CAMBRIDGE WORKING PAPERS in ECONOMIC and SOCIAL HISTORY

Working Paper No. 27 – MARCH 2017:

ALLOCATING LABOURERS TO OCCUPATIONAL (SUB-)SECTORS USING REGRESSION TECHNIQUES

Sebastian A.J. Keibek

Cambridge Group for the History of Population and

Social Structure & Queens' College

sk571@cam.ac.uk

ABSTRACT

British historical sources on occupational information such as censuses, parish registers, and probate records describe many men with the unhelpfully vague term of ‘labourer’. This paper introduces a new method to allocate these labourers to occupational (sub-)sectors, a prerequisite for creating comprehensive and accurate historical occupational structures. The new method leads to a significant correction on the allocation shares used in the national accounts literature. Its results at national level are largely in agreement with another new approach, developed by Osamu Saito and Leigh Shaw-Taylor. But it has an important advantage over that approach: it is capable of allocating labourers at all geographic levels, and can thus generate local and regional occupational estimates, rather than only national ones.

INTRODUCTION

When attempting to establish male occupational structures of pre- and early-industrial Britain, labourers present a problem. The term labourer is a very common occupational denominator; thirty per cent of the men in the Cambridge Group's quasi census of c.1817 are described as such in the baptism registers on which the quasi census is based.[1] But the term labourer is a vague one. It tells us that these men worked for an employer and were probably mainly remunerated by money wages but, on its own, it does not tell us what kind of work they did and to what occupational sector they should be attributed. In the Victorian and later censuses, the problem is manageable, as most labourers were, in fact, allocated to an occupational sector. For example, one in five working men in the 1851 census were labourers, but seventy-four per cent of these were explicitly allocated to an occupational sector, predominantly to agriculture. The problem of how to apportion these labourers to sectors is therefore limited to five per cent of the male population, for which it is furthermore clear that they were not farm labourers. But for the c.1817 quasi censuses and for earlier estimates of occupational structure, whether based on parish register, probate records, or social tables, no such contemporary allocation is available. For these estimates, therefore, historians have had to devise ways for allocating labourers to sectors themselves.

In this paper, a new method for apportioning labourers to sector is presented, based on regression techniques. The results of this allocation for England and Wales are also discussed, on national and regional scales, for the eighteenth and early nineteenth centuries. But before discussing the regression approach, it is first necessary to discuss why a new method is required at all. The answer is straightforward: because existing solutions to the labourer problem are not satisfactory. Nick Crafts, in essence, allocated all labourers to agriculture, although he also performed a sensitivity analysis, using the ratio between agricultural and non-agricultural labourers in the 1831 census to calculate a lower bound for the agricultural sector.[2] For their 1381 and 1522 estimates, Stephen Broadberry et al allocate labourers to sectors in line with the agricultural and non-agricultural shares of non-labourers, assigning all non-agricultural labourers to the secondary sector. For their 1688, 1759, and 1801 estimates, they applied the 1522 allocation percentages, apportioning 68.2% per cent of labourers to agriculture and the remainder to the secondary sector.[3] They are thus making two unlikely assumptions. Firstly, that labourers were divided across agriculture and the secondary sector in line with the agricultural/non-agricultural split for other workers, leading to the improbable result of the ratio between labourers and non-labourer being higher in the secondary than in the primary sector. Secondly, that the 1522 allocation key applies to a much more advanced economy with a very different occupational make-up nearly three centuries later. Another, lesser issue with this allocation method is that all non-agricultural labourers are presumed to have been working in the secondary

[1] Kitson *et al*, 'The creation of a "census" of adult male employment for England and Wales for 1817' (Cambridge, 2012), http://www.econsoc.hist.cam.ac.uk/docs/CWPESH%20number%204%20March%202012.pdf

[2] Crafts, *British economic growth during the Industrial Revolution* (Oxford: Clarendon, 1985), p. 14.

[3] Broadberry *et al*, *British economic growth, 1270-1870* (Cambridge: Cambridge University Press, 2015), pp. 346-7, 349-50, 352-4, 356-360. The 68.2 and 31.8 percentages for agricultural and non-agricultural labourers are actually based on the division of non-labourers for 1522 as calculated in an earlier journal paper, namely Broadberry, Campbell, and van Leeuwen, 'When did Britain industrialise? The sectoral distribution of the labour force and labour productivity in Britain, 1381–1851', *Explorations in Economic History*, 50:1 (2013), pp. 17-21. However, this was a miscalculation, which was corrected in the more recent book. Nevertheless, the allocation percentages for later time periods remain to be based on the (it would seem miscalculated) division in the earlier paper.

sector, whilst some of them are in fact likely to have been engaged in tertiary sector activities such as dock work.

A new approach to the labourer problem was recently developed by Osamu Saito and Shaw-Taylor.[4] It is still work in progress and remains unpublished, as the authors do not yet consider it satisfactory, but it represents a major improvement on earlier approaches. A version of the approach was used for Shaw-Taylor and Sir E.A. Wrigley's national estimates in the recent Cambridge Economic History of Modern Britain (CEHMB).[5] It utilises the fact that labourers were allocated with varying but generally increasingly precision to occupational groups in the censuses between 1851 and 1911. For these occupational groups, ratios between labourers and non-labourers can therefore be calculated. Such ratios from the late nineteenth and early twentieth century census were used to allocate the not-yet-specified labourers in the 1851 census more accurately to sectors. The 1851 ratios between labourers and non-labourers in the secondary and tertiary sector were subsequently used to estimate labourer-to-non-labourer ratios in earlier time periods. For c.1817, the 1851 ratio was used directly; for c.1710, two estimates were produced, one based on the 1851 ratio, and one on half that ratio. By applying these ratios to the number of non-labourers in the secondary and tertiary sectors, the number of labourers for each of these two sectors was estimated and subtracted from the total number of unspecified labourers, the remainder of whom were allocated to agriculture.

Saito and Shaw-Taylor note that using evidence from the mid nineteenth century to calculate labourer shares in c.1817 and c.1710 is 'far from ideal'. However, as they demonstrate, labourer-to-non-labourer ratios within the secondary and tertiary sectors were reasonably stable for the 1851-1911 censuses, suggesting that they likely did not vary a great deal between c.1817 and 1851 either. The c.1710 estimates are based on two sets of possible ratios since, as the authors remark, simply applying the 1851 ratios to the early eighteenth century 'would require a much greater leap of faith', particularly in light of growing proletarianisation over the intermediate years. The two ratios result in a difference of less than three percentage points in estimates of the primary sector share for c.1710.

This solution represents a significant improvement over previous approaches. As Saito and Shaw-Taylor readily admit, it is unfortunate that it has to rely on ratios derived from much later census data. Substantial employers of labourers such as the transport industry, in which most tertiary sector labourers worked, experienced fundamental changes over the c.1710-1851 period, for example moving from road to canal to rail transport, and this potentially changed employment opportunities for labourers substantially. The authors demonstrate satisfactorily, however, that the effects on the national estimates of the male occupational structure are likely modest. Another implicit drawback of the approach is that it only produces an indirect estimate for the primary sector, the largest employer of labourers.

Arguably, none of these issues is serious enough to necessitate developing yet another labourer allocation methodology. However, there is an additional problem with the Saito/Shaw-Taylor approach: as they themselves emphasize, it is only really suitable for allocating labourers on a national scale. One of the great advantages of occupational data over the national accounts approach in analysing and explaining economic developments is that the former can be used to generate estimates for particular regions or even for individual towns and villages. Regional economic

[4] The Saito and Shaw-Taylor paper is not yet publically available, but the authors kindly provided me with a draft version and with the underlying data.
[5] 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.

trajectories often differed greatly in pre-industrial Britain, and the process of regional specialisation within an integrating economy was a recognizable feature of its economy, and potentially a factor in its relatively precocious development. But estimates below the national level cannot be generated with the Saito/Shaw-Taylor methodology.

A new method was therefore developed, based on multivariate regression techniques. Contrary to the Saito/Shaw-Taylor solution, it only uses contemporary data and provides a direct, independent estimate of the primary sector. It is founded on a minimal number of straightforward assumptions, the validity and accuracy of which can be directly tested. Most importantly, it can be used to allocate labourers on local and regional as well as national scales.

PRINCIPLES OF THE NEW ALLOCATION METHOD

The approach developed in this paper is based on multivariate regression techniques, with the number of labourers (per geographical unit of analysis) as the dependent variable. The mathematical basis is straightforward: the total number of labourers (L) within a given geographic area equals the sum of the number of agricultural labourers (L_a) and non-agricultural labourers (L_{na}) in that area, the latter being composed of labourers working in mining and quarrying (L_m), in the secondary sector (L_s), and in the tertiary sector (L_t):

$$L = L_a + L_{na} = L_a + L_m + L_s + L_t \quad (1)$$

It should be noted here that a labourer's employment was not necessarily restricted to a single type of work. Labourers may have taken up varying types of work, depending on availability, perhaps working in agriculture during the peak seasons of the farming year whilst working for building tradesmen at other times. The number of labourers in a specific trade should therefore not be interpreted in terms of individuals allocated fully to that trade, but rather as the sum of fractions of individuals, allocated to that trade, with those fractions corresponding to the share of the average working year in which they provided wage labour to employers in the trade. A general labourer working eight months of the year on the land and four in building and construction is counted here as two-thirds of a farm labourers and one-third of a building labourer.

Returning to equation (1), one may logically presume that, for farms engaged in a similar type of farming, on similar terrain, and in a similar climate, there will have been a positive relationship between the average number of agricultural workers per farm and the mean farm surface. The shape of this relationship depends on potential economies of scale which larger farms may have enjoyed over smaller ones. The 1851 census provides the data to test the shape of the relationship directly, the results of which are displayed in Figure 1 for Bedfordshire and Rutland.

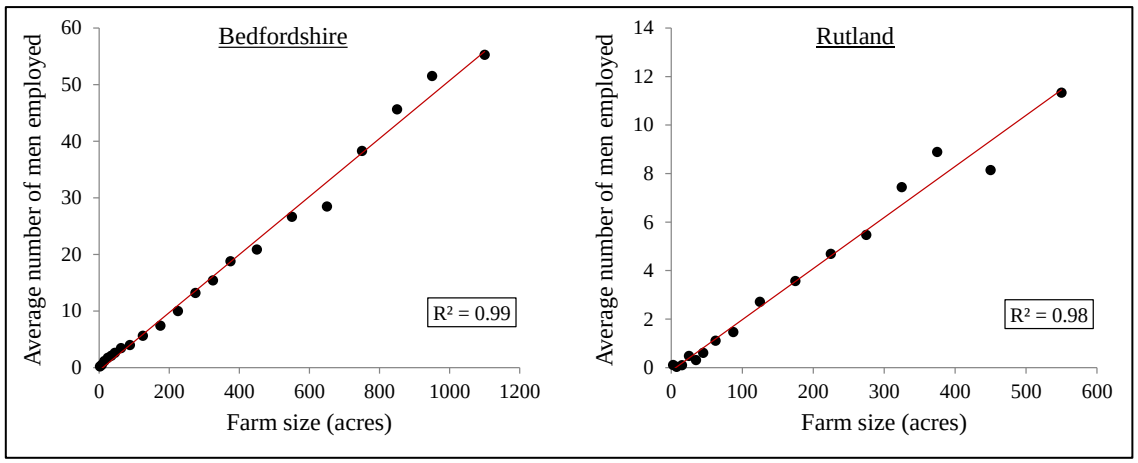


Figure 1. The average number of men employed per farm, as a function of farm size, 1851

Note: The data for these graphs were collected from the 1851 census by Leigh Shaw-Taylor for his research into the rise of agrarian capitalism, and kindly made available to me.
Sources: 1851 Census; Leigh Shaw-Taylor.

It is clear that economies of scale were minimal, and that the relationship was a simple linear one:[6]

$$\frac{L_a + NL_a}{F} = C_S \cdot \frac{S}{F} \qquad (2)$$

with L_a being farm labourers, NL_a being the other, non-labourer farm workers (farmers, family members, specialist workers, etcetera), F being the numbers of farms, S being the total surface in agricultural use, and with C_S a constant.

Equation (2) can only really be expected to hold for areas of homogeneous agricultural topography, such as within the boundaries of fairly uniform counties like Bedfordshire and Rutland. As Figure 1 shows, C_S differed considerably between these two counties, with one farm worker being added for every twenty acres in Bedfordshire, compared to one per forty-eight acres in Rutland. The value of C_S is a measure of the 'labour intensity' of agriculture, and depends on local factors such as soil quality, climate, type of agriculture, and the proximity to urban centres. For larger and more mixed counties, let alone for England and Wales *in toto*, a single constant cannot adequately describe the relationship between farm size and numbers of worker. Such heterogeneous areas have to be divided into smaller, (more) uniform areas, with equation (2) solved for each of these smaller areas separately. The smaller and the more uniform such areas are made, the more accurate the regression will be. Therefore, in this paper, the equation has always been solved at the lowest level of geographical detail for which the required data were available. For 1851, at the time these analyses were performed, this was the census registration district, as occupational information was not (yet) available for smaller geographical units.[7] Thus, for 1851, England and Wales were divided into 624 units. For 1813-20 and for earlier time periods, male occupational information derived from baptism registers was available at the much smaller level of (up to) 11,365 Anglican Registration Units, that is, the parishes and chapelries in

[6] The linearity of the relationship can also be tested directly, by trying out different curves in the regression analysis. This too shows that a linear relationship provides the best 'fit' between model and data.
[7] Via the I-CeM project, information for the 1851 census has recently become available at the level of parishes too.

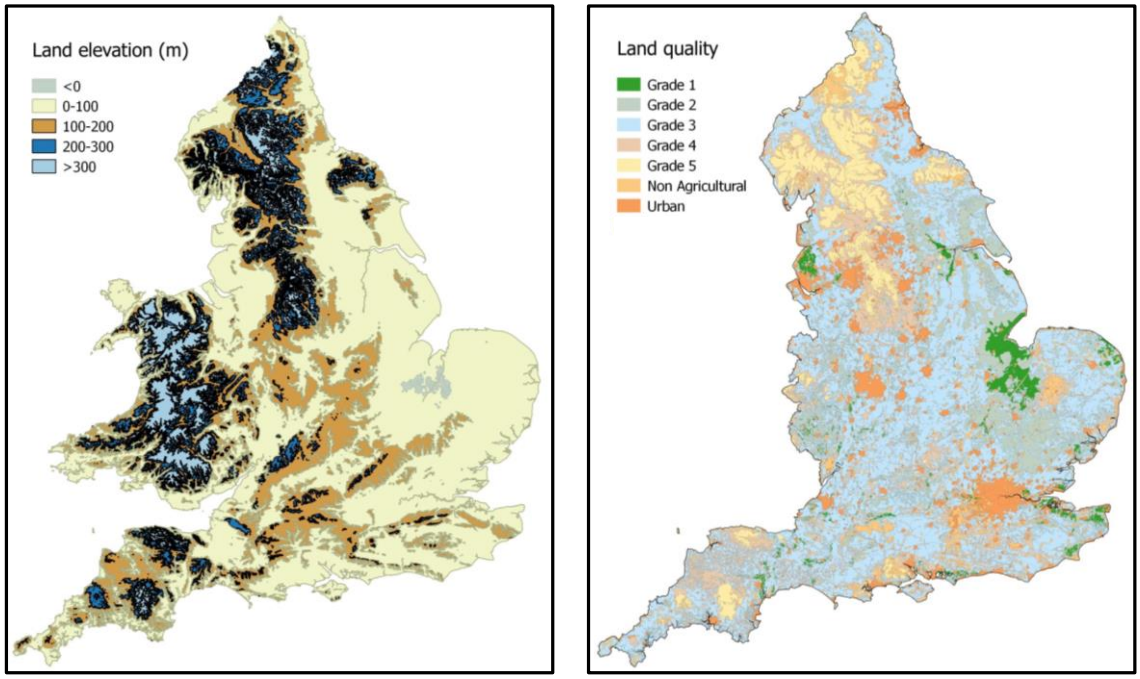
which occupational information was recorded.[8] These smaller geographical areas were therefore used as the units of analysis in these time periods.[9]

To cater for the differences in local climate and soil quality between these geographical units, a number of additional independent variables were included in the regression analyses, namely elevation, agricultural land quality, and latitude. Elevation has an obvious effect on the suitability of land for agriculture, with both climate and, generally, soil quality deteriorating with increasing altitude. The geographical units of analysis were therefore intersected with elevation contours, at hundred metre intervals, depicted on the left-hand side in map 1. Measures for agricultural land quality are not systematically and quantitatively available for the early nineteenth and eighteenth centuries. Therefore, modern-day land classification assessments were used.[10] Having to rely on modern-day data is, of course, not ideal for those parts of England in which land quality has changed significantly since the early nineteenth century, such as in the fenlands of East Anglia, but it is the best approximation currently available at low-level geographical scales. As with elevation, the geographical units of analysis were intersected with the land quality contours, depicted on the right-hand side in Map 1. For Wales, no land quality data were available at all; therefore, elevation is the only variable approximating land quality here. Latitude has a straightforward effect on local temperatures and, thereby, on conditions for agriculture. Longitude was also initially included, but found to be statistically irrelevant and, therefore, taken out of the model again.

[8] Kitson, 'The codebook of Anglican registration units' (Cambridge, 2010), http://www.campop.geog.cam.ac.uk/research/projects/occupations/britain19c/anglicanregistration.html.

[9] Benefitting from the availability of reliable contemporary population data from censuses in the early nineteenth century, some of the parishes and chapelries in the 1813-20 time period were consolidated into larger geographical units, so they could be matched up with the units at which population data were available. This reduced the number of units from 11,365 to 8,290. For the eighteenth century, the actual parishes/chapelries were used, as population data are not available anyway. The number of units here depended on the number of parishes/chapelries in which the occupation of fathers was reliably and consistently recorded in the baptism registers, which varied over time. For the early eighteenth century, 1,126 units were available as the basis for the analyses in this paper.

[10] As created in the Agricultural Land Classification surveys from 1966, and published by Natural England. See: http://publications.naturalengland.org.uk/publication/35012 for more details. In these surveys land in agricultural use was divided into five classes (Grades 1 to 5, indicating 'excellent', ' very good', 'moderate to good', 'poor' and 'very poor' land respectively) with the remainder divided into 'urban' and 'non-agricultural' land.



Map 1. Contours of elevation (left) and modern-day land quality (right), with which the geographical units of analysis were intersected

In addition to these factors affecting local conditions for agriculture, the presence of other agricultural workers can also be expected to have affected the need/room for agricultural labourers. Specialist agricultural workers employed by farmers such as ploughmen, threshers, and chaff cutters may be expected to have, to a degree, served as 'competitors' for work to more general farm labourers. The 1851 census recorded a significant number of farmers' sons which may, again, be expected to have reduced the need for external labour. Farmers themselves serve both as employers and competitors for farm labourers. Without farmers, no employment opportunities would exist at all. But farmers obviously also provided a fraction of the required agricultural labour themselves, reducing the need for external help. Men working in other agricultural occupations, such as dairymen, shepherds, and market gardeners are likely to have had very different, if any, requirements for agricultural labourers than farmers, and their numbers, differentiated by type, have been included separately in the regressions. An occupational denominator which requires special attention is the term 'husbandman' as it meant different things in different regions and time periods; since it could indicate a (small) farmer as well as an agricultural labourer, its number was included as a separate predictor variable.

Incorporating the above, equation (2) is expanded into:

$$L_A = \left[(1 + C_{Lat} \cdot Lat) \cdot \sum_{i}^{land\ types} C_{S,i} \cdot S_i\right] + \sum_{j}^{agri\ occs} C_{a,j} \cdot NL_{a,j} \qquad (3)$$

in which the land types are the intersections of the land quality and elevation contours (except for the Welsh units, where only elevation contours are available), with S_i indicating the surface area of 'type i' land in the geographic unit. Given the positive relationship between acreage (of a certain elevation, land quality, etcetera) and the demand for labour, all $C_{S,i}$ must be non-negative, and this constraint was therefore imposed on the regression analyses. The 'agri occs' refer to the non-labourer agricultural occupations, with $NL_{a,j}$ indicating their numbers; an overview is provided in Table 1, below. *Lat* indicates latitude.

For non-agricultural labourers, it is reasonable to assume that, there was a roughly linear relationship between the number of labourers (L_{na}) and the number of non-labourers (NL_{na}) in the same line of work, that is:

$$L_{na,k} \sim NL_{na,k} \tag{4}$$

for each non-agricultural occupation *k*. After all, if there were twice as many bricklayers in area X than in area Y, it is likely that area X also offered twice the amount of work for bricklaying labourers than area Y. The linearity of the labourer-to-non-labourer relationship can be tested for the, much later, 1911 census, which allocated labourers at a high level of detail, distinguishing, for example, between different types of building labourers; Figure 2 depicts the relationships between builders and building labourers per county, demonstrating linearity.

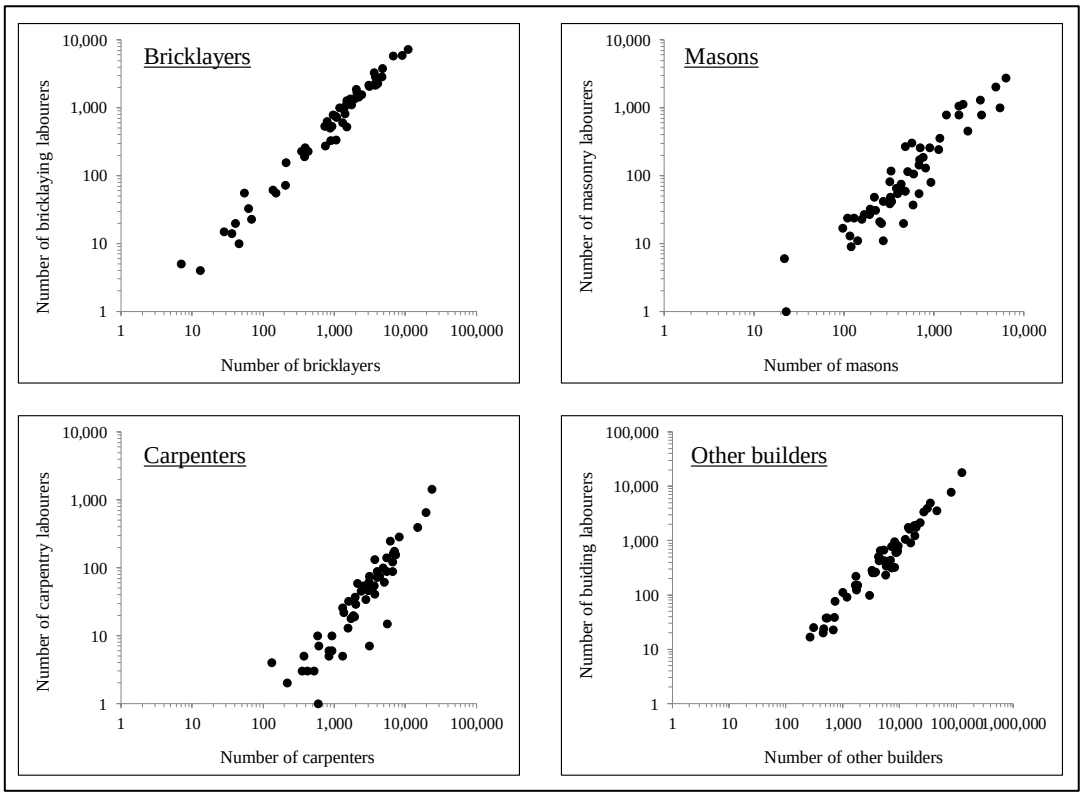


Figure 2. The relationship between building labourers and non-labourers, for different building trades, per registration county, 1911

Note: Each data point in the charts refers to a registration county.
Source: 1911 Census.

So, the number of non-agricultural labourers in a specific geographical area can be expressed as

$$L_{na} = \sum_{k}^{non-agri\ occs} C_{na,k} \cdot NL_{na,k} \tag{5}$$

with all $C_{na,k}$ constrained to non-negative values. As in the second part of equation (1), the different non-agricultural occupations can of course be grouped into sectors (mining, secondary, tertiary), but that straight-forward step has not been replicated here (see also Table 1, below).

Combining equations (3) and (5) then, the total number of labourers of all types in a given geographical unit can be expressed as:

$$L = \left[(1 + C_{Lat} \cdot Lat) \cdot \sum_{i}^{land\ types} C_{S,i} \cdot S_i \right] + \sum_{j}^{agri\ occs} C_{a,j} \cdot NL_{a,j} + \sum_{k}^{non-agri\ occs} C_{na,k} \cdot NL_{na,k} \quad (6)$$

with $C_{S,i}$ >= 0 for all *i*, and $C_{na,k}$ >=0 for all *k*. Note that for the agricultural non-labourers, it is not clear *a priori* whether they created a demand for agricultural labourers as their employers (in which case, $C_{a,j}$>0), or whether they reduced the demand for agricultural labourers by themselves providing the required labour (in which case, $C_{a,j}$<0). Therefore, no constraints were imposed on the $C_{a,j}$ coefficients. An overview of all the independent variables for the regression analyses is provided in Table 1.

Using a constrained, non-linear, multivariate regression analysis, this equation can now be solved statistically for all of the constants ($C_{...}$) using contemporary data at the level of smallest geographical unit for which occupational data are available – as discussed above.

Table 1. An overview of the regression predictor variables

Surface area, England (acres)[1]			**Surface area, Wales** (acres)[2]	**Non-labs, continued** (men/occ.)
Altitude (m)		*Land quality*	*Altitude (m)*	*Secondary sector*
Below 0	&	Grade 1	Below 0	Bricklayer
Below 0	&	Grade 2	0 to 100	Carpenter
Below 0	&	Grade 3	100 to 200	Mason
Below 0	&	Grade 4	200 to 300	Building, other - general[6]
Below 0	&	Urban	Above 300	Building, other - specialist[7]
Below 0	&	Non Agricultural		Baker
0 to 100	&	Grade 1	**Geo position** (m)	Butcher
0 to 100	&	Grade 2	Lattitude	Brewer/distiller
0 to 100	&	Grade 3		Woodworking - low skill[8]
0 to 100	&	Grade 4	**Non-labourers** (men/occ.)	Woodworking - high skill[9]
0 to 100	&	Grade 5	*Agricultural*	Iron manufacture
0 to 100	&	Urban	Yeomen/farmer	Non-ferrous raw metal
0 to 100	&	Non Agricultural	Husbandman	Blacksmith
100 to 200	&	Grade 1	Market gardener	Nail/pin manufacture
100 to 200	&	Grade 2	Sons of farmer	Other metal products[10]
100 to 200	&	Grade 3	Cattle/horse husbandry	Tanning
100 to 200	&	Grade 4	Other animal husbandry	Other leather/rope/bone
100 to 200	&	Grade 5	Agric. Manager	Textiles
100 to 200	&	Urban	Other agricultural	Clothing
100 to 200	&	Non Agricultural	*Mining*	Milling
200 to 300	&	Grade 2	Miner	Pottery
200 to 300	&	Grade 3	Quarry worker	Shipbuilding
200 to 300	&	Grade 4	*Other primary sector*	Glass production
200 to 300	&	Grade 5	All other prim. sector	Chemical industries[11]
200 to 300	&	Urban	*Tertiary sector*	Other industrial manufacture[12]
200 to 300	&	Non Agricultural	Transport - road	Other specialist manufacture[13]
300 to 400	&	Grade 3	Transport - inland water	
300 to 400	&	Grade 4	Transport - rail	
300 to 400	&	Grade 5	Transport - sea	
300 to 400	&	Urban	Maritime services[3]	
300 to 400	&	Non Agricultural	Other logistic services[4]	
400 to 500	&	Grade 3	Other tertiary sector[5]	
400 to 500	&	Grade 4		
400 to 500	&	Grade 5		
400 to 500	&	Non Agricultural		
Above 500	&	Grade 4		
Above 500	&	Grade 5		
Above 500	&	Non Agricultural		

Notes: [1]The number of elevation/land quality combinations included for England (38) is smaller than the theoretical number (7 x 8 = 56) would suggest because some intersections did not, in practice, occur (for example, grade 1 land quality above 500 metres elevation). Only non-empty intersections were included. [2]For Wales, no land quality data were available, and the highest elevation contour available was 300 metres. [3]Mostly dockworkers. [4]Mostly warehouse workers. [5]Since outside of transport, no other tertiary sector occupations were likely to employ general labourers on a significant scale, these were are lumped together. [6]Mostly consisting of men simply called 'builder' in the parish registers. [7] Mainly painters, glaziers, plasters, slaters, thatchers, tilers, and plumbers. [8]Mostly sawyers. [9]Coopers, furniture makers, basket makers, etc. [10] For example, gun makers, hinge makers, cutlery makers. [11]Salt boilers, soap makers, dye makers, etc. [12]Mostly men merely indicated by the terms 'manufacture' or 'mill man'. [13] Coach makers, clock makers, book binders, etc.

RESULTS OF THE NEW ALLOCATION METHOD

Agricultural labourers were identified as such in the 1851 census. This means that the validity and precision of the labourer allocation method described in the previous section can be tested by applying it to this census. If the approach works, it should accurately reproduce the actual division between agricultural and non-agricultural labourers. On a national scale, for England and Wales, the regression approach apportions 72.3% of all labourers to agriculture, and the remaining 27.7% to the others sectors.[11] This compares remarkably well with the actual distribution, directly derived from the census, of 71.8% agricultural and 28.2% non-agricultural labourers. Labourers working in transportation are also specified in the 1851 census, and constitute 3.6% of all labourers there.[12] Again, this is very similar to the results from the regression approach, which allocates 3.4% of all labourers to transport.

In Figure 3, the regression has been tested for smaller geographical units, namely at the level of counties. In the left-hand chart, calculated and actual numbers of agricultural labourers are compared; in the right-hand chart, the same has been done for non-agricultural labourers. Clearly, the method works very well at this lower geographical level also.

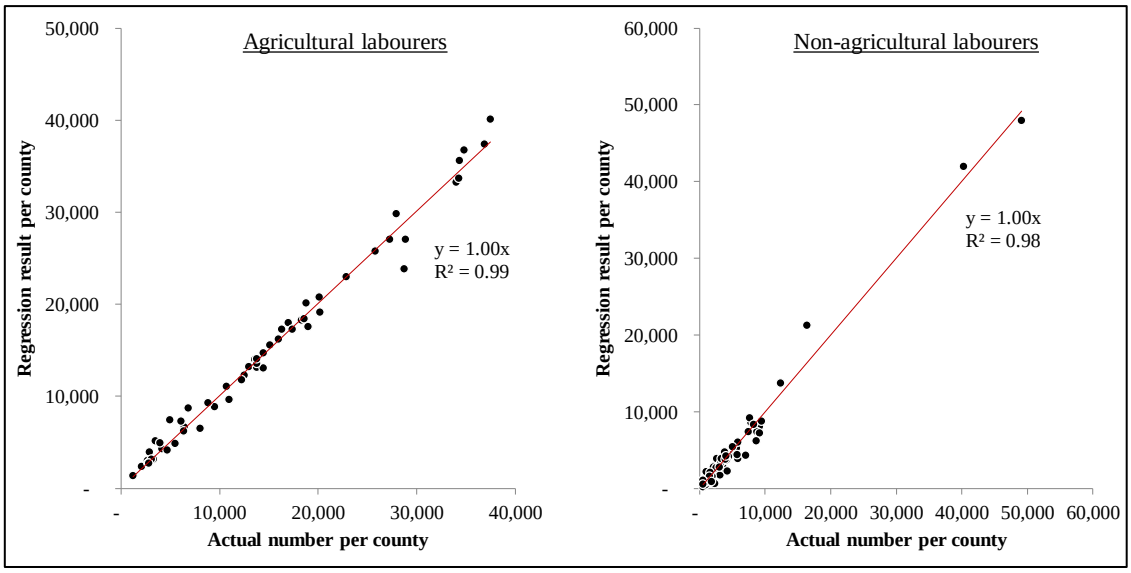


Figure 3. Comparison between the numbers of labourers calculated by the regression methodology and the actual numbers, per county (England and Wales, 1851)

Sources: 1851 census; regression-based labourer allocation approach.

Having demonstrated the accuracy of the approach, it can now be applied to data for which the labourer distribution is not known, such as the early-nineteenth century and eighteenth century parish register data, collected by the Cambridge Group. Combining these data with population figures from the 1811 and 1821 censuses – which, as discussed in footnote 9, leads to the need to consolidate some parishes/chapelries into larger geographical units – generates a dataset of 8,290 geographical units. Thirty per cent of the men in this dataset are (unallocated) labourers. The regression approach was applied to this dataset, resulting in a very good fit, as demonstrated by the R^2 of .89, and as confirmed

[11] Namely 22.3% to the secondary sector, 3.4% to transport, 2.0% to quarrying and mining. Since transportation
[12] These transportation labourers worked in maritime services and for the railways.

graphically in Figure 4, in which the actual and predicted numbers of labourers are compared at the county level.

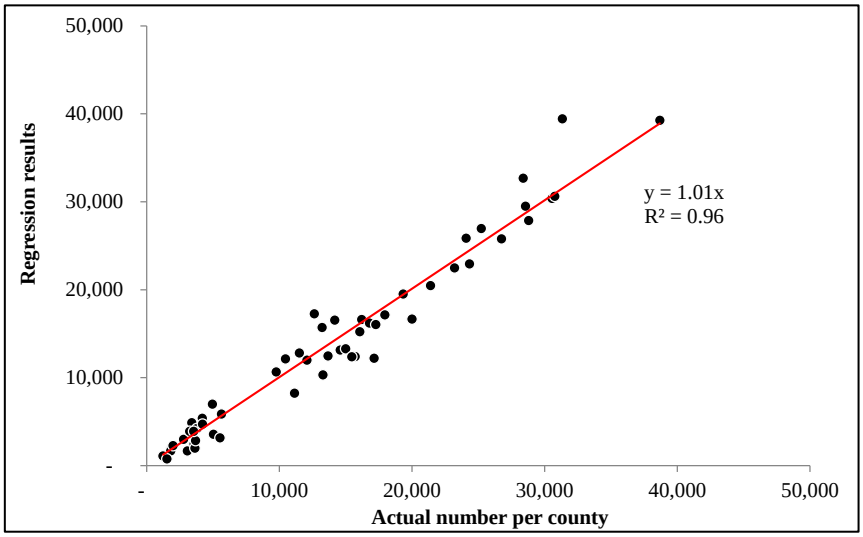


Figure 4. The total numbers of labourers per county as calculated by the regression approach compared to the actual numbers (England and Wales, 1813-20)

Sources: Parish register database for 1813-20; regression-based labourer allocation approach.

Only some of the regression coefficients turned out to be statistically significant ($p<.01$). These have been listed in Table 2, with their values.

Table 2. The coefficients resulting from the regression analysis, when applied to the parish register dataset (England and Wales, 1813-20)

Surface area, England (labs/acre)[1]			
Altitude (m)		*Land quality*	*Value*
Below 0	&	Grade 1	0.0274
Below 0	&	Grade 2	0.0464
0 to 100	&	Grade 1	0.0281
0 to 100	&	Grade 2	0.0403
0 to 100	&	Grade 3	0.0319
0 to 100	&	Grade 4	0.0219
0 to 100	&	Urban	0.0521
100 to 200	&	Grade 1	0.0472
100 to 200	&	Grade 2	0.0464
100 to 200	&	Grade 3	0.0240
100 to 200	&	Grade 4	0.0143
100 to 200	&	Grade 5	0.0144
100 to 200	&	Urban	0.0404
100 to 200	&	Non Agricultural	0.0209
200 to 300	&	Grade 3	0.0276
200 to 300	&	Grade 4	0.0085
200 to 300	&	Grade 5	0.0075
300 to 400	&	Grade 3	0.0177
300 to 400	&	Grade 4	0.0175
300 to 400	&	Grade 5	0.0036
300 to 400	&	Non Agricultural	0.0007
Above 500	&	Grade 5	0.0028

Surface area, Wales (labs/acre)[2]	
Altitude (m)	*Value*
0 to 100	0.0191
100 to 200	0.0144
200 to 300	0.0057
Above 300	0.0032
Geo position (per '000 km)	
Lattitude	-1.2048

Non-labourers (labs per non-lab)	
Agricultural	*Value*
Yeomen/farmer	0.1099
Husbandman	-1.0386
Market gardener	1.3219
Other agricultural	-0.5114
Secondary sector	
Bricklayer	0.4217
Mason	0.7295
Building, other - specialist[3]	0.5055
Woodworking - low skill[4]	1.5854
Nail/pin manufacture	0.1992
Pottery	0.0936
Shipbuilding	0.3613
Tanning	0.4199
Brewer/distiller	2.0475
Tertiary sector	
Transport - inland water	0.1252
Maritime services[5]	0.8812

Notes: [1] Land quality data and 400 and 500 metre contours available for England only; see main text for more details. [2]For Wales, no land quality data were available, and the highest elevation contour available was 300

metres. [3] Mainly painters, glaziers, plasters, slaters, thatchers, tilers, and plumbers. [4] Mostly sawyers. [5] Mostly dockworkers.

A particularly interesting regression coefficient in Table 2 is the one for husbandmen. As discussed, the term was an ambiguous one, not distinguishing clearly between small farmers and farm labourers. This term had all but disappeared from use in many parts of England and Wales by the 1813-20 period, but it is still encountered in parish registers in some corners of England, notably East Anglia, the south-west, and the north. But its value of close to minus one suggest that, by this time, the term husbandmen had become virtually synonymous with agricultural labourer. Independent confirmation of this result from the regression allocation approach can be obtained by analysing the development of the numbers of farmers between 1813-20 and 1851. Figure 5 presents something of a counter-factual in which, for the sake of the argument, it is assumed that husbandmen were farmers. As the figure clearly shows, it is remarkable then that the number of 'farmers' (including husbandmen) increased everywhere in England, except for those counties in which a substantial share of the 'farmers' were husbandmen, indicated in red in the chart. Here the number of 'farmers' typically declined, with steepness of the decline directly correlated to the husbandmen share amongst the county's 'farmers'. The conclusion must be that these husbandmen were not, in fact, farmers at all but, rather, agricultural labourers. Lancashire is the only county with a high number of husbandmen which showed an overall growth in 'farmers', which is likely explained by the particularly high growth of (actual) farmer numbers in this county resulting from the extensive drainage of the Lancashire mosses in this period.

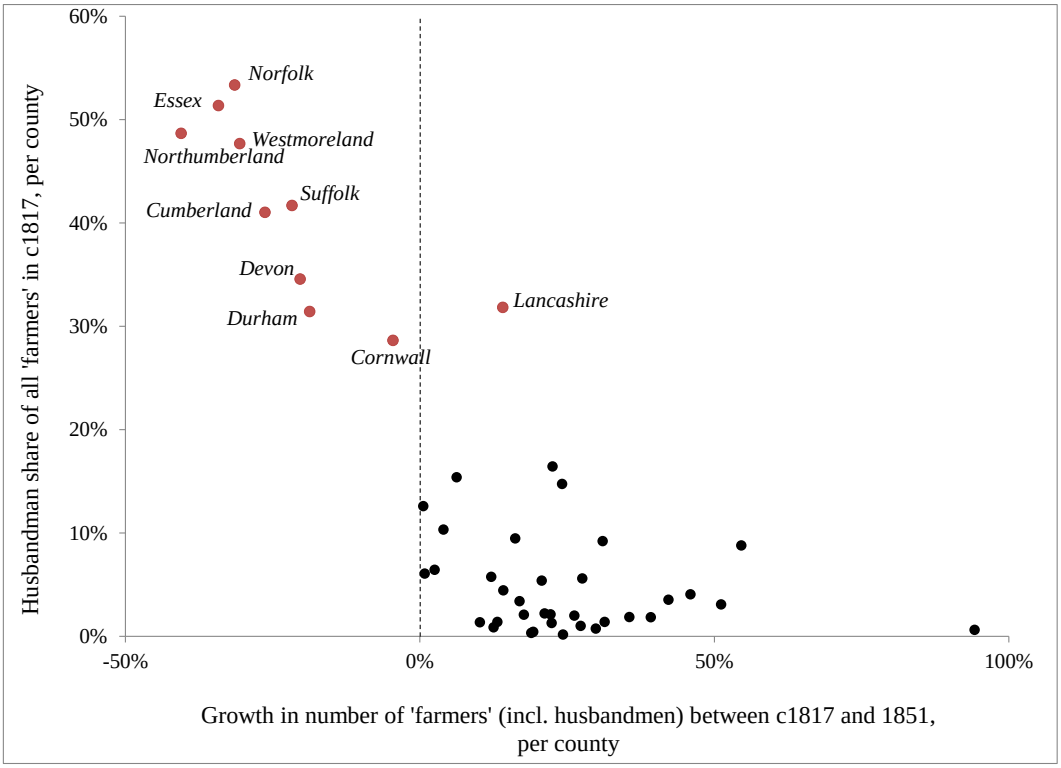


Figure 5. The development of the number of 'farmers' (including, for the purpose of the discussion, husbandmen) over the 1813-20 to 1851 period, compared to the husbandman share amongst these 'farmers' in English counties in 1813-20

The many geographic regression coefficients (elevation, land quality, latitude) have been more meaningfully summarised in Figure 6, showing the number of acres required for every additional

labourer. Unsurprisingly, the lower the elevation and the better the quality of the land, the more labourers could be employed on the same area of land.

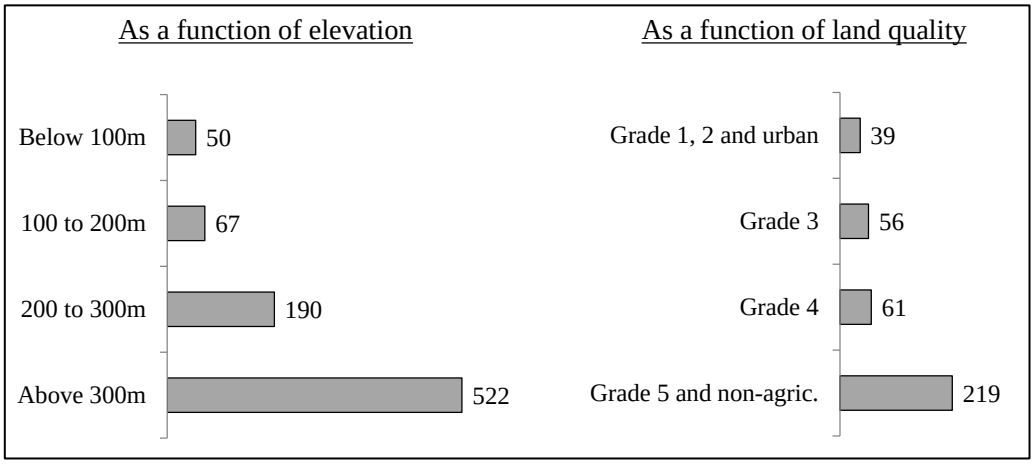
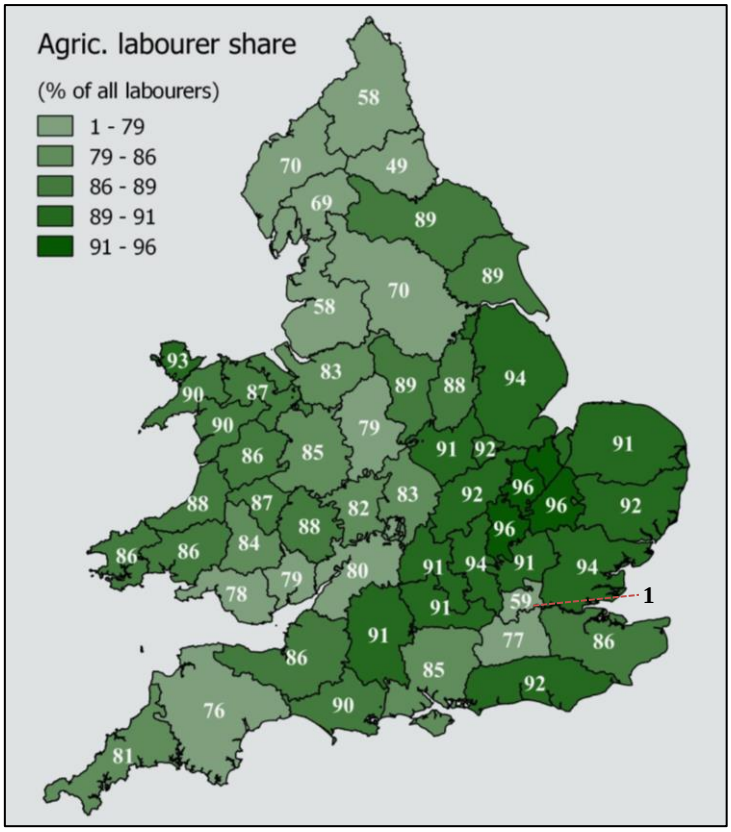


Figure 6. The number of acres per labourer as a function of elevation (England and Wales) and land quality (England only)

Note: The land quality categories are, as discussed, based on conditions today. Most of what is now urban land was in agricultural use at the time and, generally, of high quality, and has therefore been combined with grade 1 and 2 land.

Using these regression coefficients, labourers can now be allocated to sectors and, even, sub-sectors. At the national scale (England and Wales), eighty-four per cent of all labourers were agricultural, with the vast majority (80%) working on farms and the remainder (4%) employed by market gardeners. Of the sixteen per cent of labourers that were non-agricultural, fifteen per cent were employed in the secondary sector; only one per cent is to be allocated to the tertiary sector, all of them to transport. Nearly half the secondary sector labourers appear to have been employed in building and construction, with the remainder working in other sub-sectors, such as low-skill woodworking, nails and pin production, large scale brewing and tanning, and shipbuilding.

At the level of counties, the allocation varied considerably, as shown in Map 2. Unsurprisingly, in the Southern and Eastern English counties with their large farming sectors and big capitalist farms, usually over ninety per cent of labourers were employed in agriculture. The northern counties, characterised by family farms, had a relatively large share of non-agricultural labourers, particularly in industrialising and urbanising Lancashire. Unsurprisingly, in London, ninety-nine per cent of labourers were non-agricultural, the large majority working in the secondary sector, with partially urbanised Middlesex and Surrey also showing relatively low agricultural labourer shares.



Map 2. The share of labourers working in agriculture, by county (1813-20)

The results of the labourer allocation in *rural* England and Wales is interesting, albeit it perhaps not very surprising. 7,503 of the 8,290 geographical units in this data set can unambiguously be identified as rural.[13] In Figure 7, the regression-determined number of *agricultural* labourers per farmer, per county, is compared to the *total* number of labourers in the *rural* parishes in each county, as directly taken from the parish register data. With a R^2 of .98, the correspondence is almost perfect. The slope of the linear regression line shows that ninety-seven per cent of the rural labourers were employed in agriculture.

When the regression technique is applied to earlier time periods, the resulting labourer allocations turn out to be very similar to those in the 1813-20 period. The accuracy of these earlier regressions is somewhat affected by the fact that occupational information from parish registers is only available for a sample of the total number of parishes in England and Wales. But by calculating labourer numbers at the level of individual counties first, and then combining the county averages into a national one, using Wrigley's work on county populations, a national estimate can nevertheless be constructed, albeit with a margin of error.[14] Between eighty-five and eighty-nine per cent of all labourers in the early eighteenth century were employed in agriculture, with the remainder in the secondary sector, mainly in building. The figures for the mid and late eighteenth century were also in this range.

[13] By excluding all units that contain a (market) town, erring on the side of caution.
[14] Wrigley, *The early English censuses* (Oxford: British Academy Records of Economic and Social History, new series, 2011).

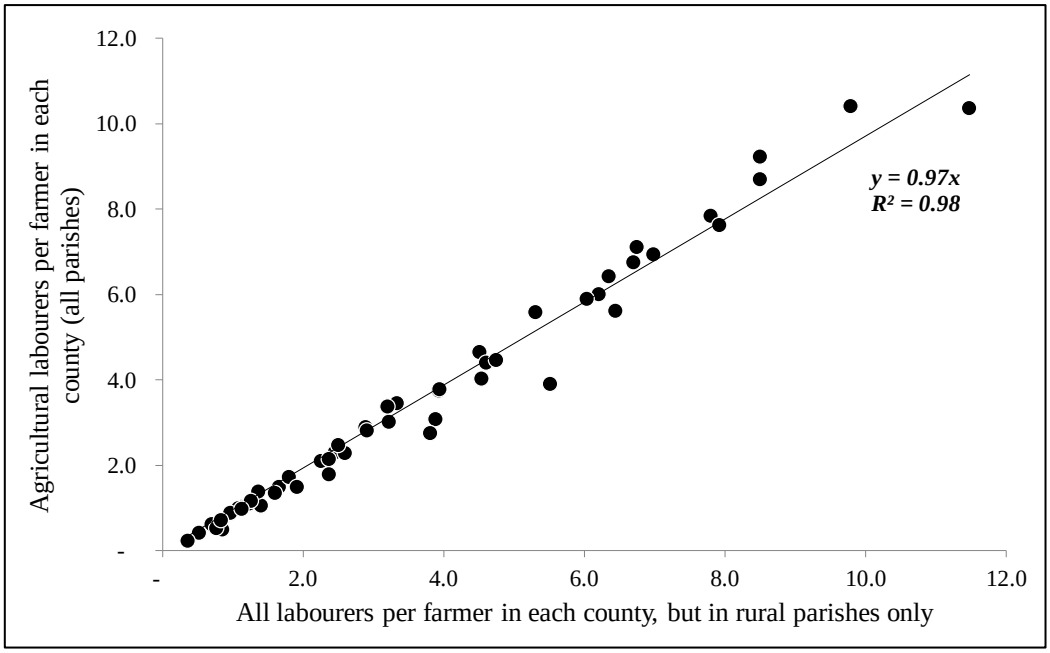


Figure 7. The number of agricultural labourers per farmer in all parishes (as determined by the regression approach) compared to the total numbers of labourers per farmer in rural parishes only (as directly derived from the parish register data set) (per county, 1813-20)

The stability in the agricultural/non-agricultural share of the labourer population over the eighteenth and early nineteenth century is not as surprising as it might, at first sight, appear. The number of agricultural labourers per farmer increased threefold over the period, from 1.0 in c.1710 to 3.2 in c.1817, as a result of the increasingly capitalist nature of farming in England and, to a lesser extent, Wales.[15] But the number of non-farmers who were likely to employ labourers increased similarly fast, in comparison to the rise in the number of farmer numbers. For every English and Welsh farmer in c.1710, there were 0.4 builders, but in c.1817 this ratio had also increased nearly threefold, to 1.0. The increase in agrarian capitalism being paralleled by a roughly similarly sized increase in the number of non-agricultural employers for labourers explains the relative stability in agricultural and non-agricultural labourer shares over the eighteenth century.

COMPARING THE RESULTS TO THOSE FROM EXISTING METHODS

Comparisons can only be made at the national scale, as existing techniques only work at that level. The national results of the regression approach (84% in agriculture, 15% in the secondary sector, 1% in the tertiary sector in c.1817) are not dissimilar to those from the Saito/Shaw-Taylor approach, with the exception of the tertiary sector. Saito and Shaw-Taylor base their allocation on the assumption that labourers formed an equal share of the secondary and tertiary sectors in c.1817 as they did in 1851, at 10.1 and 8.1 per cent respectively, with the remaining labourers presumed to have worked in agriculture. Applying these percentages to the c.1817 quasi census leads to eighty-one per cent of all labourers being allocated to agriculture, fourteen per cent to the secondary, and five per cent to tertiary sector.

[15] The c.1817 figure includes husbandmen, as the term husbandman had, by this time, become synonymous with farm labourer, as discussed above.

The much larger allocation of labourers to the tertiary sector than in the regression-based approach can be explained by comparing the composition of the tertiary sector in 1851 and c.1817. The tertiary sector labourers in 1851, as indicated by the census, worked in only two, very specific types of transportation occupations: in maritime services and the railways. These occupied a significant number of men in 1851, with 37,000 labourers and 42,000 other workers. These 42,000 other workers made up about four per cent of all tertiary sector employment for non-labourers. In c.1817, however, these occupations provided only very limited employment, with about 1,600 men working in maritime services, and (unsurprisingly) none in rail transport. These 1,600 men represent only 0.3 per cent of all tertiary sector employment in c.1817 and the tertiary sector labourer share in c.1817 must therefore have been much lower than in 1851. If the tertiary sector labourer allocation is based on the size of the maritime services and railway occupations rather than on that of the tertiary sector as a whole, the Saito/Shaw-Taylor approach would allocate of eighty-five per cent to agriculture, fourteen per cent to the secondary sector, and one per cent to the tertiary sector – virtually identical to the national distribution resulting from the regression approach. That two such different approaches arrive at such similar results is, surely, highly encouraging. It also confirms the expectation by Saito and Shaw-Taylor that their methodology would provide good results at the national level.

As discussed, the labourer allocations in the national accounts literature are flawed for a number of reasons. It is therefore not surprising that they differ significantly from the results of the regression method and the Saito/Shaw-Taylor approach. Crafts essentially allocated all labourers to agriculture. Broadberry et al apportion thirty-two per cent of all labourers to the secondary sector in the eighteenth and early nineteenth centuries; this is more than twice the percentage generated by the regression approach (and by Saito and Shaw-Taylor).

* * *

BIBLIOGRAPHY

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

Broadberry, S., B.M.S. Campbell, and B. van Leeuwen, 'When did Britain industrialise? The sectoral distribution of the labour force and labour productivity in Britain, 1381–1851', *Explorations in Economic History,* 50:1 (2013), pp. 16-27

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

Kitson, P.M., 'The codebook of Anglican registration units' (Cambridge, 2010), http://www.campop.geog.cam.ac.uk/research/projects/occupations/britain19c/anglicanregistration.html

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

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

Wrigley, E.A., *The early English censuses* (Oxford: British Academy Records of Economic and Social History, new series, 2011).