Online Supplementary Materials

Table of Contents

S1 Data Quality 1

S1.1 Distributional Analysis 1

S1.2 Error Type Subsets 5

S1.3 Clerical Exemption and Noble Evasion 9

S2 Worst Case Scenarios (WCSs) Overview 10

S3 WCS1: Maximal Modifier Rages 11

S4 Proxy Data WCS Models 12

S4.1 Taxpayer Exemption Rates 12

S4.2 WCS2: Taxpayers as Households 14

S4.3 WCS3: Taxpayers as Adult Males 15

S5 Summary of WCSs 16

S6 Sensitivity Analysis 18

S6.1 Implementation 18

S6.2 Results 19

S7 Bibliography 23

S1 Data Quality

S1.1 Distributional Analysis

In order to cross-check the validity of the 1524/5 Lay Subsidy data, we first subjected the univariate distributions of taxes and taxpayers to analysis (see Table S1, pp.6-8 below, and supplementary dataset “main_dataset.csv”). Log-normal distributions are expected for modern urban variables, and indicate that the data’s underlying generative processes are unconstrained (i.e. strongly interacting) and multiplicative (Batty, 2006; Gibrat, 1931; Limpert, Stahel, & Abbt, 2001; Montroll & Shlesinger, 1982). We might therefore expect that the 1524/5 Lay Subsidies town variables would also exhibit log-normality if they are valid, and their values (taxpayer counts, tax returns) are proportional to the variables in question (population, socioeconomic output). Figure S1 shows the distributions of the natural logs of our main sample of towns from the 1524/5 Lay Subsidy (n = 93). The distribution of the logged tax returns is unimodal, two tailed, and looks approximately Gaussian (Figure S1a). In contrast, the distribution of logged taxpayer counts is strongly right skewed—such that it could have come from a Gaussian distribution with the left tail cut off (Figure S1b).

In order to quantitatively test the probability that our samples are approximately log-normal, we use the one-sample Kolmogorov-Smirnov (KS) goodness-of-fit test for normality. As opposed to the highly-sensitive Shapiro-Wilk and Anderson-Darling normality tests with greater statistical power (Basu, Shioya, & Park, 2011: 22-24; Razali & Wah, 2011; Stephens, 1974), the KS test is more appropriate because the 1524/5 Lay Subsidy dataset is *not a random sample*. The Lay Subsidy data are a relatively small sample of complete records from surviving historical documents. Even if their survival and completeness was random, the numerous missing cases constitute a considerable degree of sampling error—especially in the form of inliers. This introduces a high risk of type II error, making extremely-efficient normality tests with high statistical power inappropriate in this context (Basu et al., 2011).

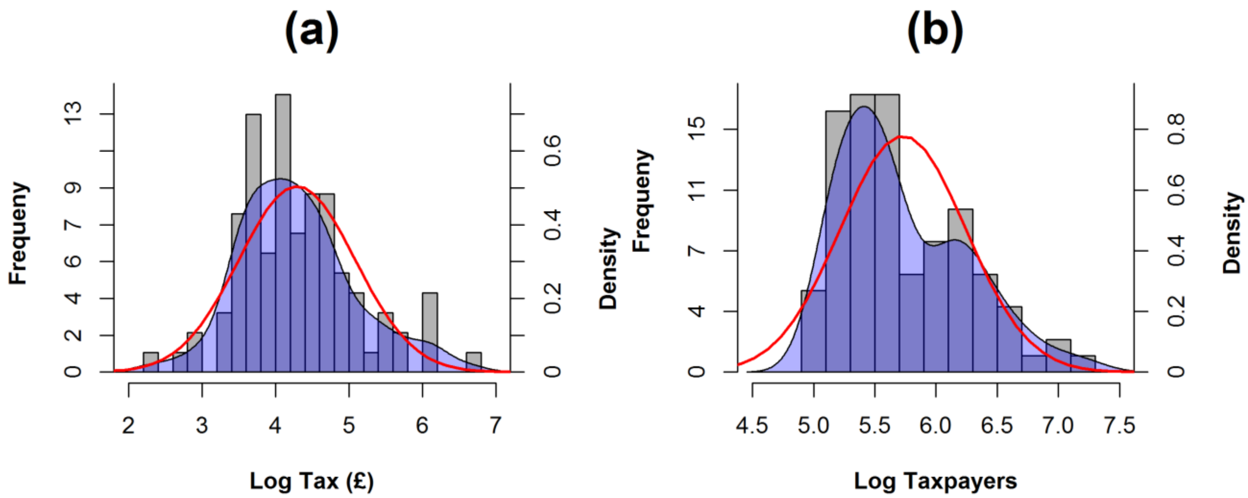

Figure S1. Histograms and kernel density PDFs of the 1524/5 Lay Subsidy main sample of towns (n = 93). (a) natural log of tax returns; **(b)** natural log of taxpayer counts. Blue distributions are kernel density PDFs of each variable, and red lines are Gaussian curves fitted to the sample means and variances.

As suggested by visual inspection of Figure S1, the two variables yielded divergent KS test results. A one-sample, two-sided KS test for normality of the logged tax returns (n = 93)—using a Gaussian null hypothesis with the mean and variance of the sample—failed to reject the null hypothesis (D = 0.08; p > 0.73). Visual inspection of the ECDF and null hypothesis (Figure S2a) also shows a fairly tight fit, with the null hypothesis (red) well within the 95% C.I. We can

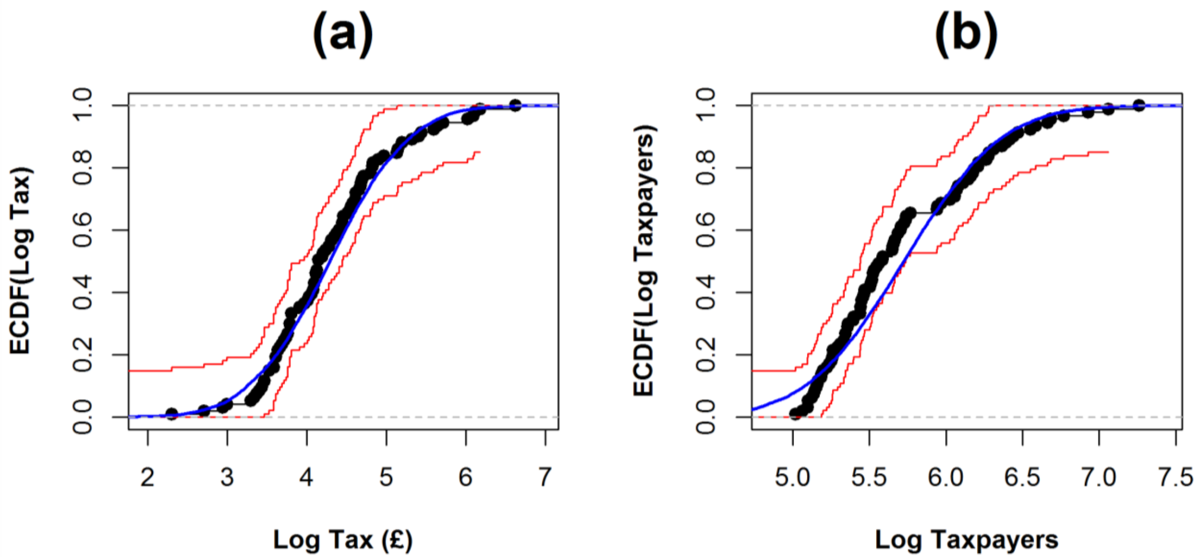

Figure S2. KS tests of the sample data with 95% C.I. envelopes. Black points and lines are the ECDF of logged sample data, thin red lines are the upper and lower 95% C.I. envelopes, and blue line is the Gaussian null hypothesis fitted to the log sample means and variances.

therefore assume that the sample of tax returns is approximately log-normal. However, the same test for the taxpayer count data (n = 93) yielded more ambiguous results (D = 0.14; p = 0.06). This is just over the standard $\alpha < 0.05$ level of confidence, and the graph of the KS test (Figure S2b) has a poor visual fit, and the null hypothesis grazes the lower 95% C.I.

As suggested above, the poor log-normal fit of the Lay Subsidy taxpayer sample appears to be the result of its arbitrary lower threshold of towns with at least 150 taxpayers. Neither the histogram of taxpayer counts (Figure S1b) nor the KS test graph (Figure S2b) show any signs of a left tail. However, the fat right tail of the logged taxpayer counts conforms to Zipf's law, with a log-log rank-size slope of roughly b = -0.35 (Figure S3). It is notable that this result was attained even without some of the largest English towns of the period, including London. Zipfian distributions have frequently been shown to be a pronounced feature of the right tails of log-normal distributions for modern urban variables with threshold size cut-offs—like our sample threshold of 150 taxpayers (Adamic & Huberman, 2002; Batty, 2006; Gabaix, 1999; Pumain, 2000). Historical case studies have also found evidence for this threshold effect among pre-modern cities (Bairoch, 1991; De Vries, 1984; Dittmar, 2011; Johnson, 1981; Pumain, 2000).

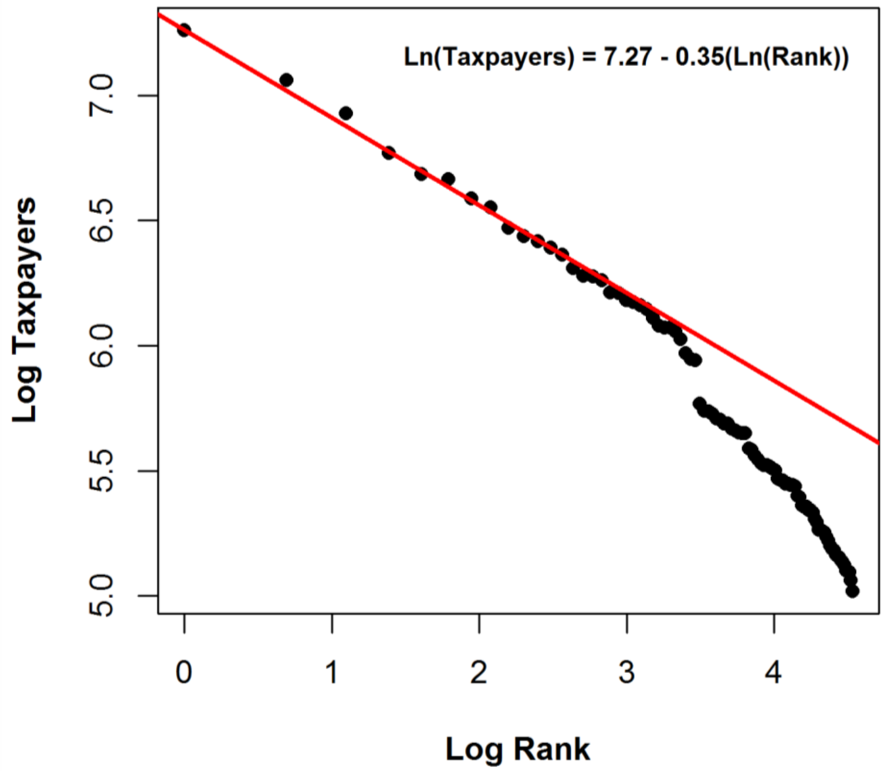

Figure S3. Zipfian Tail of Log-Log Rank Size Plot of the 1524/5 Taxpayer Counts (*n* = 93). Red line equation is given in upper right corner, estimated by visual fit of slope with intercept equal to size of largest town (in terms of Log Taxpayers).

As expected, an expanded sample of towns from the 1524/5 Lay Subsidy with more than 50 taxpayers (n = 134) is approximately lognormal. These additional town taxpayer counts come from the revised estimates of Rigby (2010: Appendix I), such that all towns with taxpayer counts between 50 and 150 (n = 41) were added to our main sample of taxpayer counts (n = 93). The expanded sample of 1524/5 taxpayer counts is available to readers in the supplementary dataset “add_taxpayers.csv”. Visual inspection of this expanded taxpayer counts dataset (n = 134) shows approximate lognormality, with a noisy but univariate and two-tailed distribution (Figure S4a). Likewise, a two-tailed KS normality test of the logged taxpayer counts (D = 0.06; p > 0.65) fails to reject the Gaussian null hypothesis, and exhibits a much better visual fit (Figure S4b).

This finding of both log-normality and Zipfian behavior for the sample data indicates that they exhibit statistical properties typical of modern urban data. This suggests that the 1524/5 Lay Subsidy tax returns and taxpayer counts are representative samples of Tudor towns c.1524/5, and therefore that the recorded values are proportional to socioeconomic output and population, respectively. Given the broad scholarly consensus on the quality of the 1524/5 Lay Subsidy data (see section 4.1 of the main text), we therefore consider these variables valid metrics of population and socioeconomic output for settlement scaling analysis.

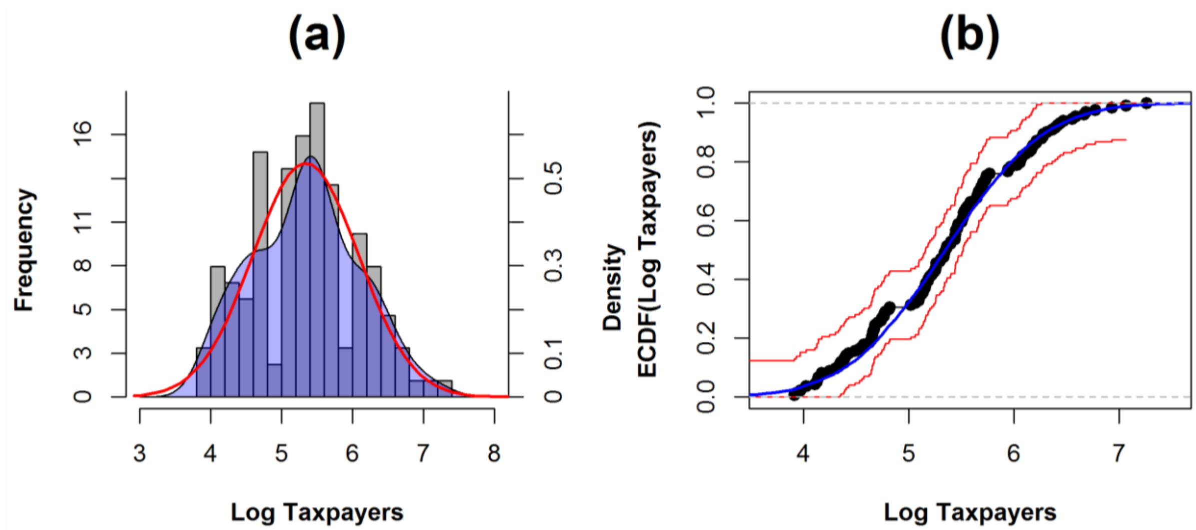

Figure S4. Analysis of the Expanded Log Taxpayer Counts Dataset of towns with at least 50 taxpayers (n = 134). (a) Histogram (grey), kernel density PDF (blue), and Gaussian curve fitted to the sample means and variances (red). **(b)** KS test of the Log Taxpayer data ECDF (black), 95% C.I. envelopes (red), and Gaussian null hypothesis fitted to the log sample mean and variance.

S1.2 Error Type Subsets

The accuracy of the 1524/5 Lay Subsidy returns have been called into question for several important reasons, which we discuss here in detail. The analyses presented here serve to place controls on the data for scaling analysis. In this section, we go through each concern methodically and evaluate their potential for either systematic or random error. The cases most susceptible to these errors are then subdivided from the dataset (n = 93) into corresponding likely error-type subsets (see Table S1, pp.6-8 below, and supplementary dataset “main_dataset.csv”). All possible combinations of these error-type subsets are analyzed in section 5 of the main text, as well as in the WCSs below, to evaluate the sensitivity of our results to each potential source of error.

Inconsistent Assessments

It has been suspected that municipalities across England were not assessed consistently. This is inevitably true to some degree, since different municipalities were assessed by different tax commissions, who themselves faced different local circumstances (Schofield, 2004; Sheail, 1968). Alan Dyer (2000b) has even suggested that the corruption of local assessors, motives for evasion, and general incompetence in the tax’s administration may mitigate the accuracy of the 1524/5 returns in general. However, this caveat only suggests random errors across all cases, which should not systematically impact our analysis. If anything, a large number of random errors should make the data noisy and depress the correlation coefficient.

It has also been argued that suspected errors in the tax returns are geographically autocorrelated because of the systematic biases of county-level tax administration (Dyer, 2000b;

Goose & Hinde, 2006). Nevertheless, we should expect a consderable degree of spatial autocorrelation in the tax returns given the strong interdependence of rural and urban economies at this time (Rigby, 2010). Indeed, multiple analyses of the geographical distribution of lay wealth in 1524/5 (at the town and county levels) have found strong spatial concentrations of economic activity in the south and east of England (Darby, Glasscock, Sheail, & Versey, 1979; Dyer, 2000b; Rigby, 2010; Schofield, 1965; Sheail, 1968). The broad consistency of the 1524/5 Lay Subsidies' geographical patterning suggests that these data primarily reflect the well-studied (sub-)regional, historical variability in demography and economic output.

Other detailed analyses of the 1524/5 Lay Subsidy returns have found them to be highly reliable. Roger Schofield's analysis of the subsidy yields of 1524-72 relative to their corresponding probate inventories indicates that the 1524/5 subsidies were the most accurate taxes of the era. Not only was the tax accurate among the lay public, even the wealthiest peers (who were otherwise most susceptible to underassessment and evasion) were strictly handled by special royal commissioners in 1524-5 (R. Schofield, 2004: 98–102, 168–77, 201–18). Moreover, the 1522 Muster survey, disguised as an assessment of military capacity, provided a dress rehearsal for the administration of the subsidies. The Muster simultaneously assessed expected lay wealth prior to the 1524/5 subsidies in order to set expectations for the 1524/5 returns (see e.g. Campbell, 1981; Cornwall, 1962a). Taken together with low spatial autocorrelation and the distributional analysis above, we find no reason to believe that these suspected errors invalidate the 1524/5 Lay Subsidy returns for scaling analysis.

Problematic Returns

A second potential source of error concerns a number of towns that have been explicitly identified as having problematic taxpayer numbers and/or tax receipts due to underassessment or deficient returns. Some towns like Lynn and Boston have incomplete tax records (Dyer, 2000a). Other cases, including boroughs directly assessed by royal commissioners, seem to have been systematically under-assessed when their taxpayer counts are compared to contextual expectations from later demographic data (the Chantry Certificates of 1548 and the Bishop's Census of 1563).[1] In addition, several counties have been recognized as having "deficient," under-assessed, or otherwise suspiciously low returns in 1524/5 when compared with later tax records (Dyer, 2000b; Goose & Hinde, 2006; Sheail, 1968). Other towns, assessed by these suspect county assessors, have likewise been identified as under-assessed.[2] As noted by Rigby (2010), the use of later tax records to judge the accuracy of the 1524/5 returns makes potentially hazardous assumptions, especially since both the population of England and the fortunes of individual towns changed considerably over the intervening period (see e.g. Dyer, 1995, 2000c; Palliser, 1992: 263–75). Despite these issues, we conservatively group these "problematic" cases (n = 16) together and set them aside as an error-type subset for use in sensitivity analysis (see Table S1).

---

[1] From our sample of towns with >150 taxpayers: Gloucester, York, and Hull are noted as underassessed by Alan Dyer (2000a, 2000b) when compared to contextual information (see Palliser, 1979, pp. 201–25), the Chantry Certificates of 1548, and the Bishop's Census of 1563. Dyer has likewise identified Nottingham and Derby as having been underassessed, and he excluded the low tax receipts of both Ely and Rochester from his listings altogether (Dyer, 2000a, 2000b).

[2] From our sample of towns with >150 taxpayers: Beverley, Pontefract, and Selby in Yorkshire; Taunton and Wells in Somerset; Bodmin in Cornwall (see Dyer, 2000a, 2000b).

Rural Municipalities and Small Town Errors

The 1524/5 tax returns are vulnerable to bias from municipalities that included both towns and their rural hinterlands (see Goose & Hinde, 2006). For example, A. Dyer (2000a) notes that the data on the municipalities of Crediton, Tiverton, Maidstone, and High Wycombe included parts of adjacent rural areas, while the return from the town of Huntingdon is conflated with the parish of Godmanchester. These problems tend to be exaggerated for smaller towns due to the uncertain spatial extent of each municipality (Goose & Hinde, 2006) and potentially less efficient assessment (Dyer, 2000b). Smaller towns were also more vulnerable to rapid population fluctuations, the magnitude of which could easily swamp the permanent local residents. These fluctuations were variously caused by high seasonal occupational mobility, high labor mobility associated with underemployment, and an outstanding number of impoverished vagrants that circulated through towns and the countryside looking for temporary work and shelter (Cornwall, 1962a, 1988: 216–30; Galley, 1998: 20–7; Palliser, 1992: 63–6). The presence of rural cottage industries or the residence of a few very wealthy merchants could also greatly distort the returns of small towns (Cornwall, 1962a, 1988: 51–61, 70–87). While large towns were also subject to these fluctuations, their greater size will have dampened their proportional impact. As such, we cautiously group all towns with under 200 taxpayers together with the five cases named above (*n* = 25 in total). Together these form the "small town/rural boundary" error-type subset for sensitivity analysis (see Table S1).

Table S1. Full 1524/5 Lay Subsidy Dataset of Towns, Indicating their Error Type Subsets. These data come from Rigby (2010), A. Dyer (2000a), and Sheail (1968), and give preference to the enumerations in that order.

	Town	Taxpayers	Tax (£)	Error Type Subset
1	Norwich	1423	749	--
2	Bristol	1166	479	--
3	Exeter	1019	441	--
4	Salisbury	799	411	--
5	Canterbury	784	269	--
6	Coventry	725	448	--
7	Colchester	701	204	--
8	Bury St Edmunds	645	180	--
9	Lincoln	625	124	--
10	Hereford	611	124	--
11	Winchester	596	86	--
12	St Albans	580	95	--
13	Shrewsbury	550	101	--
14	Oxford	533	105	--
15	Reading	531	223	--
16	Cambridge	524	97	--
17	Worcester	499	171	--
18	Yarmouth	497	125	--
19	Ipswich	484	282	--
20	Northampton	474	91	--

21	Southampton	450	101	--
22	Leicester	427	107	--
23	Newbury	414	121	--
24	Lichfield	391	36	--
25	Saffron Walden	380	61	--
26	Hadleigh	311	109	--
27	Plymouth	310	85	--
28	Beccles	307	74	--
29	Chichester	301	63	--
30	Wymondham	287	45	--
31	Basingstoke	284	67	--
32	Windsor	267	94	--
33	Alton	260	55	--
34	Barking	256	66	--
35	Wisbech	252	44	--
36	Ottery St Mary	250	79	--
37	Spalding	250	39	--
38	Kingston-on-Thames	249	62	--
39	Stamford	247	100	--
40	Cullompton	245	60	--
41	Marlborough	237	85	--
42	Colyton	236	63	--
43	Dunwich	235	40	--
44	Walsingham	232	58	--
45	Aylsham	231	34	--
46	Sudbury	231	61	--
47	Barnstaple	230	38	--
48	Totnes	220	144	--
49	Shaftesbury	213	60	--
50	Bath	212	45	--
51	Peterborough	212	44	--
52	Glastonbury	209	42	--
53	Ramsey	209	28	--
54	Lewes	207	43	--
55	Aylesbury	202	37	--
56	York	871	230	Problematic
57	Gloucester	466	134	Problematic
58	Rochester	437	59	Problematic
59	Ely	382	69	Problematic
60	Lynn	320	302	Problematic
61	Taunton	300	86	Problematic
62	Boston	295	111	Problematic
63	Nottingham	295	56	Problematic
64	Bodmin	285	37	Problematic
65	Hull	284	108	Problematic
66	Beverly	266	63	Problematic
67	Derby	232	32	Problematic
68	Wells	221	61	Problematic
69	Pontefract	181	15	Problematic and Small/Rural

70	Selby	170	10	Problematic and Small/Rural
71	Manchester	163	32	Problematic and Small/Rural
72	Maidstone	480	169	Small/Rural
73	Crediton	433	74	Small/Rural
74	Huntingdon	433	82	Small/Rural
75	Tiverton	289	53	Small/Rural
76	Lavenham	199	180	Small/Rural
77	Maldon	193	72	Small/Rural
78	St Neots	193	34	Small/Rural
79	Woodbridge	192	45	Small/Rural
80	Henley-on-Thames	191	41	Small/Rural
81	Wellingborough	188	30	Small/Rural
82	Croydon	185	37	Small/Rural
83	Modbury	179	50	Small/Rural
84	Petworth	178	29	Small/Rural
85	High Wycombe	175	44	Small/Rural
86	Holbeach	174	31	Small/Rural
87	Torrington	173	19	Small/Rural
88	Dorchester	171	77	Small/Rural
89	Wolverhampton	168	27	Small/Rural
90	Penryn	164	20	Small/Rural
91	Oundle	164	34	Small/Rural
92	Devizes	158	50	Small/Rural
93	Louth	151	38	Small/Rural

S1.3 Clerical Exemption and Noble Evasion

The high nobility was purposely taxed very strictly to prevent evasion in the 1524/5 Lay Subsidies (even including their own royally appointed assessors), the effects of which can be seen quantitatively in Schofield's finding of very high accuracy for the 1520s subsidies compared to other 16th century assessments (Schofield, 2004). Moreover, the vast majority of very wealthy merchants and nobles were located in large towns and cities (Cornwall, 1988; Palliser, 1992; Sheail, 1968). If the proportion of wealthy taxpayers was actually larger due to evasion, then this would mean deflated returns for larger towns. Correcting such a bias for larger towns would only increase the estimated scaling exponent—thereby strengthening our finding of IRS among Early Tudor towns.

Another hypothetical source of systematic bias is the exemption of the clergy and ecclesiastical institutions. The distribution of clerical wealth across English towns in the early 16th century remains unstudied (Jurkowski, 2015), but it is possible that larger towns had greater untaxed income/moveables on average than small towns (Cornwall, 1988; Thompson, 2014). The county level geographical distribution of clerical wealth in the early 16th century conformed to the overall distribution of lay wealth (Schofield, 1965). Most of the regular clergy and wealthy ecclesiastical institutions were rural (and also in sharp decline by 1524/5).

One might assume that larger towns had much greater proportions of ecclesiastical wealth, if only due to greater numbers of ecclesiastical institutions like hospitals, guilds, and cathedrals. Nevertheless, Cornwall (1988) has found that the secular clergy were middling (beneficed) or poor (unbeneficed) on average, and that town ecclesiastical institutions were often in debt or barely breaking even (i.e. net income/moveables). The extant urban data is limited, but even in a large and prosperous town like Coventry, the secular clergy was small and middling-to-poor. Here, Lay wealth made up 94.8% of all assessed wealth, the secular clergy 2.8%, and the institutional coffers of cathedral and guilds 2.4% (Cornwall, 1988). Certainly major church cities like Cambridge and York had more, but these represent the exception rather than the norm.

If we assumed that a systematic bias for large towns existed, then correcting such a bias for larger towns would again only increase estimated scaling exponents in favor of IRS. The magnitude of this hypothetical bias is unclear. If we assume that ~5% ecclesiastical wealth is a moderate-to-high proportion of total urban wealth, then the ecclesiastical bias in favor of the largest towns would be modest (5-10%, perhaps ~7.5%). Such a systematic bias is small in comparison to the magnitudes of our demographic Worst Case Scenario biases for poverty exemptions in sections S2 through S5, below (about 1/5th to 1/10th the size).

Moreover, if both systematic town size biases existed (for clerical and poverty exemptions), then they should cancel each other out to some degree. Larger fractions of exempted poor in larger towns should depress IRS, while greater fractions of untaxed ecclesiastical wealth would increase IRS. For this reason, we think that our estimates of IRS for early Tudor towns are accurate given the existing evidence.

S2 Worst Case Scenarios (WCSs) Overview

The most problematic potential source of error comes from the number of people not counted by the taxes. The important quantity is the proportion of town populations that were not counted in the 1524/5 taxpayer rolls. This can be understood as the number of uncounted people per taxpayer, or the ratio of taxpayers to non-taxpayers., Historians multiply the taxpayer counts by constant 'modifiers' to estimate the non-taxpayer population. These modifiers are derived from other available demographic data, and range from 4 to 9 depending on the specific data, methods, and assumptions of the author (see e.g. Dyer, 2000b; Goose & Hinde, 2007; Rigby, 2010). Embedded in the use of constant multipliers are the tacit assumptions that

1. Errors are random and unstructured
2. The proportion of unenumerated persons does not covary with town population

If these assumptions are valid, such that taxpayer counts are roughly proportional to population (with some degree of unstructured variability), then only the taxpayer counts are necessary to conduct scaling analysis.

However, it has been previously suggested that larger towns may have had greater proportions of exempted taxpayers due to higher poverty rates (Cornwall, 1962b; Rigby, 2010). If so, then the undercounting of taxpayers in larger towns—due to the exemption of people that would otherwise be counted as taxpayers and used to estimate population—would cause the systematic underestimation of the population of larger towns. Indeed, widespread urban poverty, unemployment, and vagrancy are well attested in the contemporary records of larger towns, which seems to have been the result of rural poor gravitating to urban centers in search of work (see e.g. Cornwall, 1988; Dobson, 1977; Palliser, 1992; Phythian-Adams, 1978, 1979, Pound, 1971, 1986). In this case, simply using raw taxpayer counts (or estimating population with a constant multiplier) would systematically bias log-linear regression results by increasing their superlinearity (i.e. increasing the slope of the OLS line by increasing β and decreasing α). This feature of the town Lay Subsidy data has not previously been addressed by historians in any systematic way.

To test this possibility, we employ several alternative methods to model "worst case scenarios" (WCSs) in the following sections. In these WCSs, town populations are modeled such that (1) the proportion of exempt taxpayers directly covaries with the number of 1524/5 taxpayers, and (2) the magnitude of this covariation is modeled to the greatest extent possible from the existing evidence. Then, in section S6, each of these WCSs datasets are evaluated (alongside the four error-type subsets) to test the sensitivity of increasing returns to scale to these extreme methodological assumptions.

S3 WCS1: Maximal Modifier Rages

The simplest and most straightforward way of modeling WCSs is to use estimated population modifier ranges from the published literature. The set of WCSs that use this method are collectively referred to as 'WCS1'. Modifier ranges are often intended as uncertainty ranges for individual towns, although some are estimates of modifier variability among all towns (Table S2). As seen in Table S2, modifier ranges estimated by historians are all ≤ 2, the specific magnitudes of which are constrained by methodological assumptions. In general, lower ranges (e.g. 4-5) interpret taxpayers as adult males, while higher ranges (e.g. 6-7) interpret taxpayers as households (see Dyer, 2000b; Goose & Hinde, 2006, 2007; Rigby, 2010). By extending each of the estimated ranges in Table S2 to an "exaggerated range" of 2, their differential exemption proportions will be much greater than suggested in the literature. This magnitude is quantified by the b/a ratio (Table S2), which indicates proportionally how many times larger the largest town's multiplier is than that of the smallest town.

Table S2. Estimated 1524/5 Lay Subsidy Taxpayer-Population Multiplier Ranges

Source	Estimated Ranges [*a*, *b*]	*b*/*a* ratio	Intended Scope	Extended Ranges [*a*, *b*]	*b*/*a* ratio
Goose (1984, pp. 242–52)	[4.4, 5]	1.136	All Towns	[4, 6]	1.5
Keene (1985, p. 368)	[5.3, 6.7]	1.264	Winchester	[5, 7]	1.4
Dyer (2000b, p. 271)	[5.5, 7.5]	1.364	All Towns	[5.5, 7.5]	1.364
Britnell (1986, p. 201)	[6, 8]	1.333	Colchester	[6, 8]	1.333
Pound (1976, p. 129)	[7.7, 9]	1.16	Norwich	[7, 9]	1.285

It should be noted that the vast majority of estimated modifiers in the literature are single point estimates (see Rigby, 2010), and our exaggerated ranges used in WCS1 are greater than or equal to those that have been estimated by scholars. The original "estimated ranges" in Table S2 were also intended as ranges of unstructured variability—such that a small town might have a high modifier, while a large town might have a small modifier. By making these ranges directly covary with the number of taxpayers, our WCS1 estimates are considerably more extreme than those estimated in the literature. Furthermore, these direct, linear covariance models are especially pessimistic because reality was almost certainly much noisier—as opposed to the perfect 1-to-1 relationship between taxpayers and the magnitude of the population modifier implied by WCS1. As such, our method is particularly appropriate for WCSs.

To model town populations from the WCS1 modifier ranges, we can linearly rescale the extended modifier ranges from Table S2 to fit the range of taxpayers in the dataset [151, 1423] using the equation

$$M_i = \frac{(b-a)(T_i - \text{min})}{\text{max} - \text{min}} + a \qquad \text{(S1)}$$

where M_i is the multiplier for town i, T_i is its number of taxpayers, a is the minimum and b is the maximum of the modifier range $[a, b]$, and 'min' and 'max' are the minimum and maximum number of taxpayers in the sample (151 and 1423). The product of M_i and T_i is thus the modeled population of town i. Taking the modifier range [6, 8] in Table S2 as an example, the smallest town will have a population of 906 (= 6 × 151) and the largest town will have a population of 11,384 (= 8 × 1423). The population estimates produced by the WCS1 modifier ranges in Table S2 are available in the file "WCS.csv", and their calculation is replicated in the adjoining "Data_Replication_File_Cesaretti.r" script.

S4 Proxy Data WCS Models

Alternatively, we can use available demographic data from Early Modern English towns to estimate the proportion of would-be taxpayers who were exempted in the 1524/5 Lay Subsidies. These exemption proportions can subsequently be turned into WCS modifier ranges using estimates of mean household size, average adult sex ratios, and average age structures. In theory, the taxpayer rolls included all men over the age of 15, as well as female heads of household, whose assessed wealth was worth £1 or more (Schofield, 2004, pp. 102–10). Historians have commonly assumed average taxpayer exemption rates ranging from 10% to 30%, the variability of which has generally been assumed to be unstructured (Campbell, 1981; Goose & Hinde, 2007; Patten, 1978, p. 101; Wrigley & Schofield, 1981, p. 567). Only in exceptional cases like Coventry and York are larger exemption proportions used (Palliser, 1979; Phythian-Adams, 1979), and Goose has even argued that 30% is excessive for most towns (Goose & Hinde, 2007).

S4.1 Taxpayer Exemption Rates

Two rough proxies are available for estimating the proportion of exempted taxpayers in Tudor towns. First, we can use surviving data from the 1522 Muster Rolls, which counted the number of persons assessed at under £1. As seen in Table S3, the proportion of taxpayers under £1 in 1522 extends considerably beyond the 10-30% range, especially for larger towns, and is therefore ideal for our WCS. Second, detailed studies of the 1524/5 taxpayer rolls indicate that a considerable number of poor wage laborers (assessed at or just above the £1 exemption threshold) were excluded in either 1524 or 1525—but not taxed in both years like their wealthier counterparts (Cornwall, 1962a, 1988; Goose, 1984, p. 60). We can therefore also use the total number of taxpayers that were not named in both 1524 and 1525 as an additional proxy for the 1524/5 exemption rate. These two proxies are not identical, they come from different years, and the 1522 methods were also different. Nevertheless, they both measure the same category of persons and produce rather extreme numbers. They are therefore useful proxies for modeling our WCSs.

Table S3. Proxies for Exemption Proportions in English towns, 1522-1525

Town	Taxpayers 1524/5*	Estimated % Exempt	Proxy	Source
Beaconsfield†	87	19.6%	% under £1 in 1522	Cornwall (1962a, p. 60, 64)
Oakham†	104	39.8%	% under £1 in 1522	Cornwall (1962a, p. 60, 64)
Stony Stratford†	106	32.9%	% under £1 in 1522	Cornwall (1962a, p. 60, 64)
Buckingham†	111	11%	% under £1 in 1522	Cornwall (1962a, p. 60, 64)
Amersham†	123	20.5%	% under £1 in 1522	Cornwall (1962a, p. 60, 64)
High Wycombe	175	26.2%	% under £1 in 1522	Cornwall (1988, p. 58)
Lavenham	199	14.6%	% under £1 in 1522	Cornwall (1988, p. 53)
Aylesbury	202	20.3%	% under £1 in 1522	Cornwall (1962a, p. 60, 64)
Sudbury	231	20.4%	% under £1 in 1522	Cornwall (1988)
Walsingham	232	32.7%	% under £1 in 1522	Cornwall (1988, p. 59)
Newbury	414	25%	% under £1 in 1522	Cornwall (1988, p. 59)
Leicester	427	33.3%	% under £1 in 1522	Hoskins (1956, p. 17)
Yarmouth	497	36.3%	% under £1 in 1522	Cornwall (1988, p. 259)
Cambridge	524	37%	% undercounted 1524/5	Goose (1984, p. 60)
Reading	531	44%	% undercounted 1524/5	Goose (1984, p. 60)
Colchester	701	27%	% undercounted 1524/5	Goose (1984, p. 60)
Coventry	725	54.4%	% under £1 in 1522	Cornwall (1988, p. 259)
Exeter	1019	48%	% under £1 in 1522	Cornwall (1988, p. 267)

* Taxpayer figures from Table S1 dataset, from Rigby (2010), A. Dyer (2000a), and Sheail (1968), as well as Cornwall (1962a, p. 60, 64) for towns $\leq$ 150 taxpayers
† These towns are below the taxpayer count threshold for our sample ($\geq$ 150), but provide a useful frame of reference for possible exemption proportions as they extend below our sample.

To be sure, the sample in Table S3 is very small (n = 18) and noisy, making the application of parametric statistical models invalid. Moderate positive correlation between the number of taxpayers 1524/5 and the exemption proxies (R^2 = 0.46) suggests that these are broadly greater in larger towns than in smaller towns, with a very high range of variability. We also know from detailed historical studies that several of the towns from Table S3 have unrepresentative proxy data. Coventry (54.4%) was then undergoing an exceptional recession, with exceptionally high proportions of poor and transients (Phythian-Adams, 1979). Lavenham (14.6%) was a small, unincorporated cloth manufacturing town at the nexus of a booming cottage industry and wealthy textile merchants (Britnell, 2000; Cornwall, 1988, pp. 63, 73, 102). Buckingham's (11%) percent undercounted 1524/5 was 36%, suggesting that the 11% under £1 from the 1522 muster returns is not representative (Cornwall, 1962a, p. 60). Thus, it would appear that those cases with exempt proportions >50% and <20% should be treated as outliers.

For the purposes of constructing a WCS population model, we can therefore linearly extrapolate this 20-50% taxpayer exemption range onto their corresponding range of 1524/5 taxpayers. Since 50% exemption is reached at approximately the 1000 taxpayer level, Coventry is an outlier, and we do not have data for Bristol and Norwich, we apply the 50% exemption proportion to all towns with over 1000 taxpayers. As such, we use the linear rescaling equation (Eq. S1) from above with an [a, b] range of [1.25, 2], and a min-max range of [151, 1000+], which leads to a b/a ratio of 1.6.[3] This narrowing of the taxpayer range from [151, 1423] to [151, 1000+] will further reduce the superlinearity of estimated log-linear scaling coefficients, making it particularly appropriate as a WCS assumption. Perhaps more importantly, this range is already considerably greater than any estimated ranges in Table S3, as well as being a perfect linear relationship.

However, even if we assume that the exemption model estimated here is useful for WCSs, we still need to convert the number of taxpayers into absolute population numbers for them to be compared to population estimates (e.g. the maximal modifier ranges of WCS1). This can be done in two different ways using alternative assumptions about what taxpayers represent demographically—either adult males or households. First, figures on adult males or households must be converted into population numbers using fragmentary historical data on either mean household (MHS) or age ratios and sex ratios. Then, we must also account for how demographic statistics (MHS or age/sex ratios) should also covary with the exempt population due to poverty, given that exempted taxpayers are by-definition very poor.

S4.2 WCS2: Taxpayers as Households

The first method assumes that taxpayers represent households, which is here referred to as WCS2. In general, the mean household size (or 'MHS') of urban households in Tudor towns is thought to have ranged from about 4-4.75 (see Dyer, 2000c; Patten, 1978: 100-1; cf. Goose & Hinde, 2006, 2007). It is well documented that poor households were much smaller on average than those of better-off and wealthy households due to greater numbers of children, domestic servants, and apprentices (Cornwall, 1988; Palliser, 1992; Phythian-Adams, 1979). While the statistical data are biased towards the 17th and 18th centuries, statistical changes in MHS increased over this period remain within the range 4-5. The data are relatively scarce, but they are corroborated by a wealth of qualitative data that indicates continuity of the Early Modern pattern in Tudor sources (see Laslett, 1972; Palliser, 1992; Pound, 1971; Wall, 1972). In Laslett's (1972: 154) sample of 100 English communities from 1574 to 1821, mean household size ('MHS') varied directly with social status: 3.74 for Paupers, 4.34 for Laborers, 4.72 Tradesmen and craftsmen, 5.39 for Husbandmen, and 7.54 for Gentlemen. At Coventry in 1523, MHS directly covaried with household wealth, ranging from 2.8 among Nil assessments to 11.8 in households assessed at >£100 (Phythian-Adams, 1979: 241).[4] Goose's (1980) analysis of Cambridge parishes in the 1619-32 finds clear positive correlation between occupational status and household size, ranging from 3.65 for laborers, to 4.25-4.9 for various craftsmen, to 5.36 for prof/office, to 8.63 for gentlemen.

Thus, if the proportion of impoverished households (with assessments under £1) directly covaries with both MHS and population size, then MHS will covary with the population of towns. To do so, we use the MHS of Coventry (3.83) and England as a whole (4.75) as the possible MHS range, and make MHS directly covary with the proportion of exempt taxpayers in 1524/5. This yields the population equation

$$N_i = T_i M_i H_M$$

[3] For a = 20%, the non-exempt rate is (1-0.2) = 0.8, and 1/0.8 = 1.25 is thus the total taxpayer multiplier. For b = 50%, the non-exempt rate is (1- 0.5) = 0.5, and 1/0.5 = 2 is thus the total taxpayer multiplier. Thus, a = 1.25; b = 2.
[4] The same is true for rental categories, ranging from 2.6 to 9.2 (*ibid*., p. 239).

where N_i is the population of town '*i*', M_i is the exemption multiplier for town *i*, T_i is its number of taxpayers, and H_M is MHS given the exemption multiplier for town *i*. Linearly extrapolating the MHS range of [3.83, 4.75] to raw exemption modifier range [1.25, 2] from above produces a final WCS2 modifier range of [5.94, 7.66], a range of 1.72, and a *b*/*a* ratio of 1.29.[5]

These modeled population estimates comprise one of the datasets in WCS2. The other WCS2 dataset also uses these modeled population estimates, but instead uses the exemption proportion proxy data from Table S3 for those towns with exemption estimates in Table S3 (n = 15). This second dataset in WCS2, that includes real proxy exemption proportion data, is referred to as "WCS2 honoring the data" in the following analyses. The population estimates produced by the WCS2 modifier range, and the WCS2 model honoring the data, are available the file "WCS.csv", and their calculation is replicated in the adjoining "Data_Replication_File_Cesaretti.r" script.

S4.3 WCS3: Taxpayers as Adult Males

Our second method—that we call WCS3—assumes that taxpayers represent adult males. For the purposes of extrapolating the 1524/5 Lay Subsidy returns into population figures, the adult sex ratios of Tudor towns have been assumed to be somewhere around 100,[6] and average age structures have been assumed to be around 35-40% children (e.g. Goose, 1980; Goose & Hinde, 2006, 2007). However, as with MHS, there is substantial evidence that both adult sex ratios and the proportion of children were systematically smaller in towns with larger populations. Most servants were young women, and servants gravitated towards to larger towns in search of work—leading to greater proportions of large town populations being unmarried women that did not contribute to the fertility rate (see Clark & Souden eds., 1987; Galley, 1998: 11–30, 116–49; Pound, 1971, 1986; Souden, 1984). While reported rural servant adult sex ratios range from 75 to 201 (Laslett, 1972, p. 152; Souden, 1984, p. 150), these ranged from 53 to 85 in the towns of sixteenth and seventeenth century England (Galley, 1998: 22–5; Phythian-Adams, 1979: 221; Souden, 1984: 150). Similarly, the sex ratios of towns in the late seventeenth century averaged 90.4 for small towns and 83.4 in large towns, compared to 91-109 in the overall population of England (Laslett, 1972; Souden, 1984). Due to great abundance of adult women servants, *adult* sex ratios (the metric necessary for extrapolating town populations from taxpayer counts, rather than total sex ratios) were often much lower than overall sex ratios. For example, the adult sex ratio was 72 in both 1523 Coventry and 1696 Lichfield (Galley, 1998: 22–5; Phythian-Adams, 1979: 200; Souden, 1984). With respect to the age structure, Wrigley and Schofield (1981: 528) estimate 36.2% of the entire English population was under 15 in 1550, and Laslett's (1972: 148) mostly rural sample averages 41.8% children—both of which are considerably higher than the 26.8% at Coventry in 1523 (Phythian-Adams, 1979: 221). Poorer households had consistently fewer children on average (Goose, 1980; Laslett, 1972; Phythian-Adams, 1979), and larger towns had greater proportions of unmarried/childless servants (Phythian-Adams, 1979; Souden, 1984).

Thus, these sixteenth and seventeenth century trends indicate that age structure and adult sex ratios also directly covaried with both town population and the proportion of exempt persons. Here we can use the adult sex ratios and percent children under 16 from Coventry (72:100,

---

[5] (1.25 exemption multiplier) * (4.75 MHS) = 5.94; (2 exemption multiplier) * (3.83 MHS) = 7.66

[6] i.e. 100 males per 100 females (or, 50% male)

26.8%) as the values for the largest town (Phythian-Adams, 1979: 200, 221). For the smallest town, we can use the average sex ratio of 1696 small towns (90:100; Souden, 1984) and the percent children of Laslett's (1972:148) sample of towns (~40%). To estimate the population of Tudor towns using these ranges, we can use the equation

$$N_i = T_i M_i S_M C_M$$

where N_i is the population of town 'i', M_i is the exemption multiplier for town i, T_i is its number of taxpayers, S_M is the sex ratio multiplier (a linear function of the exemption multiplier), and C_M is the child proportion multiplier (a linear function of the exemption multiplier). The sex ratio multiplier is given by the aggregate divided by the male percent (S_M = 100% / %M = 100% / (M_frac / (M_frac + F_frac))), and the child proportion multiplier is given by the aggregate divided by the adult fraction (C_M = 100/(100-%Children)). For the largest towns we arrive at S_M = 1/0.418 = 2.39 and C_M = 1/(1-0.268) = 1.37. For the smallest towns we arrive at S_M = 1/0.473 = 2.18 and C_M = 1/(1-0.4) = 1.66. This produces a modifier range of [4.52, 6.55], a range of 2.03, and a *b*/*a* ratio of 1.45.[7]

These modeled population estimates comprise one of the datasets in WCS3. As for WCS2, the other dataset in WCS3 uses the exemption proportion proxy data from Table S3 for those towns with exemption estimates in Table S3 (*n* = 15). This second dataset in WCS3, that includes real proxy exemption proportion data, is referred to as "WCS3 honoring the data" in the following analyses. The population estimates produced by the WCS3 modifier range, and the WCS3 model honoring the data, are available in the file "WCS.csv", and their calculation is replicated in the adjoining "Data_Replication_File_Cesaretti.r" script.

S5 Summary of WCSs

Table S4 presents the nine (*n* = 9) individual WSC datasets. These 9 unique WSC datasets are categorized into their 3 WSC groups (separated by horizontal black lines in Table S4). WCS1 includes the 5 extended modifier ranges from section S3 and Table S2. WCS2 are the proxy exemption proportion estimates using mean household size as the basis of population estimation from section S4.2 WCS3 are the proxy exemption proportion estimates using sex and age ratio as the basis of population estimation from section S4.3. WCS2 and WCS3 also include datasets that "honor" the exemption proportion proxy data from Table S3 for their respective towns (*n* = 15). By 'honoring' the available proxy exemption data from Table S3, we thereby include outliers with exemption proportions >50% and <20%—which extends the modifier range, and therefore further exaggerates the WCS.

Table S4. Summary of Methods for all Worst Case Scenarios (WCSs)

Worst Case Scenario Group	Modifier Range	Range (*b*/*a* ratio)	Exempt Range	Extrapolation Range*	Method
	[4, 6]	2 (1.5)	*NA*	[151, 1423]	*NA*
WCS1	[5, 7]	2 (1.4)	*NA*	[151, 1423]	*NA*
Extended Modifier	[5.5, 7.5]	2 (1.36)	*NA*	[151, 1423]	*NA*
Ranges	[6, 8]	2 (1.33)	*NA*	[151, 1423]	*NA*

[7] (1.25 exemption multiplier) * (2.18 adult sex ratio multiplier) * (1.66 percent children multipier) = 4.52;
(2 exemption multiplier) * (2.39 adult sex ratio multiplier) * (1.37 percent children multipier) = 6.55

	[7, 9]	2 (1.28)	*NA*	[151, 1423]	*NA*
WCS2 Proxy Exemption	[5.94, 7.66]	1.72 (1.29)	[0.8, 0.5]	[151, 1000+]	Avg. households (3.83-4.75)
Household Taxpayers	[5.5, 9.05] ‡	3.55 (1.65)	[0.8, 0.5]	[151, 1000+]	Honor the data
WCS3 Proxy Exemption	[4.52, 6.55]	2.03 (1.45)	[0.8, 0.5]	[151, 1000+]	sex ratio† (72-90) children (27-40%)
Adult Male Taxpayers	[4.22, 7.42] ‡	3.2 (1.76)	[0.8, 0.5]	[151, 1000+]	Honor the data

† *adult* sex ratio

* Note that "1000+" in the range [151, 1000+] indicates that towns with greater than 1000 taxpayers have a 50% taxpayer exemption rate. See section S4.1 for details.

‡ These modifier ranges are the maximum and minimum modifiers when the data from Table S3 is honored, which correspond to Coventry (54.4% exempt; 725 taxpayers) and Lavenham (24.6% exempt; 199 taxpayers) rather than Norwich (50% exempt; 1423 taxpayers) and Louth (20% exempt; 151 taxpayers).

The WCSs modeled here should be a definitive test of the settlement scaling hypothesis for several reasons. First, it must again be noted that all prior analyses of the 1524/5 Lay subsidies have used taxpayers or constant modifiers as a proxy for population. As such, the original taxpayer data itself may be taken as a valid test of settlement scaling theory. Second, most estimated modifiers in the literature are single point estimates, many are intended as margins of error, and none of them have implemented direct modifier covariation with taxpayer number. Third, the exaggerated ranges used here are greater than or equal to those that have been estimated by scholars. Fourth, the published modifier ranges were intended to have internal variability—such that a small town might have a higher modifier, while a large town might have a lower modifier. By making these ranges directly covary with the number of taxpayers, our estimates are substantially more extreme than any population estimates in the literature. Likewise, proxy-estimated exemption proportions are also made to directly covary with taxpayers, even though our small sample of proxy data on exemption proportions exhibits a very wide range of variability. For example, proxy exemption proportions of small towns range from 11-40%, and large towns range from 27-54% (Table S3). The resulting direct, linear covariance models are especially pessimistic because reality was almost certainly much noisier—as opposed to the perfect 1-to-1 relationship between taxpayers and the magnitude of the population modifiers implied by WCSs. Finally, we used every method yet employed or suggested in the literature to estimate the population of Tudor English towns from the 1524/5 subsidies, including modifier ranges, household sizes, age structure, and sex ratios, using detailed historical data.

Together, this means that the WCS datasets used in our sensitivity analysis deliberately make outlandish assumptions with respect to the accepted mainstream methods of historical demography. This was done with the specific intention of testing the robustness of settlement scaling theory. The WCS Tudor town populations were modeled in such a way as to intentionally destroy the superlinearity of the estimated log-linear regression scaling (β) coefficients. This was accomplished by making the modeled population of towns directly and linearly covary with the number of 1524/5 taxpayers. The magnitude of this covariation was modeled to the largest extent possible from the existing evidence, and in many cases it extends considerably beyond realm of possibility.

S6 Sensitivity Analysis

S6.1 Implementation

Twelve (n = 12) different population datasets are evaluated in the following sensitivity analysis. These include the nine unique WSC datasets (n = 9) from Table S4, explained in sections S3 and A4. These include the modifier ranges in WCS1 (n = 5), the exemption proportion household-based estimates of WCS2 (n = 2), and the exemption proportion adult male-based estimates of WCS3 (n = 2). In addition to these, two population datasets based on Dyer's (2000b) population estimates (n = 2) and the original taxpayer dataset (n = 1) are added for the sake of comparison. Each of these 12 datasets is further evaluated for each of the four (n = 4) error-type subsets in Table S1, resulting in a total of 48 separate regressions (12 × 4 = 48). These 12 datasets are summarized in Table S5.

Table S5. The Twelve Datasets Analyzed in this Sensitivity Analysis

Number	Analysis	Group	Modifier Range	Estimation Basis
1	A	Taxpayers	NA	Raw Data
2	B	Dyer (2000c)	NA	Pop Estimates
3	C	Dyer (2000c) Extended	NA	Pop Estimates
4	D	WCS1	[4, 6]	Modifier Ranges
5	E	WCS1	[5, 7]	Modifier Ranges
6	F	WCS1	[5.5, 7.5]	Modifier Ranges
7	G	WCS1	[6, 8]	Modifier Ranges
8	H	WCS1	[7, 9]	Modifier Ranges
9	I	WCS2	[5.94, 7.66]	Households
10	J	WCS2	[5.5, 9.05]	Households
11	K	WCS3	[4.52, 6.55]	Adult Males
12	L	WCS3	[4.22, 7.42]	Adult Males

The population estimates of Alan Dyer (2000b) use a basal modifier of 6, but are 'corrected' case-by-case using contextual information and assumed retrodictions from both the later Chantry Certificates of 1548 and the Bishop's Census of 1563. Dyer's (2000b) datatset excludes numerous small towns, and therefore only provides 1524/5 estimates for 63 of the towns in our total dataset (n = 93). Therefore, in addition to Dyer's (2000b) original dataset (n = 63), we also create another dataset using his modifier of 6x to extend his dataset to our remaining 30 small towns. In this way, we evaluate two datasets based on Dyer's (2000b) estimates—one that strictly includes only his published estimates, and the other with 30 added small towns added to his published estimates. Both of these Dyer (2000b) datasets are evaluated for all 4 error-type subsets. Finally, the original taxpayer dataset (evaluated in section 5 of the main text) is analyzed again here to facilitate its comparison with the Dyer datasets and the WCSs.

The scaling model is implemented just as in section 5 of the main text. Scaling exponents and prefactors were estimated through log-linear OLS regression of the two variables (Eq. 11 in the main text). Parameters were estimated in R using the 'sandwich' package (R Core Team, 2016; Zeileis, 2004). As in the main text, the parameter of greatest interest is β. Estimated β coefficients > 1 indicate increasing returns to scale, while estimated β coefficients ≈ 7/6 indicate increasing returns to scale predicted by settlement scaling theory (see section 3 of the main text).

S6.2 Results

Table S6 provides a statistical summary of all WCS regressions ($n = 36$) performed in this sensitivity analysis. It is clear that the estimated parameters, correlation coefficients, and confidence intervals are each tightly clustered. In addition to being estimated with controls on heteroskedasticity (Zeileis, 2004), the log-linear models explained between 68% and 76% of the variability in the data, indicating that they are not mis-specified. More importantly, the estimated WCS scaling coefficients were superlinear ($\beta > 1$) in all 36 cases—in spite of our deliberate attempt to eliminate their superlinearity. The estimated WCS scaling coefficients range from 1.06 to 1.17, with a mean of 1.12. The distribution of estimated scaling coefficients is shown in Figure S5, which is basically unimodal, two-tailed, and symmetrical. The black-shaded histogram bars denote the β estimates of the 'no rural, no problem' error type subset ($n = 55$), which should provide the most reliable, or 'least-error-prone', estimates. This 'no rural, no problem' error-type subset produced a notably higher mean of 1.135, and a range of 1.09 to 1.17.

Table S6. Statistical Summary of all WCS Regressions (*n* = 36). This involves WCSs 1-3 and Analyses D-L in Table S5. The individual summarized regressions are No. 13-48 in Table S9.

	R^2	β	α	β 95% C.I.	
				Lower	**Upper**
Mean	0.71	1.12	0.019	0.96	1.27
St.Dev.	0.03	0.03	0.007	0.03	0.04
Min	0.68	1.06	0.009	0.91	1.20
Q1	0.68	1.10	0.013	0.95	1.25
Median	0.70	1.12	0.017	0.96	1.27
Q3	0.74	1.15	0.025	0.98	1.30
Max	0.76	1.17	0.033	1.02	1.34

Interestingly, our WCS β distribution in Figure S5—which includes the most pessimistic assumptions possible—overlaps with the settlement scaling theory prediction of 7/6 (1.167). Moreover, this distribution also overlaps with the majority of the IRS scaling coefficients empirically-observed for the socioeconomic rates of modern industrialized cities (e.g. GDP, income, and wages in Europe, the United States, and Japan), which have ranged from 1.08-1.17 (see Bettencourt, 2013: supplement; Bettencourt & Lobo, 2016). It is notable that three-quarters of the WCS β estimates are at least 1.10, indicating solid IRS superlinearity, and only 14% of the WCS β estimates (5/36) were less than 1.08. We take this as a strong indication that the finding of IRS for Tudor English towns is very robust to a wide range of extreme demographic assumptions about the 1524/5 Lay Subsidies.

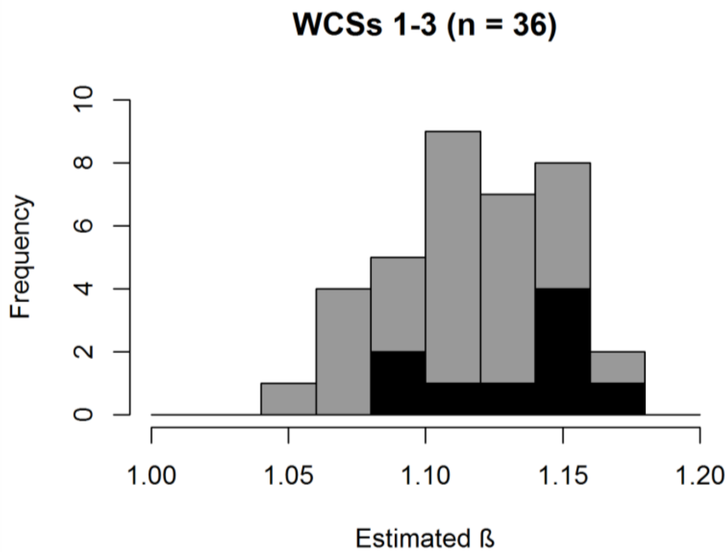

Figure S5. Sensitivity Analysis Histogram for all WCS *β* coefficients (*n* = 36). Black-shaded histogram bars are those *β* estimates corresponding to the "no rural, no problems" error-type subset.

The error-type subsets had a noticeable impact on both the correlation coefficients and estimated parameter values. This effect is illustrated in Table S7, which summarizes the mean impacts of the error-type subsets on all robustness analysis datasets. The error type-subsets have the same average effects as in in the case of the raw taxpayer dataset (Table 1 of the main text)—such that the highest β and R^2 values are for the subset "without rural or problematic" cases, followed closely by the "all cases" subset. The "municipal/rural" and "problematic" error-type subsets performed about equally on average. We suspect that the "without rural or problematic" performs the best because numerous error-prone cases and outliers are shed from the data (see section S1.2). Importantly, we suspect that the high performance of the "all cases" subset is caused by the greater abundance of data, which strongly suggests that the superlinear trend in the 1524/5 Lay Subsidy data is genuine.

Table S7. Results summarized by Error Type Subset

Error Type Subset	*n**	Mean R^2	Mean β Estimate	Mean β 95% C.I.**	Mean α Estimate
All cases	12	0.67	1.15	[0.99, 1.31]	0.018
Without Municipal/Rural	12	0.68	1.13	[0.96, 1.30]	0.022
Without Problem Cases	12	0.72	1.13	[0.97, 1.29]	0.023
Without Rural or Problem Cases	12	0.75	1.17	[0.99, 1.34]	0.018

* This is the number of individual datasets evaluated, not individual towns
** Calculated by separately taking the means of the upper and lower 95% C.I. limits estimated for each dataset.

Table S8 summarizes the results of the sensitivity analysis using the mean estimated correlation coefficient and parameter values for each of the twelve (*n* = 12) individual datasets. In other words, the regression results of all four (*n* = 4) error-type subsets are averaged in Table S8 for each of the twelve (*n* = 12) datasets in Table S5 (above). Because of the small range of variability of the estimated parameters and correlation coefficiets in each of the twelve (n = 12) datasets, these dataset averages closely reflect the results of the individual regressions. The results of all individual regressions (n = 48) are presented in Table S9, at the end of the appendix, for interested readers. All data and scripts used in analysis are available in the Supplementary Materials file "WCS.csv" and "Data_Replication_File_Cesaretti.r" script.

Table S8. Regression Results Summarized by Dataset

No.	Dataset	Sensitivity Analysis Group	n*	Mean R^2	Mean β Estimate	Mean β 95% C.I.**	Mean α
1	A	Taxpayers	4	0.71	1.26	[1.08, 1.44]	0.055
2	B	Dyer (2000c)	4	0.67	1.23	[1.02, 1.44]	0.007
3	C	Dyer (2000c) Ext	4	0.67	1.19	[1.00, 1.39]	0.009
4	D	WCS1 [4, 6]	4	0.71	1.10	[0.94, 1.24]	0.029
5	E	WCS1 [5, 7]	4	0.71	1.12	[0.97, 1.28]	0.019
6	F	WCS1 [5.5, 7.5]	4	0.71	1.13	[0.97, 1.29]	0.016
7	G	WCS1 [6, 8]	4	0.71	1.14	[0.98, 1.30]	0.014
8	H	WCS1 [7, 9]	4	0.71	1.15	[0.99, 1.32]	0.010
9	I	WCS2 [5.94, 7.66]	4	0.71	1.13	[0.98, 1.29]	0.014
10	J	WCS2 [5.5, 9.05]	4	0.71	1.09	[0.93, 1.23]	0.027
11	K	WCS3 [4.52, 6.55]	4	0.72	1.12	[0.97, 1.28]	0.016
12	L	WCS3 [4.22, 7.42]	4	0.72	1.08	[0.93, 1.22]	0.029

* This is the number of individual datasets evaluated, not individual towns
** Calculated by separately taking the means of the upper and lower 95% C.I. limits estimated for each dataset.

The high performance of the contextually-modified expert population estimates of Dyer (2000b) provide further evidence for the accuracy and validity of the IRS scaling results. Like the raw taxpayer data, the β estimates for Dyer (2000b) datasets B and C are strongly superlinear, with averages ranging from 1.19 to 1.23 (Table S8). We suspect that these results are the most accurate, valid, and representative results for Tudor towns. The 95% confidence intervals of the Dyer datasets are also almost entirely superlinear, with the lower limit hovering around 1.0. Despite this, the Dyer (2000b) estimates had lower R^2 values than any other groups, which was caused by Dyer's modified population estimates for certain larger towns. However, the scaling result was only slightly affected by these contextual modifications. This further suggests that the strong IRS exhibited by the 1524/5 Lay Subsidy data are robust to random perturbations in the values of the small number of large towns (to which OLS is most sensitive).

In general, Table S8 shows that WCS1 and WCS2 mostly produced similar results, while WCS3 produced the lowest β estimates and C.I.s. However, this was not caused by the specific methods of WCS3, which assumed taxpayers to be adult males, and used age and sex ratios to estimate modifier ranges. More careful analysis of Table S9 reveals that Dataset D of WCS1 and both Datasets K and L of WCS3 produced the lowest β estimates and C.I.s. This resulted from their higher b/a ratios, which is a product of their lower modifier 'a' values (i.e. the lower bound of the [a, b] modifier range was between 4 and 5]. In general, this b/a ratio explains about half of the variability in the estimated β coefficients. *This was directly caused by the direct linear extrapolation of modifier ranges to taxpayer numbers*, such that higher b/a ratios inflated the relative disparity between larger and smaller towns. In order to destroy the superlinearity of the estimated beta coefficient, the modifier range b/a ratio of this *direct linear extrapolation* would need to be about 1.8. This is far beyond the modifier ranges estimated by historical scholars in Table S2, the b/a ratios of which hover between 1.13 and 1.36. Most importantly, the ranges in Table S2 were not intended to be linearly extrapolated to taxpayers as done for our WCS, and most expectations for taxpayer exemption rates range from only 10-30% (Goose & Hinde, 2006, 2007). As such, we argue that the lowest possible scaling coefficients for the towns of Tudor England are those produced by Dataset F of WCS1, with a *b*/*a* ratio of 1.36, and estimated β ranging from 1.11 to 1.15. We take this to be the absolute lowest range possible given the available evidence, such that the most probable values range from 1.11 to 1.28 (the latter being maximum value estimated using the raw taxpayer data; No. 4 in Table S9). Again, we suspect that the Dyer population estimates produce the most accurate results, which reinforces our conclusion that the most probable values of β are above the range produced by Dataset F ($1.15 < \beta < 1.28$).

Table S8 also illustrates that the estimated β 95% confidence intervals are quite high for all WCS, and both their range and variability are constant across analyses (see also Table S9). As the estimated β coefficients are gradually pushed downwards by higher *b*/*a* ratios, the lower bound of their 95% C.I.s are gradually pushed downwards by the same magnitude. This direct proportionality indicates that all β 95% C.I.s with lower bounds less than 1.0 are caused the over-inflation of unrealistic WCS b/a ratios. Even Dataset F—our lowest possible result—has an average β 95% C.I. of [0.97, 1.29], which is just barely lower than 1.0. Given the large size and high quality of the 1524/5 Lay Subsidy sample data, this strongly suggests that that the probability of a linear or sublinear scaling coefficient for Tudor towns is very low.

Taken together, our sensitivity analysis was not able to falsify the increasing returns to scale hypothesis for the provincial towns of Tudor England. Analyses of taxpayers and the estimates of Dyer (2000b) both suggest strong superlinearity greater than the theoretical expectation of 7/6. Our WCSs—intentionally created to exaggerate the covariation between taxpayer exemption proportions and town populations—failed to eliminate the increasing returns to scale exhibited by the 1524/5 Lay Subsidy data. Close scrutiny of the WCS results indicate solid superlinearity, with possible values ranging $1.11 < \beta < 1.28$, and most probable values of $1.17 \leq \beta \leq 1.26$ (the average Dyer dataset range). All of these values are consistent with those of modern cities (Bettencourt, 2013; Bettencourt & Lobo, 2016).

Table S9. Full Regression Results of the Sensitivity Analysis

No.	**Anal**	**Group**	**Subset**	**n**	**R^2**	**β**	**β 95% C.I.**	**α**
1	A	Taxpayers	all	93	0.68	1.27	[1.10, 1.44]	0.051
2	A	Taxpayers	no rural	68	0.68	1.25	[1.07, 1.44]	0.056
3	A	Taxpayers	no problem	77	0.72	1.23	[1.05, 1.41]	0.065
4	A	Taxpayers	no rural, no problem	55	0.74	1.28	[1.09, 1.48]	0.047
5	B	Dyer (2000c)	all	67	0.60	1.21	[1.00, 1.42]	0.007
6	B	Dyer (2000c)	no rural	62	0.64	1.19	[0.98, 1.40]	0.009
7	B	Dyer (2000c)	no problem	54	0.72	1.26	[1.06, 1.47]	0.005
8	B	Dyer (2000c)	no rural, no problem	50	0.73	1.26	[1.06, 1.46]	0.005
9	C	Dyer (2000c) Ext	all	93	0.62	1.17	[0.99, 1.35]	0.010
10	C	Dyer (2000c) Ext	no rural	68	0.64	1.19	[0.98, 1.39]	0.009
11	C	Dyer (2000c) Ext	no problem	77	0.70	1.18	[1.00, 1.36]	0.010
12	C	Dyer (2000c) Ext	no rural, no problem	55	0.71	1.23	[1.03, 1.44]	0.007
13	D	WCS1 [4, 6]	all	93	0.68	1.11	[0.97, 1.25]	0.025
14	D	WCS1 [4, 6]	no rural	68	0.68	1.08	[0.93, 1.23]	0.031
15	D	WCS1 [4, 6]	no problem	77	0.72	1.08	[0.93, 1.22]	0.033
16	D	WCS1 [4, 6]	no rural, no problem	55	0.75	1.11	[0.95, 1.27]	0.026

17	E	WCS1 [5, 7]	all	93	0.68	1.14	[0.99, 1.28]	0.016
18	E	WCS1 [5, 7]	no rural	68	0.68	1.11	[0.95, 1.27]	0.020
19	E	WCS1 [5, 7]	no problem	77	0.72	1.10	[0.95, 1.25]	0.022
20	E	WCS1 [5, 7]	no rural, no problem	55	0.75	1.14	[0.97, 1.30]	0.016
21	F	WCS1 [5.5, 7.5]	all	93	0.68	1.15	[1.00, 1.29]	0.014
22	F	WCS1 [5.5, 7.5]	no rural	68	0.68	1.12	[0.96, 1.28]	0.017
23	F	WCS1 [5.5, 7.5]	no problem	77	0.72	1.11	[0.96, 1.26]	0.018
24	F	WCS1 [5.5, 7.5]	no rural, no problem	55	0.75	1.15	[0.98, 1.31]	0.014
25	G	WCS1 [6, 8]	all	93	0.68	1.15	[1.00, 1.31]	0.012
26	G	WCS1 [6, 8]	no rural	68	0.68	1.13	[0.97, 1.29]	0.014
27	G	WCS1 [6, 8]	no problem	77	0.72	1.12	[0.97, 1.27]	0.016
28	G	WCS1 [6, 8]	no rural, no problem	55	0.75	1.16	[0.99, 1.32]	0.012
29	H	WCS1 [7, 9]	all	93	0.68	1.17	[1.02, 1.32]	0.009
30	H	WCS1 [7, 9]	no rural	68	0.68	1.14	[0.98, 1.31]	0.011
31	H	WCS1 [7, 9]	no problem	77	0.72	1.13	[0.98, 1.29]	0.012
32	H	WCS1 [7, 9]	no rural, no problem	55	0.74	1.17	[1.00, 1.34]	0.009
33	I	WCS2 [5.94, 7.66]	all	93	0.68	1.15	[1.00, 1.30]	0.012
34	I	WCS2 [5.94, 7.66]	no rural	68	0.68	1.12	[0.96, 1.28]	0.016
35	I	WCS2 [5.94, 7.66]	no problem	77	0.72	1.11	[0.96, 1.27]	0.016
36	I	WCS2 [5.94, 7.66]	no rural, no problem	55	0.75	1.15	[0.98, 1.32]	0.012
37	J	WCS2 [5.5, 9.05]	all	93	0.68	1.10	[0.96, 1.24]	0.024
38	J	WCS2 [5.5, 9.05]	no rural	68	0.68	1.07	[0.92, 1.22]	0.030
39	J	WCS2 [5.5, 9.05]	no problem	77	0.73	1.07	[0.92, 1.21]	0.031
40	J	WCS2 [5.5, 9.05]	no rural, no problem	55	0.75	1.10	[0.94, 1.25]	0.024
41	K	WCS3 [4.52, 6.55]	all	93	0.68	1.13	[0.98, 1.29]	0.014
42	K	WCS3 [4.52, 6.55]	no rural	68	0.69	1.11	[0.95, 1.27]	0.017
43	K	WCS3 [4.52, 6.55]	no problem	77	0.73	1.10	[0.95, 1.26]	0.018
44	K	WCS3 [4.52, 6.55]	no rural, no problem	55	0.76	1.15	[0.98, 1.31]	0.013
45	L	WCS3 [4.22, 7.42]	all	93	0.68	1.09	[0.94, 1.23]	0.026
46	L	WCS3 [4.22, 7.42]	no rural	68	0.69	1.06	[0.91, 1.21]	0.032
47	L	WCS3 [4.22, 7.42]	no problem	77	0.73	1.06	[0.91, 1.20]	0.032
48	L	WCS3 [4.22, 7.42]	no rural, no problem	55	0.76	1.09	[0.94, 1.25]	0.025

S7 Bibliography

Adamic, L. A., & Huberman, B. A. (2002). Zipf's law and the Internet. *Glottometrics*, *3*, 143–150.

Bairoch, P. (1991). *Cities and economic development: from the dawn of history to the present*. University of Chicago Press.

Basu, A., Shioya, H., & Park, C. (2011). *Statistical Inference: The Minimum Distance Approach*. Boca Raton, FL: CRC Press.

Batty, M. (2006). Hierarchies in Cities and City-Systems. In D. Pumain (Ed.), *Hierarchy in Natural and Social Sciences* (pp. 143–168). Dordrecht: Springer.

Bettencourt, L. M. A. (2013). The origins of scaling in cities. *Science*, *340*(6139), 1438–1441. http://doi.org/10.1126/science.1235823

Bettencourt, L. M. A., & Lobo, J. (2016). Urban scaling in Europe. *Journal of The Royal Society*

Interface, *13*(116), 20160005. http://doi.org/10.1098/rsif.2016.0005

Britnell, R. H. (1986). *Growth and Decline in Colchester, 1300-1525*. Cambridge: Cambridge University Press.

Britnell, R. H. (2000). The Economy of British Towns 1300–1540. In *The Cambridge Urban History of Britain, Volume I: 600-1540* (pp. 313–334). Cambridge: Cambridge University Press.

Campbell, B. M. S. (1981). The population of early tudor England: a re-evaluation of the 1522 Muster Returns and 1524 and 1525 Lay Subsidies. *Journal of Historical Geography*. http://doi.org/10.1016/0305-7488(81)90117-1

Clark, P., & Souden, D. (Eds.). (1987). *Migration and society in early modern England*. London: Hutchinson.

Cornwall, J. (1962a). English Country Towns in the Fifteen Twenties. *The Economic History Review, New Series*, *15*(1), 54–69.

Cornwall, J. (1962b). The people of Rutland in 1522. *Transactions of the Leicestershire Archaeological and Historical Society*, *37*, 7–28.

Cornwall, J. (1988). *Wealth and society in early sixteenth century England*. London: Routledge.

Darby, H. C., Glasscock, R. E., Sheail, J., & Versey, G. R. (1979). The changing geographical distribution of wealth in England: 1086-1334-1525. *Journal of Historical Geography*, *5*(3), 247–262. http://doi.org/10.1016/0305-7488(79)90071-9

De Vries, J. (1984). *European Urbanization, 1500-1800*. New York: Routledge.

Dittmar, J. E. (2011). *Cities, Markets, and Growth: The Emergence of Zipf's Law*. Princeton.

Dobson, B. (1977). Urban decline in late medieval England. *Transactions of the Royal Historical Society*, *27*, 1–22.

Dyer, A. (1995). *Decline and Growth in English Towns 1400-1640* (Vol. 13). Cambridge: Cambridge University Press.

Dyer, A. (2000a). Appendix: Ranking Lists of English Medieval Towns. In D. M. Palliser (Ed.), *The Cambridge Urban History of Britain, Volume I: 600-1540* (pp. 741–770). Cambridge: Cambridge University Press.

Dyer, A. (2000b). "Urban Decline" in England, 1377-1525. In T. R. Slater (Ed.), *Towns in Decline, AD 100-1600* (pp. 266–88). Aldershot: Ashgate.

Gabaix, X. (1999). Zipf's Law for Cities: An Explanation. *The Quarterly Journal of Economics*, *114*(3), 739–767.

Galley, C. (1998). *The Demography of Early Modern Towns: York in the Sixteenth and Seventeenth Centuries*. Liverpool: Liverpool University Press.

Gibrat, R. (1931). *Les inégalités économiques: applications: aux inégalitês des richesses, à la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc: d'une loi nouvelle: la loi de l'effet proportionnel*. Librairie du Recueil Sirey.

Goose, N. (1980). Household size and structure in early-Stuart Cambridge. *Social History*, *5*(3), 347–385. http://doi.org/10.1080/03071028008567485

Goose, N. (1984). *Economic and social aspects of provincial towns: a comparative study of Cambridge, Colchester and Reading, c. 1500–1700*. University of Cambridge.

Goose, N., & Hinde, A. (2006). Estimating local population sizes at fixed points in time: Part I - General principles. *Local Population Studies*, (77), 66–74.

Goose, N., & Hinde, A. (2007). Estimating local population sizes at fixed points in time: Part II - Specific sources. *Local Population Studies*, (78), 74–88.

Hoskins, W. G. (1956). English Provincial Towns in the Early Sixteenth Century. *Transactions of the Royal Historical Society*, *6*, 1–19. http://doi.org/10.1017/S0080440100019630

Johnson, G. A. (1981). Monitoring complex system integration and boundary phenomena with settlement size data. In S. E. van der Leeuw (Ed.), *Archaeological approaches to the study of complexity* (pp. 143–188). Amsterdam: A. E. van Giffen Instituut voor Prae-en Protohistorie.

Jurkowski, M. (2015). The History of Clerical Taxation in England and Wales, 1173–1663: The Findings of the E 179 Project. *The Journal of Ecclesiastical History*, *67*(01), 53–81.

Keene, D., & Rumble, A. R. (1985). *Survey of medieval Winchester*. Oxford: Oxford University Press.

Laslett, P. (1972). Mean household size in England since the sixteenth century. In P. Laslett & R. Wall (Eds.), *Household and family in past time* (pp. 125–158). Cambridge: Cambridge University Press.

Limpert, E., Stahel, W. A., & Abbt, M. (2001). Log-normal Distributions across the Sciences: Keys and Clues. *BioScience*, *51*(5), 341–352.

Montroll, E. W., & Shlesinger, M. F. (1982). On 1/f noise and other distributions with long tails. *Proceedings of the National Academy of Sciences*, *79*(10), 3380–3383.

Palliser, D. M. (1979). *Tudor York*. Oxford: Oxford University Press.

Palliser, D. M. (1992). *The Age of Elizabeth*. London: Longman.

Patten, J. (1978). *English Towns 1500-1700*. Folkestone: Dawson.

Phythian-Adams, C. (1978). Urban decay in late medieval England. In P. Abrams & E. A. Wrigley (Eds.), *Towns in Societies: Essays in Economic History and Historical Sociology* (pp. 159–185). Cambridge: Cambridge University Press.

Phythian-Adams, C. (1979). *Desolation of a city: Coventry and the urban crisis of the late Middle Ages*. Cambridge: Cambridge University Press.

Pound, J. F. (1971). *The Norwich census of the poor, 1570*. Norfolk Record Society.

Pound, J. F. (1976). The social and trade structure of Norwich 1525-1575. In P. Clark (Ed.), *The Early Modern Town, A Reader* (pp. 129–47). New York: Longman.

Pound, J. F. (1986). *Poverty and vagrancy in Tudor England* (2nd Editio). Longman.

Pumain, D. (2000). Settlement Systems in the Evolution. *Geografiska Annaler, Series B: Human Geography*, *82*(2), 73–87. http://doi.org/10.1111/j.0435-3684.2000.00075.x

R Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Razali, N., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, *2*(1), 21–33.

Rigby, S. H. (2010). Urban population in late medieval England: the evidence of the lay subsidies. *The Economic History Review*, *63*(2), 393–417. http://doi.org/10.1111/j.1468-0289.2009.00489.x

Schofield, R. (1965). The geographical distribution of wealth in England, 1334–1649. *The Economic History Review*, *18*(3), 483–510.

Schofield, R. (2004). *Taxation Under the Early Tudors 1485-1547*. *Taxation Under the Early Tudors 1485-1547*. Oxford: Blackwell. http://doi.org/10.1002/9780470758298

Sheail, J. (1968). *The regional distribution of wealth in England as indicated in the 1524/5 lay subsidy returns*. University of London.

Souden, D. (1984). Migrants and the population structure of later seventeenth-century provincial cities and market towns. In P. Clark (Ed.), *The Transformation of English Provincial Towns 1600-1800* (pp. 133–68). London: Hutchinson.

Stephens, M. A. (1974). EDF Statistics for Goodness of Fit and Some Comparisons. *Journal of the American Statistical Association*, *69*(347), 730–737.

Thompson, J. A. F. (2014). *The Early Tudor Church and Society, 1485-1529*. London: Routledge.

Wall, R. (1972). Mean houshold size in England from printed sources. In P. Laslett & R. Wall (Eds.), *Household and family in past time* (pp. 159–204). Cambridge: Cambridge University Press.

Wrigley, E. A., & Schofield, R. (1981). *The Population History of England, 1541-1871: A Reconstruction*. Cambridge, Mass.: Harvard University Press.

Zeileis, A. (2004). Econometric Computing with HC and HAC Covariance Matrix Estimators. *Journal of Statistical Software*, *11*, 1–17. http://doi.org/10.18637/jss.v011.i10