lucky and finds that, for some reason, almost all households were inventoried in the geographical area and period that is being analysed – as Gloria Main found for late seventeenth-century colonial Maryland.⁵

Such ‘universal coverage’ is, unfortunately, not generally encountered. In early modern Britain, the focus of so much probate-based research into the agricultural, consumer and industrious revolutions, only a minority of households actually left probate inventories, especially in the eighteenth century. Greg Clark recently calculated that less than half the male decedents in Essex, Kent, Buckinghamshire and Suffolk were probated in 1600, declining to less than one in five in 1700.⁶ Clark’s figures are based on men who were probated, many of which left a will but not an inventory, so the share of inventoried male decedents must have been even lower. The 1690-1730 inventory record of Cheshire will serve as a basis of data for some of the key calculations in this paper, so it is worthwhile to try and estimate its coverage of contemporary householders. Jon Stobart has argued that around forty per cent of the 1700-1760 Cheshire population was probated, but it is unclear how this figure was derived.⁷ Indeed, as will now be shown, a comparison between male deaths and probate record numbers would suggest a much lower figure. Using age-dependent demographic data from family reconstitution and Tony Wrigley’s recent work on county populations, it is possible to generate reliable estimates for the number of male householders’ deaths in contemporary Cheshire.⁸ The number of male householders who were probated or, more specifically, inventoried can be obtained from the indexes for the probate documents of Cheshire decedents proved at the Chester diocesan consistory court and for the higher consistory courts of York and Canterbury.⁹ By combining the thus derived numbers of deaths and probate records, inventory coverage can be determined. In table 1, the results of this procedure are shown in decennial intervals, covering the hundred years between the Restoration and the Industrial Revolution.

⁵ Main, *Tobacco Colony: Life in Early Maryland, 1650-1720* (Princeton: Princeton University Press, 1982), pp. 282-92.

⁶ 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>, fig. 4.

⁷ Stobart, ‘The Economic and Social Worlds of Rural Craftsmen-Retailers in Eighteenth-Century Cheshire’, *The Agricultural History Review*, 52:2 (2004), p. 144. Stobart refers here to Stobart, ‘Geography and Industrialization: The Space Economy of Northwest England, 1701-1760’, *Transactions of the Institute of British Geographers*, 21:4 (1996), pp. 682-84, but this too provides no clear derivation of the forty per cent estimate.

⁸ Wrigley *et al*, *English Population History from Family Reconstitution, 1580-1837* (Cambridge: Cambridge University Press, 1997); Wrigley, *The Early English Censuses* (Oxford: British Academy Records of Economic and Social History, new series, 2011).

⁹ Electronic indexes are available from the Cheshire Record Office, the National Archives, and – via Origins.net – the Borthwick Institute for Archives.

Table 1. *Calculation of probate and, more specifically, inventory coverage amongst adult male householders in Cheshire, 1661-1760*

	1661-70	1671-80	1681-90	1691-00	1701-10	1711-20	1721-30	1731-40	1741-50	1751-60	
National	Average age of first marriage (men) ¹	27.4	28.0	27.7	27.1	27.4	27.3	27.0	26.9	26.1	
	Section of population above this age	48.5%	49.3%	50.6%	49.2%	47.2%	46.4%	47.3%	48.3%	48.1%	
	Mortality rate for this pop. section (per '000) ²	33.3	32.8	42.1	30.9	33.2	32.0	37.7	29.2	27.4	24.8
Cheshire	Population size (decennial mean) ³	91,560	90,438	88,880	90,164	94,592	100,691	107,183	114,094	121,450	135,565
	Male householders' deaths (decennial) ⁴	7,396	7,329	9,487	6,864	7,407	7,492	9,545	8,043	8,039	8,084
	Probated male decedents (decennial) ⁵	2,519	2,443	2,633	1,983	1,826	1,733	2,755	1,937	1,645	1,530
	Probate coverage for male householders	34%	33%	28%	29%	25%	23%	29%	24%	20%	19%
								24%			
	Inventoried male decedents (decennial) ⁵	2,105	2,174	2,178	1,191	1,030	895	1,217	590	315	212
	Inventory coverage for male householders	28%	30%	23%	17%	14%	12%	13%	7%	4%	3%
							14%				

Sources: Database of probate records from the Cheshire Record Office; 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:

- ¹ Bachelor and spinster marriages.
- ² Age specific mortality rates ($_{10}M_x$) derived from probabilities of dying per age interval ($_{10}q_x$) as provided by Wrigley *et al*, employing the relationship $_{n}q_x = 2 \times n({}_nM_x) / [2 + n({}_nM_x)]$.
- ³ Cheshire specific estimates are only available for 1600, 1700, 1750 and 1761. Intermediate years before 1700 were interpolated using national totals as a guide, since Cheshires overall population growth between 1600 and 1700 was in line with the national total. After 1700, when Cheshire's population development clearly differed from the national, an exponential interpolation was employed, that is, constant population growth rates were applied between the 1700, 1750 and 1761 estimates.
- ⁴ Calculated from the rows above, assuming fifty per cent of adult deaths to be male.
- ⁵ As obtained from the probate record databases for the diocesan consistory court of Cheshire and the higher consistory courts of York and Canterbury.

As table 1 shows, the number of deaths of male householders – here taken as all men above the average age of marriage – which led to some kind of probate document declined over the period from thirty-four to nineteen per cent. Rather than Stobart's forty per cent, less than a quarter of Cheshire's adult male population appears to have been probated in 1700-60. Table 1 also shows that the number of *inventoried* householders was even lower, and declined much faster. Over the 1690-1730 period, only one in seven male householders' deaths resulted in a probate inventory. Similar calculations for other counties show that this low probate inventory coverage was not exceptional. For the same period, the inventoried share of male householders in Nottinghamshire, Wiltshire, and County Durham appears to have been fifteen, nine and six per cent, respectively.¹⁰

In other early modern geographies, inventory coverage was generally higher than in England, but, again, far from universal. William Davisson calculated that about one in four mid-seventeenth-century household heads in Essex County, Massachusetts left an inventory.¹¹ By carefully comparing the exceptionally high-quality death records for the first parish of Hingham, Massachusetts with the probate record, Daniel Smith showed that forty-two per cent of adult men who died there between

¹⁰ The probate record and inventory numbers required for these calculations were derived from electronic indexes kindly made available by the respective record offices, complemented with indexes to the probate documents proved at the higher consistory courts of York and Canterbury (see previous footnote).

¹¹ Quoted in Smith, 'Underregistration', p. 101.

1726 and 1786 were inventoried.¹² Alice Hanson Smith calculated that whilst nearly seventy per cent of the wealth holders in the Middle Colonies who died in 1774 were inventoried, this was the case for only thirty-three per cent of wealth holders in New England.¹³ By comparing the probate record with tax lists, Main calculated that the share of taxable households which left an inventory differed considerably between towns in seventeenth-century Massachusetts, ranging from as low as thirty-seven to as high as eighty-five per cent; it should be noted, however, that tax lists may themselves be a wealth-biased source and exclude poor households, so actual inventory coverage may have been lower.¹⁴

In continental Europe too, inventory coverage differed widely. In central Sweden, only ten per cent of deceased adults were inventoried in 1770, rising to more than forty per cent by the 1830s.¹⁵ Anton Schuurman calculated for the nineteenth-century Zaanstreek in Holland that an inventory exists for only a third of deceased which left minors, and who were therefore, in theory, legally obliged to be inventoried.¹⁶ Erkki Markkanen estimated that despite the legal obligation of every adult dead to be inventoried, in practice this only happened in a quarter of the cases in late-nineteenth-century Finland.¹⁷ Sheilagh Ogilvie *et al* found that for the small town of Wildberg in Württemberg, the local presence of bureaucrats combined with clear financial incentives stimulated these officials to enforce the legal obligation on their fellow-citizens to draw up inventories at several moments during their lifetime, leading to over eighty per cent of male taxpayers by 1695 to be linkable to at least one such inventory. But inventory coverage appears to have been much patchier in the early seventeenth century, when it may have been as low as one in three male decedents.¹⁸

Patchy coverage does not, in theory, preclude representativeness of the inventories that *were* made but, as suggested above, in practice, patchy coverage was at least partly the *result* of the low probability of some groups in society to be inventoried. Although there might be legal exemptions from being inventoried for certain small groups in society, such as for certain high status groups in early modern Württemberg, often all household heads were, in principal, legally obliged to leave an inventory at death.¹⁹ In practice, however, things were often considerably more relaxed and inventories were not always drawn up, particularly if there was little to inventory. In early modern England, the authorities responsible for managing the probate process even had certain financial incentives to *discourage* the poor to draw up inventories.²⁰

¹² Ibid, p. 104.

¹³ Jones, 'Wealth Estimates for the New England Colonies About 1770', *The Journal of Economic History*, 32:1 (1972), pp. 115-16.

¹⁴ Main, 'Probate Records as a Source for Early American History', *The William and Mary Quarterly*, 32:1 (1975), p. 98. On the dangers of using colonial-American tax registers as a source for calculating probate coverage, see Lockridge, 'A Communication', *The William and Mary Quarterly*, 25:3 (1968), p. 517, fn. 4.

¹⁵ Lindgren, 'The Modernization of Swedish Credit Markets, 1840-1905: Evidence from Probate Records', *The Journal of Economic History*, 62:3 (2002), p. 818.

¹⁶ Schuurman, 'Some Reflections on the Use of Probate Inventories as a Source for the Study of the Material Culture of the Zaanstreek in the Nineteenth Century' in Van der Woude and Schuurman (eds), *Probate Inventories: A New Source for the Historical Study of Wealth, Material Culture and Agricultural Development* (Utrecht: HES, 1980).

¹⁷ Markkanen, 'Das Finnische Erbschaftsinventarmaterial' in Van der Woude and Schuurman (eds), *Probate Inventories*.

¹⁸ Ogilvie, K pker, and Maegraith, 'Household Debt in Seventeenth-Century Württemberg: Evidence from Personal Inventories', *Cambridge Working Papers in Economics*, 1148 (2011), <http://www.econ.cam.ac.uk/dae/repec/cam/pdf/cwpe1148.pdf>, p. 15.

¹⁹ Ibid, p. 11

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

The result is a severe underrepresentation of, in particular, the less wealthy members of society. In figure 1, the relative probability of leaving an inventory has been depicted as a function of average wealth by occupational group in early-eighteenth-century England. Probabilities were calculated by comparing the number of probate inventories in each occupational to the occupational structure, derived from parish register data.²¹ Wealth per occupational group was approximated by the median inventory total for all inventories of that occupation in the dataset.²² The relationship is not perfect, and cannot be expected to be since figure 1 ignores the wealth distribution within occupations, and since the probability function governing the chance of leaving an inventory as a function of wealth was not a linear one, as will be discussed in section 3 of this paper. Nevertheless, figure 1 confirms the importance of wealth as a determinant in whether or not a decedent would leave a probate inventory.

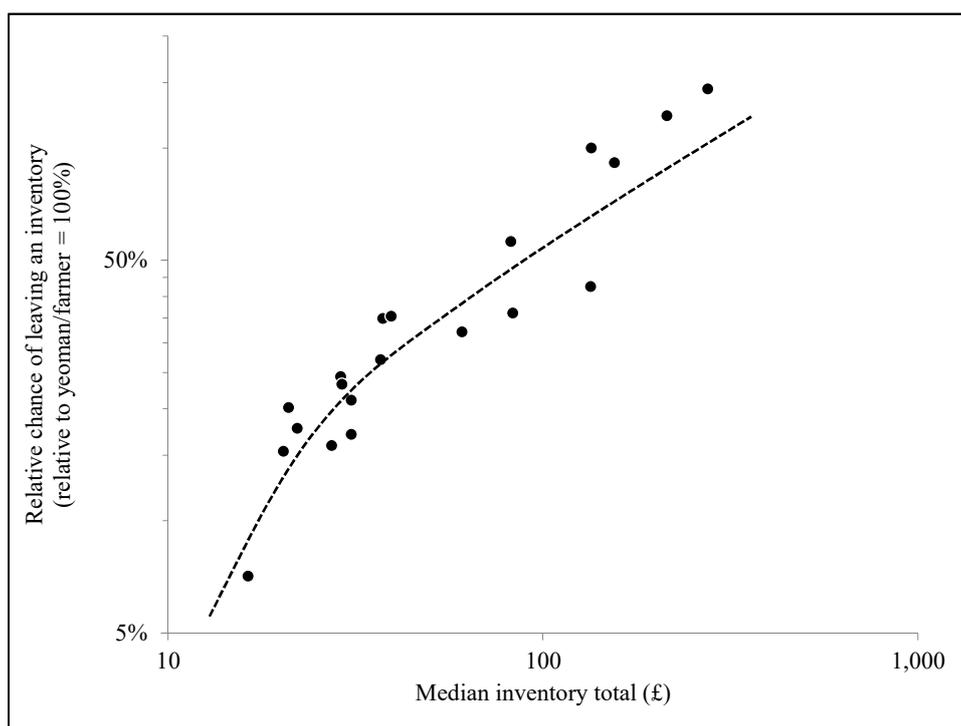


Figure 1. The relationship between median inventoried wealth and the (relative) chance of being inventoried, by occupational group (early-eighteenth-century England)

Notes: Each point in the chart represents an occupational group. Inventories were taken from six English counties/regions, namely Cheshire, Lancashire, Staffordshire, Northamptonshire, Salisbury Diocese, and Lincolnshire.

Sources: probate inventory dataset; parish register data collected by the Cambridge Group; indexes to probate data from several county record offices.

²¹ A similar procedure as the one described in Keibek, 'Using probate data for estimating historical male occupational structures', *Campop paper*, <http://www.campop.geog.cam.ac.uk/research/projects/occupations/abstracts/paper27.pdf>.

²² A description of the inventory dataset can be found at Keibek, 'By-employments and occupational structure in pre-industrial England', *Campop paper*, <http://www.campop.geog.cam.ac.uk/research/projects/occupations/abstracts/paper27.pdf>, table 1, pp. 14-5.

In areas where a much higher share of decedents left an inventory than in early modern England, wealth bias is nevertheless an issue. Despite high inventory coverage, Ogilvie *et al* found that the 23.8% of Wildberg's couples which had less than 34 gulden in wealth at the time of marriage left only 2.1% of the marriage inventories, and that the 8.3% of individuals with less than 34 gulden in wealth at death left only 3.7% of the death inventories.²³ For Hingham, Massachusetts, Smith found that the wealthiest forty per cent of inhabitants was five times more likely to be inventoried than the poorest twenty per cent, and that the average value in real property of inventoried households was more than three times that of the non-inventoried.²⁴ By comparing the burial tax registers with the extant probate record for Amsterdam in the 1701-10 period, Johannes Faber showed that less than five per cent of the citizens belonged to the three highest tax classes but these left more than a quarter of the inventories.²⁵ Using the same analysis, Thera Wijzenbeek calculated that the highest three burial tax classes were overrepresented with roughly a factor three in the collection of early-eighteenth-century Delft orphanage inventories.²⁶ Overton *et al* found for two Cornish parishes that thirty-three per cent of the local households were exempted from the hearth tax because they were considered too poor, but these left only one per cent of the inventories.²⁷ In Kent, they found that the forty-five per cent of households with just one hearth left only twenty-one per cent of the inventories, whereas the thirty-two per cent of the households that had three hearths or more left fifty-six per cent of the inventories.²⁸ For late-seventeenth-century Warwickshire, Tom Arkell found that thirteen per cent of householders liable to the hearth tax could be matched to probate records, compared to only one per cent of the non-liable.²⁹

Labourers, who not only owned few household goods because of poverty, but few work-related goods as well – as these would have been owned by their employers – are often particularly underrepresented. Anyone who has used the printed or electronic catalogues of probate records at an English record office will immediately have noticed how difficult it is to find more than a handful of probate inventories for which the deceased's occupation is listed as 'labourer', particularly compared to the ubiquitous 'yeoman' farmers. Less than one per cent of the inventories in Weatherill's large national sample on early modern England were labourers.³⁰ Labourers form the bottom left data point in figure 1, above. The underrepresentation of labourers was often equally pronounced in continental Europe. From Carl-Johan Gadd's data on mid-eighteenth-century probate inventories from rural southern Sweden, it can be calculated that peasants were twelve times more likely to leave an inventory than crofters, cottagers, servants and other 'totally landless' men.³¹ And Schuurman's data on the nineteenth-century Zaanstreek show that labourers were underrepresented in the inventory record there by a factor of three (whilst farmers were overrepresented by a factor of four).³²

²³ Ogilvie *et al*, 'Household Debt', pp. 12-13.

²⁴ Smith, 'Underregistration', pp. 105-106.

²⁵ Faber, 'Inhabitants of Amsterdam and Their Possessions, 1701-1710' in Van der Woude and Schuurman (eds), *Probate Inventories*.

²⁶ Wijzenbeek, 'Delft in the Eighteenth Century' in Van der Woude and Schuurman (eds), *Probate Inventories*.

²⁷ Overton *et al*, *Production and Consumption*, p. 23.

²⁸ *Ibid*, p. 25 (table 2.5).

²⁹ Arkell, 'The Incidence of Poverty in England in the Later Seventeenth Century', *Social History*, 12:1 (1987), p. 33.

³⁰ Weatherill, *Consumer Behaviour and Material Culture in Britain, 1660-1760*, 2nd edn (London: Routledge, 1996), pp. 210-11.

³¹ Gadd, 'Swedish Probate Inventories, 1750-1860', calculated by comparing the inventory coverage of peasants and their wives with that of all the other groups in table 1, 1748-57 data, p. 230.

³² Schuurman, 'Some Reflections', pp. 182-83.

Before discussing potential solutions for the wealth bias issue, it should be remarked here that there are other potential biases in the probate record. There may have been reasons why one decedent was inventoried whilst another one was not in addition to the value of the decedent's estate. It seems logical that large families had a clearer need for an inventory of possessions than small ones, as they had more potential heirs, complicating the inheritance. Statistical analysis shows that estate value and household size were very strongly correlated.³³ This strong correlation is not surprising. Large families required more household goods, amongst which were very expensive goods like beds. Also, the earnings potential of large families was greater than that of small ones, and large households typically list many more work-related goods in their inventories.³⁴ Given the strong correlation between estate wealth and household size, wealth bias in the probate record encapsulates the bias towards larger households. The same can be said about a potential third reason for the likelihood of an estate being inventoried: the complexity of its debts arrangements with outsiders. Again, statistical analysis shows that debts were very strongly correlated to the combined value of the material goods in the estate.³⁵ Therefore, it can be argued that an approach that successfully resolves the problem of wealth bias in the probate record will, to a large degree, also resolve issues of other potential biases in the data.

II

Historians have approached the wealth bias problem in several ways. A first type of approach is the one taken by Overton *et al*, who have simply accepted wealth bias as a *fait accompli*, perceiving themselves 'forced to define the statistical population that we study from the sample that we have'.³⁶ This means that it is explicitly admitted that all conclusions drawn from the probate analyses are not valid for society *in toto*, but only for the *inventoried section* of that society. There is of course nothing methodologically wrong with this approach. However, if not society, what is it exactly that is being analysed here? The 'inventoried section' is a rather elusive collection of contemporary households. It is often roughly equated with the 'middling sorts'.³⁷ However, as Overton *et al* acknowledge, that is not an accurate assessment of the population defined by the probate record, which covers a much wider range of people than the term 'middling sorts' can meaningfully imply, however vague and flexible that term is in itself. After all, the poor may be underrepresented in the probate record, but

³³ For the probate dataset from the Diocese of Chester used in this paper, the Pearson coefficient is .660 ($p < .001$), as calculated for log-transformed inventory wealth. The log transform was necessary since inventory wealth is not normally distributed (so significance and size of correlation cannot accurately be judged), whereas its logarithm is.

³⁴ Particularly because large households were often by-employed, with different members working in different occupations, as discussed in 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>, pp. 26-30.

³⁵ For the probate dataset from the Diocese of Chester used in this paper, the Pearson coefficient is .812 ($p < .001$), as calculated for log-transformed total of material goods and total debts. The log transform was necessary since material inventory wealth and debts are not normally distributed (so significance and size of correlation cannot accurately be judged), whereas their logarithms are.

³⁶ Overton *et al*, *Production and Consumption*, p. 29.

³⁷ For example, by Maxine Berg in Berg, *Luxury and Pleasure in Eighteenth-Century Britain* (Oxford: Oxford University Press, 2005); and by Peter Earle in Earle, *The Making of the English Middle Class: Business, Society and Family Life in London, 1660-1730* (London: Methuen, 1991).

they are not absent, and neither are the gentry.³⁸ So, not only does such an approach limit the validity of the analytical conclusion to merely a sub-section of society, it is not actually possible to equal that sub-section with a recognisable historical ‘entity’.

A second approach has been to use inventories mainly for trend analyses. Daniel Smith, for example, has argued that ‘in general, probate records are a much better source for the analysis of change over time within a small area than for the study of differences between regions and classes.’³⁹ The argument is that in this way, one set of households is compared to another set of essentially the same composition, thus providing a fair degree of insulation from the wealth bias problem. But, as Smith realized, such an approach is only valid if the inventoried share of society remained similar in size and composition over long time intervals. However, the share of households that were inventoried actually changed strongly over time, rising from very low levels during the late sixteenth century to a maximum in the final decades of the seventeenth century, and then declining and finally disappearing altogether during the eighteenth century. Secondly, Smith’s approach implicitly supposes that the trends observed in the inventoried households are reflected in very similar trends in non-inventoried households. Indeed, without this assumption, there would be little reason to choose this approach over the ‘sample defines the population’ approach discussed above. Such an assumption is not only untestable, as that would require information on the non-inventoried section of society, but it is also questionable, given the significant differences in wealth, occupational composition and social status between inventoried and non-inventoried households.

A third approach has been to supplement probate inventories with inventory collections specifically covering the poor, as in Peter King’s work on English pauper inventories.⁴⁰ Unfortunately, such inventory collections are very rare and generally quite small, limiting their application and statistical power; King’s dataset, for example, consisted of a mere fifty-one inventories. Furthermore, it is far from clear how to combine results from these ‘pauper’ datasets with those from probate inventories. Although such datasets allow one to peek at a section of early modern society that is difficult to observe through probate inventories, an integral view of society remains an elusive prospect.

A fourth approach has been to try to complete the picture of historical society by combining known analytical results for inventoried households with, for non-inventoried households, the ‘time honored [method] of the educated guess’.⁴¹ For example, when attempting to discover the wealth distribution of American colonial society, Alice Hanson Jones assumed that non-probated individuals were x per cent as wealthy as those that were probated, with x dependent on the inventoried share of the population. If that share was high, such as in the Middle Colonies, where seventy-one per cent of wealth holders left an inventory, she assumed x to be twenty-five per cent; if it was low, as in New England, where only one in three wealth holders was inventoried at death, she assumed x to be fifty per cent.⁴² Such an approach obviously suffers from the arbitrariness of the choice of x . Furthermore, it is not at all clear that the average wealth difference between probated and non-probated households would be related in this way to the inventoried share of the population. It is possible that the comparatively high share of probated households in the Middle Colonies was caused by the fact that

³⁸ Overton *et al*, *Production and Consumption*, p. 26.

³⁹ Smith, ‘Underregistration’, p. 106.

⁴⁰ King, ‘Pauper Inventories and the Material Lives of the Poor in the Eighteenth and Early Nineteenth Centuries’ in Hichcock, King, and Sharpe (eds), *Chronicling Poverty: The Voices and Strategies of the English Poor, 1640-1840* (Basingstoke: MacMillan, 1997), pp. 155-91.

⁴¹ Shammass, ‘Constructing a Wealth Distribution from Probate Records’, *The Journal of Interdisciplinary History*, 9:2 (1978), p. 298.

⁴² Jones, ‘Wealth Estimates’, pp. 116-17.

all but the very poorest were inventoried, thus lending credence to the low value of x assumed by Jones. But, the high inventory coverage could also be taken to suggest that many poor households must have been inventoried too, depressing the average wealth of inventoried households, leading to a *high* value of x . The point here is not which interpretation is correct, but that it is not possible to choose between them.

A fifth approach also relies on access to rare historical data, namely a non-probate-based source on the wealth distribution within society. Smith, in his study on early modern Hingham, Massachusetts, possessed a detailed tax list of inhabitants which could be matched to the probate record, thus allowing him to estimate the average wealth of non-inventoried relative to inventoried households⁴³ In effect, this removes the need to rely on an ‘educated guess’ for the x in the previous method; for Hingham, Smith could calculate x to have been 32.8 per cent. This approach only works when studying a small community, like Hingham, where it is feasible to match individuals between the several historical records. And even when limiting themselves to local studies, few historians are in the fortunate circumstance of having tax lists, probate records and death registers of sufficient detail to allow such a calculation to be made, as Smith readily acknowledged.⁴⁴ Furthermore, non-probate sources of wealth distribution data typically suffer from social bias themselves; early modern tax lists usually omit poor households and those with little or no real property as they were exempt from paying taxes. Also, the basis for determining the wealth of individuals in such sources is often so different from that in probate inventories that a meaningful ‘match’ is impossible, as Jones found in an experiment for Philadelphia County.⁴⁵ Tax lists, for example, are often based on real estate values which, in the Anglo-American case, are usually not included in probated wealth. And for Britain, the entire approach is unfeasible, as Lindert argued, because property and income taxes were partitioned into ‘unlinkable schedules’ or levied on occupiers rather than owners.⁴⁶

A sixth approach is based on correcting the probate record for age bias, assuming that such a correction will also remove wealth bias. The underlying argument for this assumption is that since probate inventories were taken at the end of someone’s life, probated individuals were on average older and therefore likely to be wealthier than the average living person, having had a life time to accumulate that wealth. By correcting for age bias, that is, by reconstructing the society of the living from the records of the recently deceased, this age-related wealth bias would be removed. Since probate records do not themselves provide information on the age of the deceased, this approach, like the previous one, relies on detailed information on individuals, allowing one to link individual probate records to data on the decedent’s age at death. Like the previous approach therefore, it is only really applicable in local studies, and even then only in the happy circumstances that the required detailed information on local individuals is available. Yet, one might extrapolate from such local studies to regional or even national probate records, if one can find evidence to support the assumption that the effects of the local age correction are likely to be representative for that larger geographical area. The real problem with this approach lies in the assumption that age correction will, as a fortunate by-product, result in a practically complete wealth correction too. Whereas it may well result in *some* share of wealth bias being removed, it is unclear *what* share.⁴⁷ Main and Jackson have argued that it is a large share. Based on comparisons between tax lists and probate inventories, they wrote that ‘as a

⁴³ Smith, ‘Underregistration’, p. 106.

⁴⁴ Ibid, p. 110.

⁴⁵ Jones, ‘Wealth Estimates’, p. 118, referencing here a paper titled ‘Wealth Distribution in the American Middle Colonies in the Third Quarter of the Eighteenth Century’, presented at the annual meeting of the Organization of American Historians, New Orleans (April, 1971), pp. 31-34.

⁴⁶ Lindert, ‘Algorithm’, p. 661.

⁴⁷ Shammass, ‘Wealth Distribution’, p. 297.

result of these efforts, we feel reasonably confident that the only bias afflicting the records for most of the colonial period is the familiar and natural one of age', yet provided no evidence to back this claim up.⁴⁸ Smith, using the exceptionally good Hingham records, was able to carefully test the assumption. He found that 'age per se had little to do with a man's leaving a will or having an estate inventoried. The pronounced differential [between the probated and non-probated] arises not from the age of the decedents, but from their wealth.'⁴⁹ Furthermore, if Overton *et al*'s study of the English parish of Milton is correct, age bias in the probate record was slight, meaning that an age correction would remove only a sliver of the inventories' wealth bias.⁵⁰

A seventh and more promising approach has been to divide both the collection of probate inventories and contemporary society into several sections which may be expected to have differed in average wealth. These sections can then be 'reweighted' within the collection of probate inventories, using their share of *all* households as the 'weight'. Such an approach has been proposed by Lindert, using occupational groups as the sections, and by Carole Shammas, using eight sections based on a combination of occupational status and age.⁵¹ However, as Lindert realized, the average wealth of, say, probated artisans is not actually representative of the wealth of *all* artisans, as the former were a relatively wealthy subset of the latter. In other words, this reweighting approach will remove the effects of the over- and underrepresentation of entire sections of society in the probate record, but not those of wealth bias *within* these sections.⁵² Or, to phrase this in the terms of the by-employment analyses, the same cause that made farmers or tanners likely to leave an inventory, namely the possession of livestock or expensive stocks of raw materials, would have made manufacturers who were by-employed in farming, or farmers who also worked as tanners more likely to leave an inventory than their non-by-employed colleagues. This means that the by-employed will be overrepresented within each of the occupational sections in the inventory record. A corrected inventory set, reweighting occupational sections in their entirety, would thus provide only a limited improvement over the 'raw' set, as the main cause of overrepresentation of the by-employed in the probate record lay *within* the occupational sections.

In short then, none of the above methodologies can adequately correct for the probate record's inbuilt wealth bias. But Lindert and Shammas's method does provide a starting point for an approach which can, as will be discussed in the next section.

⁴⁸ Main and Jackson, 'Economic Growth and the Standard of Living in Southern New England, 1640-1774', *The Journal of Economic History*, 48:1 (1988), p. 125.

⁴⁹ Smith, 'Underregistration', p. 105.

⁵⁰ Overton *et al*, *Production and Consumption*, p. pp. 27-28, 208. See also the discussion in this dissertation, pp. 13-4.

⁵¹ Lindert, 'Algorithm', pp. 662-63; Shammas, 'Wealth Distribution', pp. 298ff.

⁵² Lindert, 'Algorithm', p. 664.

III

This new approach can best be explained by first examining the historical process which led to the current probate record. The process has been sketched in illustration 1, which should be read from top to bottom.

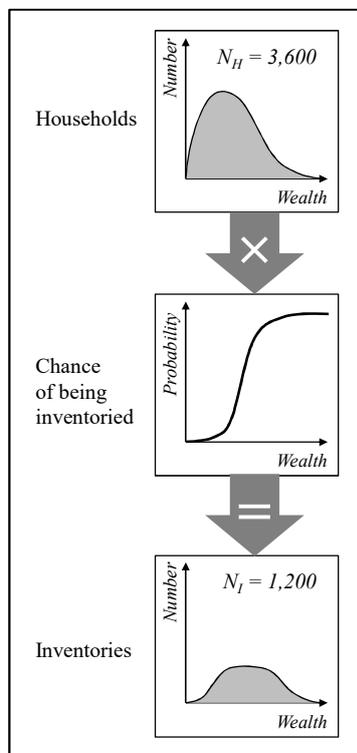


Illustration 1. Creation of the probate record

Suppose an (imaginary) historical ‘society’ which consisted of N_h households which were distributed as a function of their wealth as depicted at the top. Suppose, further, that the probability that a householder of a given wealth was inventoried was dictated by the probability function depicted in the middle – with poor household having a low and wealthy household a high chance of being inventoried. Then, the wealth distribution of the probate record, depicted at the bottom, is nothing more than the result of a multiplication of the two curves above it. Since the probability of leaving an inventory positively depended on wealth, the process ‘moved’ the original, household wealth distribution to the right (higher wealth), explaining why the household and inventory wealth distributions differ. The more the original household wealth distribution was skewed towards the left (low wealth), the fewer households left an inventory. The ratio between the number of households (N_h) and the number of inventories (N_i) thus depends both on the household wealth distribution and on the form of the probability function.⁵³ In the hypothetical example depicted here, on average one in three householders was probated.

In illustration 2, the (fictional) population of households from illustration 1 has been divided into three occupational subsets: farmers, manufacturers and labourers. The probability function was only dependent on the wealth of the individual householder, not on his occupation. But the household wealth distribution differed per occupation and so did, therefore, the average share of households that left an inventory. Relatively more farmers were inventoried since wealthy householders were relatively more numerous amongst them – the farmers’ wealth distribution was skewed to the right. The opposite was the case for labourers. As a result, in the hypothetical example here, two in every three farmers’ households left an inventory, compared to in one in every four manufacturers and one in every twelve labourers. This explains why farmers are overrepresented in the probate record compared to manufacturers, and even more so compared to labourers.

⁵³ All this can perhaps be more clearly expressed in a mathematical fashion. The probability that deceased householder j with a wealth W_j was inventoried was a function of wealth $P(W_j)$. If the population of households numbers is N_h and the number of inventories is N_i then the relationship between those two numbers can be expressed straightforwardly as follows:

$$N_i = \sum_{j=1}^{N_h} P(W_j).$$

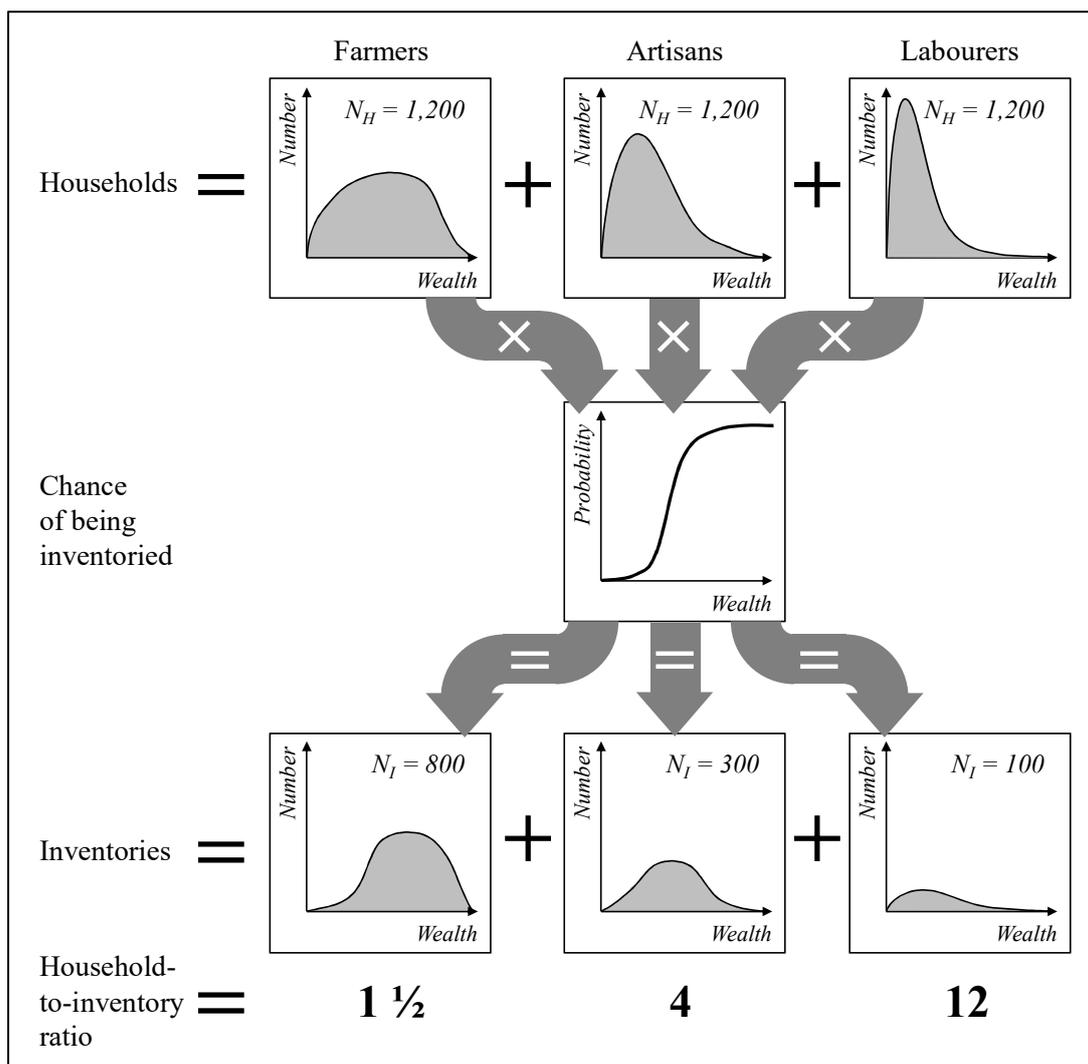


Illustration 2. An explanation of the variation in household-to-inventory ratios in occupational groups with different wealth profiles

Of course, both the household's wealth distribution and the probate probability function are, in practice, unknown. But the wealth distribution of the inventories is *knowable* – it can be accurately uncovered by analysing a sufficiently large and representative sample amongst all extant inventories. If we also knew the probate probability function, the above-described historical process could be reversed, and the original household wealth distribution could be uncovered as well.

Actually, the probate probability function is not entirely unknown. Certain logical constraints can be imposed upon it, as depicted in illustration 3. The chance of leaving an inventory must have been negligible for households of near-zero wealth. Also, the wealthier a household was, the higher would have been the probability of being probated, so the probate probability function must have increased unceasingly and without peaks – in more technical terms, it must have been monotonically increasing and non-modal. It would also have been a continuous function, that is, it is highly unlikely that there would be household wealth levels at which the probability being inventoried suddenly,

discontinuously ‘jumped’ upwards.⁵⁴ And it would have been asymptotic in shape, as above a certain wealth level, chances of being inventoried would only increase marginally with even more wealth. It is possible that the probability of being inventoried also increased only marginally with wealth at the other end of the spectrum, for households with very low wealth. This would have led to the curvature of the probability function changing from positive to negative at a wealth level somewhere in between households with very low and very high wealth – but there is no reason to expect that the function could have had more than one such an inflection point. It can therefore be concluded that the probability function must have been either straightforwardly asymptotic or S-curve shaped.

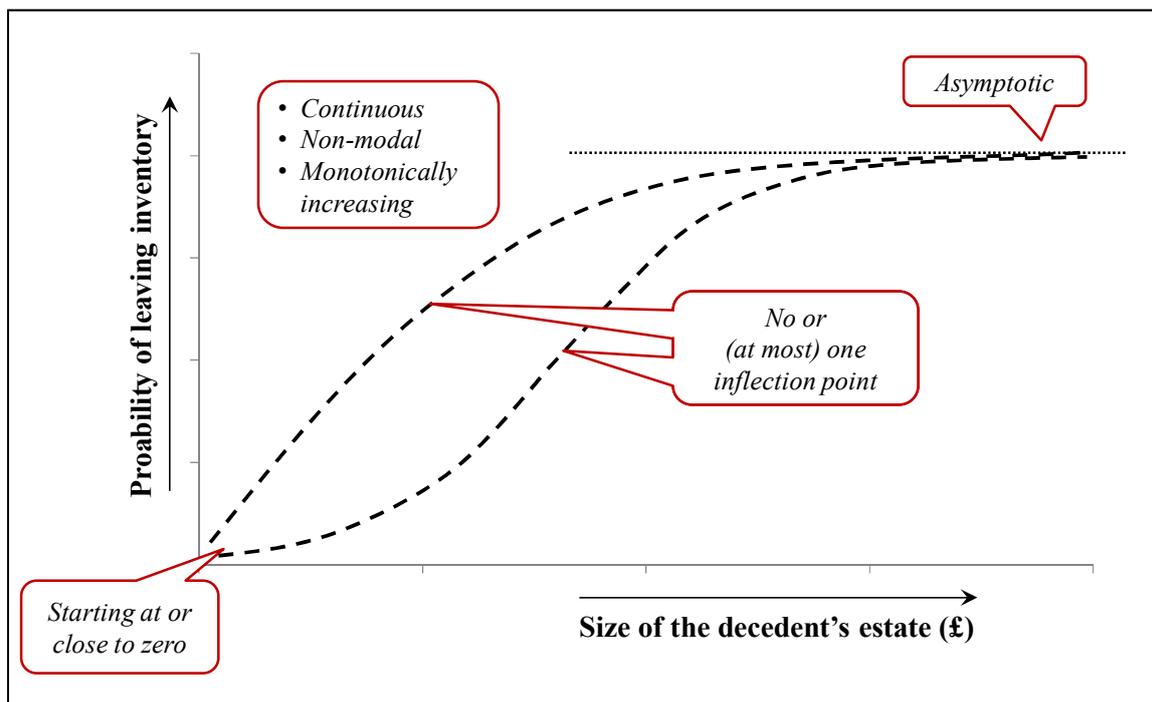


Illustration 3. Constraints on the shape of the potential probability function

The household-to-inventory ratios in illustration 2 can be calculated by comparing that occupation's share within the occupational structure with its share within the inventory record. Figure 2 presents household-to-inventory ratios for the Diocese of Chester in the early eighteenth century. The direct probate jurisdiction of the Diocese of Chester covered the ancient counties of Cheshire and Lancashire, south of the river Ribble, corresponding to what Stobart has called the world's 'first industrial region'.⁵⁵

⁵⁴ The most probable candidate for a potential discontinuity would have been the infamous £5 'threshold' in English probate inventories but, as Cox and Cox report, this threshold did not actually exist. See: Cox and Cox, 'Probate, 1500-1800: A System in Transition', in Arkell *et al* (eds), *When Death Do Us Part*, p. 26, fn. 65.

⁵⁵ Stobart, *The First Industrial Region : North-West England, C.1700-60* (Manchester: Manchester University Press, 2004).

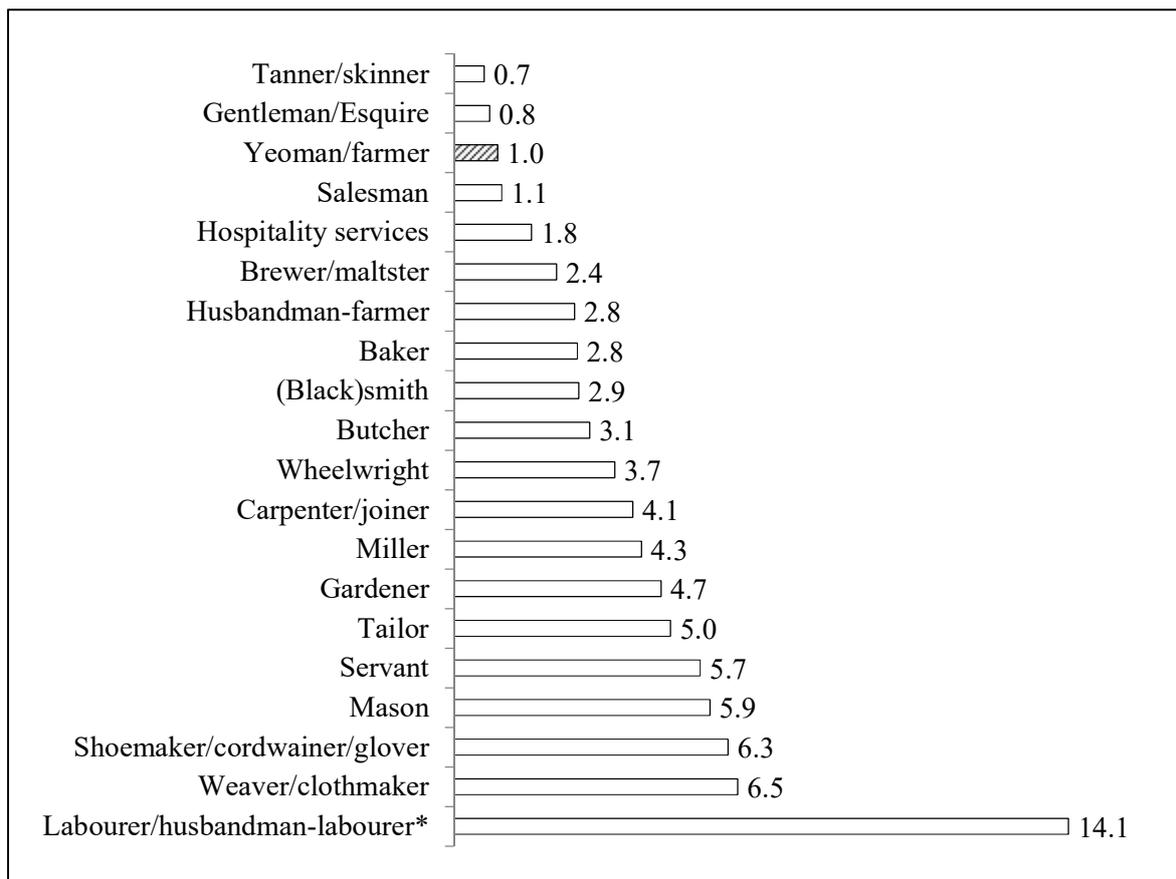


Figure 2. Relative household-to-inventory ratios; Chester Diocese, c.1725
(relative to yeoman/farmer = 100%)

Note: *A comparison between burial registers and probate data shows that one in seven decedents called 'husbandman' in their probate documents was called 'labourer' in the burial registers. Multivariate regression analyses, used as a means to allocate labourers to occupational sectors, also shows that 15% of all husbandmen in Cheshire and Lancashire were not farmers but labourers. The figures in this and following charts take this into account.

General remark: the probate data were restricted to the same set of parishes for which occupational information was available, to ensure an optimal one-to-one match.

Sources: probate database for Chester Diocese (i.e. Cheshire and Lancashire south of the river Ribble); parish records for the same geographic area, c.1725, collected by the Cambridge Group.

Armed with reliable information on the inventory wealth distribution, the (relative) household-to-inventory ratios for different occupations, and the approximate shape of the probability function, it is now possible to attempt to actually determine that function. In illustration 4, the historical process depicted in illustrations 1 and 2 has been reversed. Reading from bottom to top, the number of inventories per sector is used to calculate the number of households per sector as implied by the probability function. The probability function is determined via an iterative four-step approach, depicted in red, and described in more detail below.

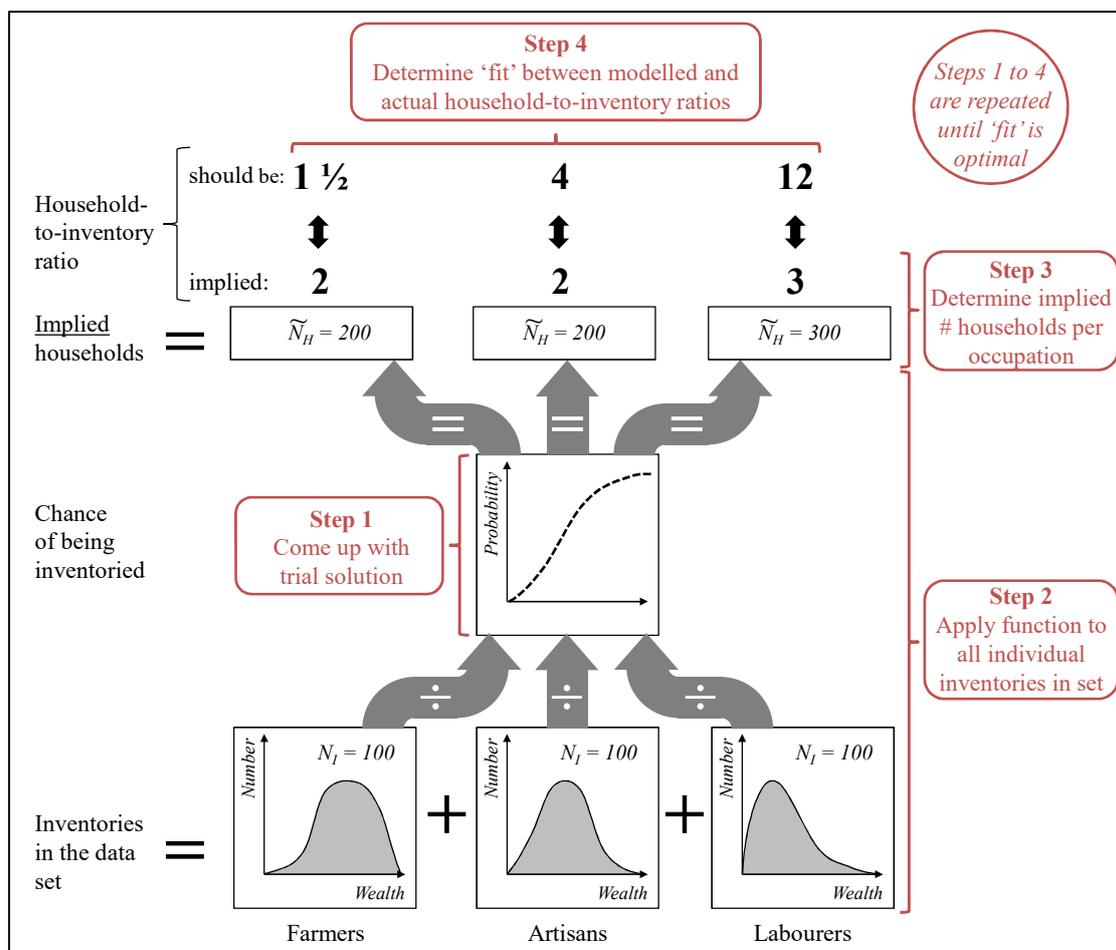


Illustration 4. A conceptual representation of the iterative process followed to recover the historical relationship between the deceased's wealth and the probability of his/her estate being inventoried

Step 1 consists of designing a trial version for the probability function. In step 2, this trial function is applied to every single inventory in the dataset to calculate the number of households that, on average, would have together left that one, single inventory *if the trial function were correct*. For example, suppose that inventory x in the dataset has an estate value of £10. Suppose furthermore that the trial function states that household heads which owned £10 in goods had a twenty per cent chance of being inventoried at death. The trial function then implies that inventory x corresponds statistically with five historical households. In step 3, these implied household numbers per individual inventory are combined per occupation, leading to implied household-to-inventory ratios per occupation. These are then, in step 4, compared to the *actual* ratios for these occupations, as determined from the occupational structure as presented in figure 2. The better these two sets of ratios compare, the more the trial function resembles the actual, historical probability function. The approach now returns to step 1, where a new version of the trial function is designed which, hopefully, improves the 'fit' between model and reality. This new trial function forms the basis of a new iteration, and these iterations are repeated until the 'fit' can no longer be improved. In practice, of course, the four steps are automated using a fairly standard iterative computer algorithm, such as the generalized reduced

gradient optimisation process which was used for the analyses in this research.⁵⁶ The end result of the iterative approach is the best possible approximation of the actual, historical probability function. Running this approach on the Chester Diocese inventory set results in a very good match between model and historical reality, as figure 3 demonstrates.

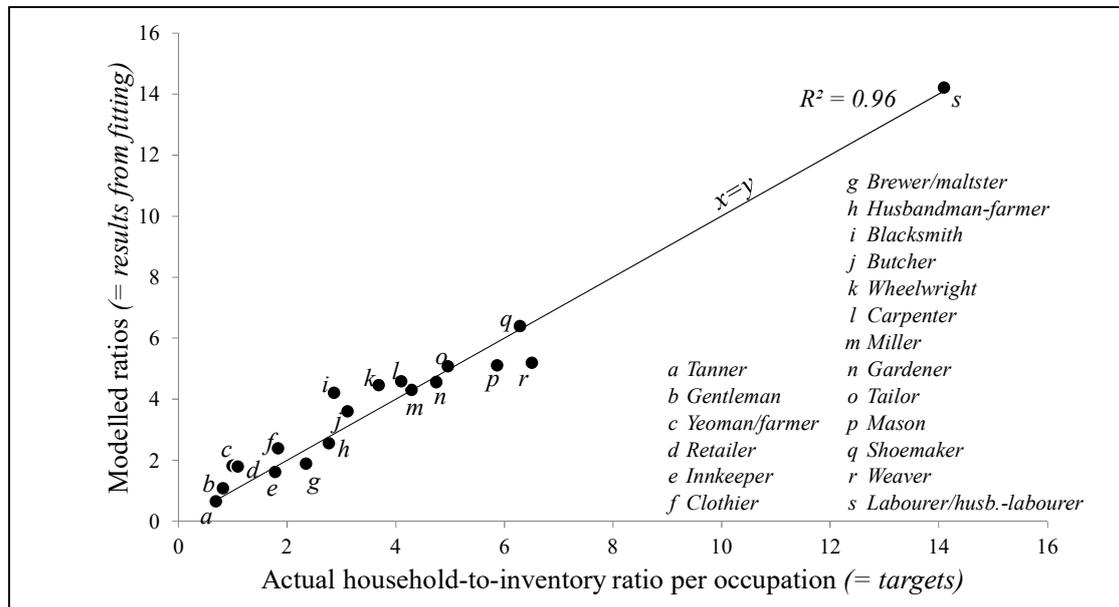


Figure 3. Household-to-inventory ratios – a comparison between actual figures (horizontal axis) and those resulting from the iteratively determined probability function (vertical axis); Chester Diocese, c.1725

The probability function thus determined for Chester Diocese in the early eighteenth century is depicted in figure 4.

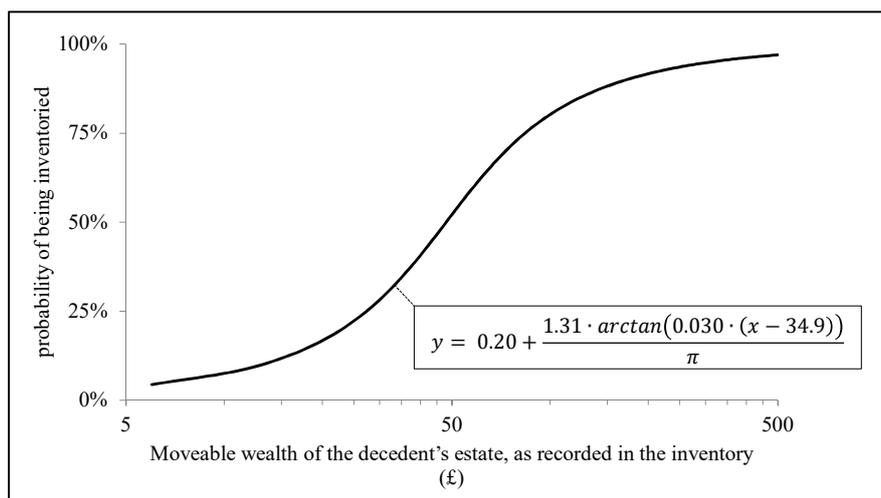


Figure 4. The probability of leaving an inventory as a function of wealth (Chester Diocese, c.1725)

⁵⁶ Lasdon *et al*, 'Solving the Pooling Problem Using Generalized Reduced Gradient and Successive Linear Programming Algorithms', *SIGMAP Bull.*:27 (1979), pp. 9-15.

Armed with the probability function from figure 4, the historical process, as depicted in illustration 1, can now be reversed, as has been done in figure 5 for all agricultural inventories in Chester Diocese. Unsurprisingly, the household wealth distribution is, on average, considerably poorer than the wealth-biased inventory wealth distribution, containing a much larger share of inventories at the lower end of the wealth spectrum, and a much smaller share at the upper end.

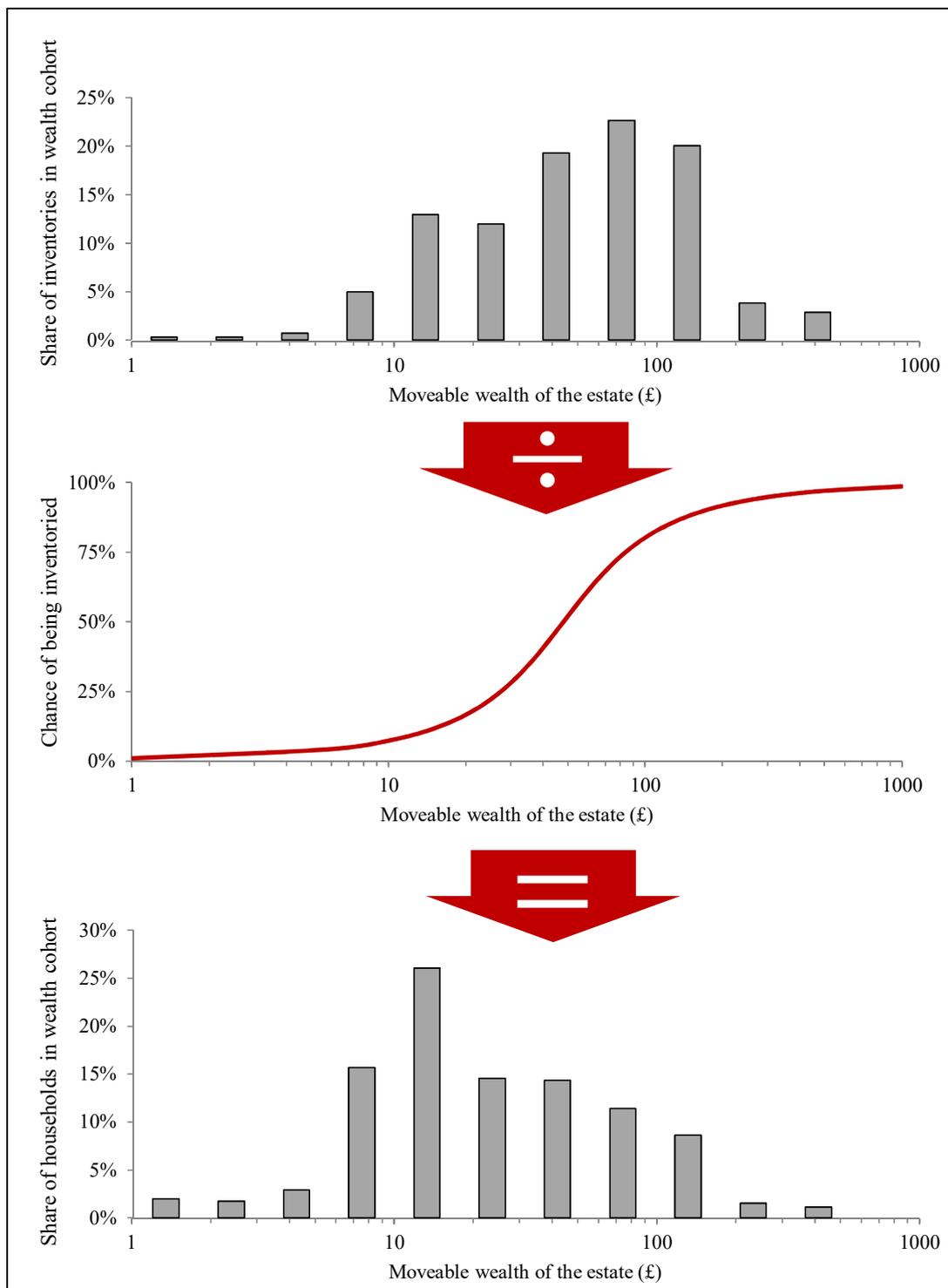


Figure 5. The reconstruction of the population of agricultural households (yeomen, husbandmen, labourers) from the probate data (Chester Diocese, c.1725)

IV

As discussed in the introduction to this paper, probate inventories are a key source of statistical, quantitative information, underpinning a number of important historical discussions. But their wealth bias is a severe handicap in most of these discussions. For example, the presumed ubiquity of by-employments in early modern Britain is almost entirely based on evidence from probate inventories. But for the same reason that farmers and brewers are overrepresented in the probate record, namely their capital-intensity and, therefore, their relatively high-value estates, so are households with a by-employment in farming or brewing. The degree to which inventories exaggerate by-employments can be demonstrated by comparing by-employment incidence in the raw probate evidence with that in the wealth-bias-corrected probate data. As figure 6 demonstrates, this exaggeration is substantial. Rather than a clear majority, less than one in three secondary-sector households was actually by-employed in agriculture. And rather than one in seven, only one in fourteen such households were by-employed in an additional manufacturing activity. I have explored the effects and implications of a wealth-bias correction on early modern households in much greater breadth and detail in a separate working paper.⁵⁷

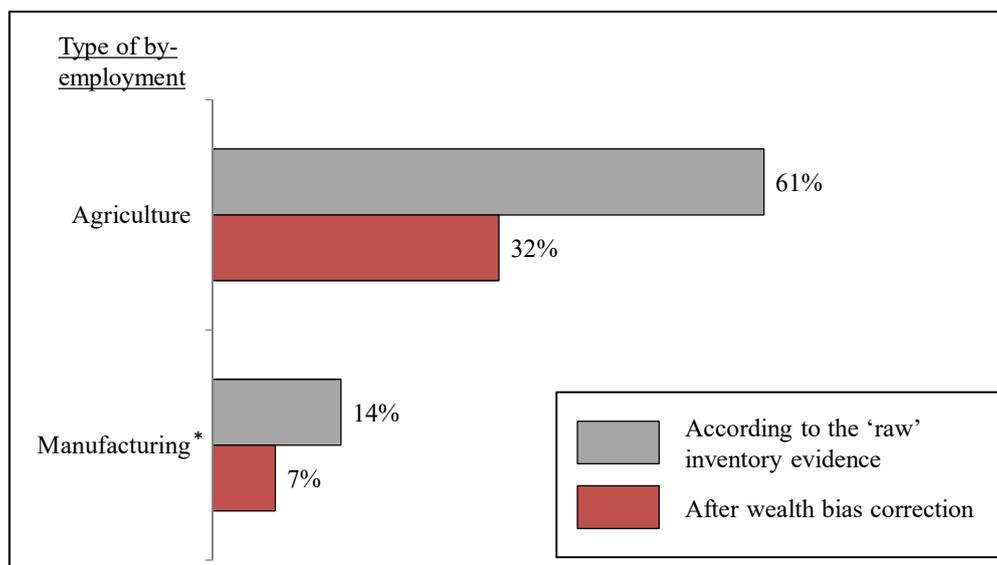


Figure 6. By-employment incidence amongst secondary-sector households before and after wealth-bias correction (Chester Diocese, c.1725)

Note: *excluding spinning.

Sources and calculations: 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>.

⁵⁷ 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>.

Probate inventories also play a central role in the discussions surrounding the rise of consumption and developments in material culture in the early modern world. The consumer revolution thesis is, to a substantial degree, built on the evidence of the increasing presence of new consumer goods in probate inventories. The consumption component of Jan De Vries's industrious revolution concept similarly depends on it, and not merely in the inventories of the middling sorts but also amongst the poorer households; it is the desirability of the new consumer goods for these households which drives the rising industriousness from which the concept gets its name. However, these poorer households are severely underrepresented in the probate data which, therefore, is likely to significantly exaggerate the penetration of the new consumer goods in early modern society as a whole. Correcting the probate evidence for wealth bias exposes the degree of exaggeration. As the example in figure 7 shows, the 'raw' probate evidence seriously inflates ownership of clocks, particularly amongst poorer households; rather than one in five, merely one in twenty labourer's households is estimated to have possessed them in early-eighteenth-century Cheshire and Lancashire.

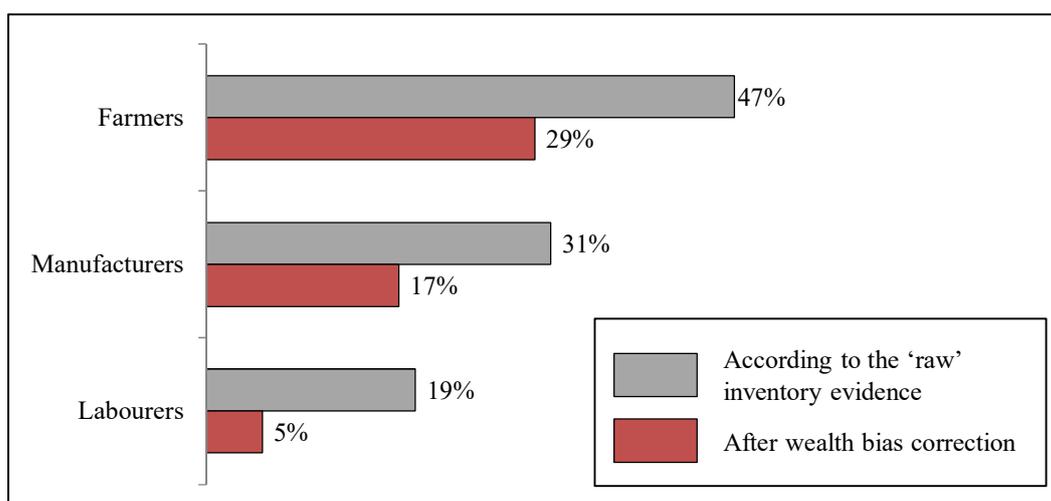


Figure 7. Clock ownership as indicated by probate inventories before and after correction for wealth bias (Chester Diocese, c.1725)

Comparing the agricultural assets suggested by the probate evidence with that from independent and unbiased sources provides a final example of the need (and validity) of a wealth bias correction. As shown in a recent article, co-authored with Shaw-Taylor, the probate evidence, taken at face value, strongly exaggerate average livestock numbers per household and, thereby, the total number of livestock in a given area.⁵⁸ As shown in figure 8, the number of cattle in Cheshire and Lancashire in the mid eighteenth century as suggested by the probate evidence greatly exceeds the actual numbers. But the numbers derived from the wealth-corrected probate evidence are in full agreement with the figures derived from independent, unbiased sources.

⁵⁸ Keibek and Shaw-Taylor, 'Early Modern Rural by-Employments: A Re-Examination of the Probate Inventory Evidence', *Agricultural History Review*, 61:2 (2013), pp. 274-7.

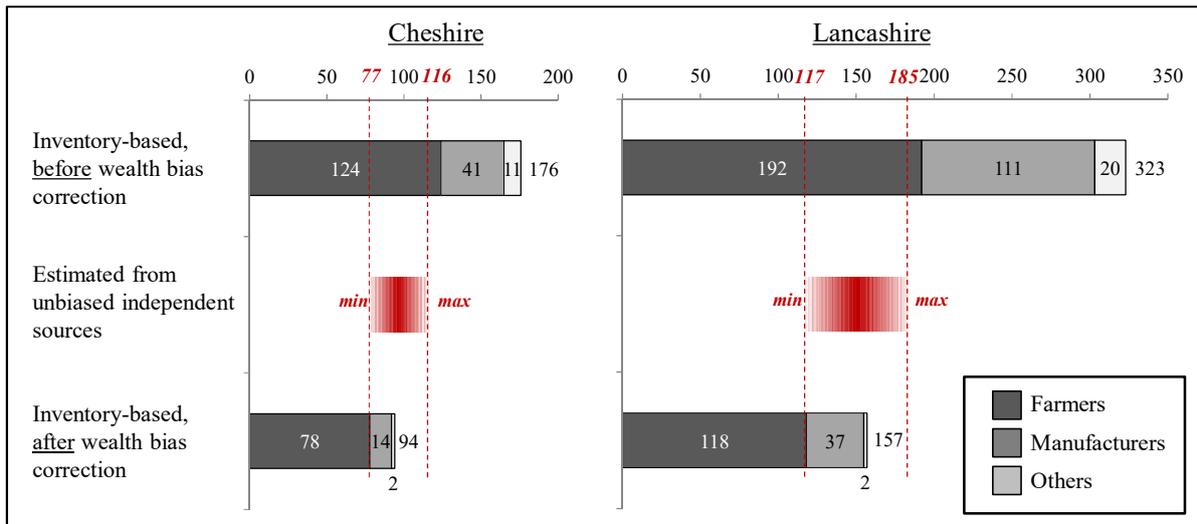


Figure 8. Total numbers of cattle in the counties of Cheshire and Lancashire, c.1760 as derived from probate data, compared to independent, unbiased estimates

Notes: For underlying assumptions and calculations of the independent and probate-based estimates, see Keibek and Shaw-Taylor, 'Rural by-employments', pp. 274-7.

Sources: Probate inventory dataset; agricultural censuses for 1871 and 1901; Wrigley, Early censuses; Wrigley and Schofield, Population history; Holland, Cheshire; Wedge, Palatine of Chester; Holt, Lancaster; Rothwell, Lancashire; 'Occupational Structure' project (parish record counts for Cheshire and Lancashire); Turner, 'Counting sheep; waking up to new estimates of livestock numbers in England c.1800', *AgHR* 46:2 (1998), pp. 142-61.

As the examples above show, wealth bias correction of the probate data provides a powerful means of overcoming the critical defect of this important historical data source, and to re-examine the many historical analyses, conclusions, and theories based on the uncorrected data. The examples in this paper are limited to the Cheshire and Lancashire, but parish register data that can serve to locally calibrate the probate record are available for much of pre-industrial England and Wales, going back to the final decades of the seventeenth century. As I have shown elsewhere, household-to-probate ratios are, for most occupations, fairly constant over time and place, so it may actually be possible to push the wealth-correction methodology further back in time, and to areas in Britain for which no calibration data exist.⁵⁹ More work is required to substantiate this possibility, as well as to examine the methodology's potential in other early-modern countries.

⁵⁹ Keibek, 'Using probate data for estimating historical male occupational structures', *Campop paper*, <http://www.campop.geog.cam.ac.uk/research/projects/occupations/abstracts/paper27.pdf>, pp. 13-7.