

**WestminsterResearch**

<http://www.westminster.ac.uk/westminsterresearch>

**The fourteenth-century poll tax returns and the study of English  
surname distribution**

**Parkin, D.H.**

This is an Accepted Manuscript of an article published by Taylor & Francis in Historical Methods: A Journal of Quantitative and Interdisciplinary History, 48 (1), pp. 1-12. The final definitive version, with images, is available online:

<https://dx.doi.org/10.1080/01615440.2014.946985>

© 2015 Taylor & Francis

---

The WestminsterResearch online digital archive at the University of Westminster aims to make the research output of the University available to a wider audience. Copyright and Moral Rights remain with the authors and/or copyright owners.

---

Whilst further distribution of specific materials from within this archive is forbidden, you may freely distribute the URL of WestminsterResearch: (<http://westminsterresearch.wmin.ac.uk/>).

In case of abuse or copyright appearing without permission e-mail [repository@westminster.ac.uk](mailto:repository@westminster.ac.uk)

**The Fourteenth-Century Poll Tax Returns and the  
Study of English Surname Distribution**

Harry Parkin

Research Associate - Family Names of the United  
Kingdom (FaNUK) research project

University of the West of England, Bristol, UK

[harry.parkin@uwe.ac.uk](mailto:harry.parkin@uwe.ac.uk)

## **The Fourteenth-Century Poll Tax Returns and the Study of English Surname Distribution**

**Abstract.** The modern-day distributions of English surnames have been considered in genealogical, historical, and philological research as possible indicators of their origins. However, many centuries have passed since hereditary surnames were first used, and so their distribution today does not necessarily reflect their original spread, misrepresenting their origins. Previously, medieval data with national coverage have not been available for a study of surname distribution but, with the recent publication of the fourteenth-century poll tax returns, this has changed. By presenting discrepancies in medieval and 19<sup>th</sup>-century distributions, it is shown that more recent surname data may not be a suitable guide to surname origins, and can be usefully supplemented by medieval data in order to arrive at more accurate conclusions.

**Keywords:** surname, by-name, distribution, medieval period, poll tax

The study of a surname's distribution has long been recognized as a useful approach in the investigation of its origins, ever since Guppy (1890, 6) set out "to ascertain the homes of familiar surnames and to ascertain the characteristic surnames of each county." It is known that much migration in England did not tend to occur over great distances, and as a result, "even now, so many English surnames are found close to where they originated, or in areas to which they had moved early in their history" (Redmonds 1997, 16). As Clark (2002, 102) has pointed out, "naming respects ... the linguistic and cultural divisions within the larger realm," meaning many of England's surnames have specific regional characteristics, borne out of such factors as local place-names, given-names and occupations, while also representing historical regional dialects in their spellings. It follows that the study of a surname's distribution can provide a good starting point for the discovery of its linguistic or geographical origin, and is therefore a useful methodological approach for the historians, genealogists and philologists who analyze surname data.

While uncovering the distribution of certain surnames was once a very time-consuming process, requiring careful analysis of many local records, recent software, like that created by Steve Archer (2011), allows quick access to such information. Archer's (2011) *British 19<sup>th</sup> Century Surname Atlas* can be used to generate distribution maps of all surnames recorded in the UK 1881 census, which can help to confirm or reassess established suggestions on name

origins, made at a time when such data was not so readily accessible, or to show the possible origins of names which have not been tackled in previous research. However, as most of the population had adopted hereditary surnames as early as 1350 in the south of England and 1450 in the north (see McKinley 1990, 31–32), the reliance on recent name distribution for the study of their origins is potentially problematic.

This approach uses the recent distribution of a name to uncover information about its medieval origins, but it would be expected that there have been many different contributors to change in a name's distribution since the period of its formation, such as widespread ramification. Up until now, it has been difficult to confirm this suspicion, with a greater number of studies aiming to show continuity in distribution since the fourteenth century, rather than change; see, for example, Schürer's (2004, 56) finding that the "broad regional distribution of the three surnames [*Fuller, Tucker and Walker* in 1881] is similar to that of the early fourteenth century." However, thanks to the recent publication of the 1377, 1379 and 1381 poll tax returns (Fenwick 1998, 2001, 2005), such investigation is now possible, using the first available collection of medieval name data with national coverage. The aim of this paper is to establish that the recent distribution of surnames is no safe guide to their geographical origins, and that there is therefore value in a distribution study of medieval name data, by showing differences between the fourteenth century and more recent periods.

Previous analyses have studied surname distribution at times earlier than 1881, but have been unable to look any further into the past than the sixteenth century with any great confidence, due to a lack of data with national coverage. In 1538 “a mandate of 5 September” (FitzHugh 1988, 213) ordered weekly recordings of baptisms, marriages and burials in parish registers. Some have studied the names of the sixteenth century using these parish registers, including Leeson (1989) and Barker et al. (2007) (also see Viereck 2005, 2008a, 2008b, 2009, outlining the important work carried out by Barker et al., which he co-authored), the latter scholars being aided greatly by the online database of parish registers provided by the Church of Jesus Christ of Latter-Day Saints, known as the International Genealogical Index (IGI) (FamilySearch 2014), among other sources. The IGI is not a perfect resource, containing some duplicate entries and omitting certain religious groups, but for a general picture of surname distribution patterns it is certainly suitable. While previous reviews of the IGI have questioned the reliability and consistency of its transcriptions, Hanks, Coates, and McClure (2012, 48) believe “these deficiencies have been overstated.” Most other surname distribution studies have investigated names of more recent periods, usually from the nineteenth century onwards, presumably because the data from this period provide better national coverage and statistical reliability in their greater numbers.

Accounts of by-name and surname distribution using medieval records do exist,<sup>1</sup> but mostly as part of more general works on the history of English surnames (see, for example, Reaney 1967, 321–356 and McKinley 1990, 177–187), and are therefore less detailed than distribution studies for later periods. It seems that the only analysis of medieval names in a work dedicated solely to their distribution has been carried out by Rogers (1995, 144–224), who also investigates names from modern-day and other post-medieval records. The comparative lack of research into the medieval national distribution of English by-names and surnames appears to be due to the nature of the data, with most collections of records providing inconsistent coverage of the country because many entries have become damaged or lost. Furthermore, the majority of medieval tax records did not include those people who were “too poor to be taxed” (FitzHugh 1988, 160), and considering that

if the surnames or by-names in use in English communities in the period from approximately 1100 to 1400 are analysed in class terms, it can be seen that there were sharp differences between one class and another in the nature of the names in use (McKinley 1990, 201),

this means that an analysis of medieval name distribution using such records will not consider those particular names that were predominantly borne by people from lower social classes. These, and other, issues mean there are

greater methodological problems in medieval name distribution research, when compared with the study of later periods, as Rogers (1995, 161) states in an introduction to his analysis, referring to fourteenth-century records:

Not for another two hundred years do we begin to have a regular series of sources for surname distribution, making the investigation of hypotheses about the intervening development and movement of surnames very difficult to undertake. There is no doubt that the fourteenth-century scan which follows is therefore a much more opaque indication of the presence of surnames than its seventeenth- and twentieth-century counterparts, and the surname detective tracking down individual names must have recourse to a much wider range of sources.

At the time Rogers' (1995) study was published, such a wide range of sources was also required in order to gather a representative sample, containing a suitable number of names for reliable conclusions on their distribution to be made; as Hanks (1992, 91) points out, "distribution only gets interesting when there is a large enough number of bearers for patterns to be measured." Rogers (1995, 224) closes his work by stating "it is ... clear that, the rarer the name, the less likely it is that the distribution of its early examples will be visible in the fourteenth-century sources until the Poll Tax becomes



widely available.” He is referring to the fourteenth-century poll tax returns (from now on referred to as the PT) which, as has been mentioned previously, have now been made available in a published collection of transcribed material, making it possible to form more reliable conclusions on the distribution of medieval by-names and surnames than it has been before.

There is general consensus on the importance of the PT returns to anthroponomastic study. Rogers (1995, 149) notes that the records “include the names and payments of some 60 per cent of the whole population, several times more than may be found in the earlier Lay Subsidies,” and that they are “by far the best source available” for the analysis of medieval names. McKinley (1990, 32) believes that “the late fourteenth-century poll tax returns ... give a more complete view of the names then in use than any other source for the same period,” and Hey (2000, 46–47) states that even though they “are unsatisfactory in their coverage, because they simply do not survive for many counties or are incomplete, ... they are the best source that we have for identifying distribution patterns for surnames close to the period of formation.” This is in part due to the fact that the PT attempted to assess the entire population of the country, resulting in the most complete records, in terms of people from all social classes, of their time; as explained above (see McKinley 1990, 201), other records which do not include all classes are likely to misrepresent the true nature of the names used by the entire population at that particular time.

It is clear that the anthroponomastic importance of the PT returns is well appreciated, and now that the extant records from the entire country have been made available by Fenwick (1998, 2001, 2005), we can investigate the medieval names of England in greater detail than has been previously possible, in order to contribute to our knowledge of medieval by-name and surname distribution. This is not to say, however, that the PT returns are a perfect resource. While preferable to other medieval records, they still have some deficiencies which are methodologically problematic, and must be considered before any analysis of their data is carried out.

An initial look at the PT returns shows that there are some counties for which there are no surviving records. Names from Cheshire, County Durham, Hertfordshire, Huntingdonshire, and the City of London do not appear in Fenwick's transcribed PT volumes, and so the counties are labelled "nd" (no data) in the maps presented below. It is possible that some of the names included in a section titled "Unidentified" (Fenwick 2005, 580–599), where Fenwick has gathered all records which have been damaged to the point that their origin cannot be determined, could be from some of these counties, but this cannot be known (though suggestions could be made following comparative linguistic analysis of names in these documents with names in the other PT returns which have a known geographical origin). Perhaps of greatest concern are the missing data from the City of London, as the pull of the capital is known to have had a significant effect on population movement, drawing in

migrants from all over England in considerable numbers. For example, the average annual population increase in London between 1650 and 1750 was 2,750 (Wrigley 1967, 46). Corresponding figures from the time of the PT are not known, but London is sure to have played some role in surname distribution change since the 14<sup>th</sup> century up to the present day, and so it is unfortunate that the surname data from the city are unavailable. Even so, there are extant PT returns for the remaining 35 English counties,<sup>2</sup> as well as for the city of York, still providing a wealth of data for analysis. Furthermore, the missing counties are by no means the largest, nor do they lie in the same areas of England, and so the absence of their PT returns is unlikely to affect the national distribution findings in any major way, except for any names which only occurred in any of these missing counties.

Greater methodological concerns are the differing levels of damage to the returns from each county, and of coverage provided for each county by the extant records. Worcestershire, for example, is poorly represented, with only the city of Worcester having any surviving returns. There are a number of returns missing from certain vills in many other counties,<sup>3</sup> and every county has damaged returns to some extent, all of which are identified by Fenwick in her introductions to each collection of county PT records. Unfortunately, an approximate count of the number of names unsuitable for analysis cannot be easily arrived at from this information alone, which is more of a general guide to the level of damage to the returns for each county. Only by analysing all

entries in the extant returns from all counties can the methodological implications of their damage be fully understood.

An account of this damage, and indeed of the national distribution of the names in the PT returns, gained by reading through Fenwick's printed volumes, would have required many decades of work. However, she kindly supplied her transcriptions of all available PT returns in database form, greatly speeding up the process. Using this database in Microsoft Excel, it has been possible to create a pivot table which groups all names of identical form together and counts their occurrences, and can also be filtered to show the names from individual counties, or from groups of counties.<sup>4</sup> This table contains all names which have been transcribed by Fenwick, including those which she marks with "{f}", meaning faded, and "{t}", meaning torn; these can appear as part of a transcribed name, indicating that some of its letters were indiscernible, or as an individual entry, indicating that no part of the name could be determined.

Entries transcribed as {f} or {t} are by far the two most common in all of the PT returns. There are a number with {f} or {t} as part of a transcribed name, but these are comparatively few and make up a small proportion of the total 189,220 surname entries in Fenwick's volumes.<sup>5</sup> 7,209 entries are faded and 3,359 are torn to an extent that Fenwick has been unable to transcribe any part of the name; these make up 3.81% and 1.78% of all surname entries respectively. This level of damage is not, however, evenly distributed

throughout the country. For example, 44.09% of all entries for Westmorland are {f}, while the corresponding figure for Surrey is only 0.04%, and 6.48% of all entries for Suffolk are {t}, but there are none for Derbyshire. Considering these differences, any comparison of name distribution between different counties must bear in mind that their PT returns may contain different amounts of damaged names, which could misrepresent the true proportions of certain names in individual counties during the fourteenth century. It would not, however, be appropriate to exclude these damaged entries from the dataset, otherwise the relative frequency of other names might be overestimated. If, for example, a name appears to have a marked distribution or a particularly high frequency in a certain county, even when a large number of damaged entries make up a considerable proportion of the names in that county, there can be confidence in that pattern's validity. However, if damaged entries were excluded and all of them were actually different names from the one being analyzed, the name under investigation would appear proportionately more frequent than it might have been.

As mentioned above, the different levels of coverage provided by the extant PT returns for each county also pose a problem to distribution analysis, with some surviving county records containing many more names than others. As a result of the creation of the pivot table, and by filtering it for the names from each county individually, it has been possible to see what proportions of all names in the PT appear in each county, and therefore which county records

can be more reliably analyzed than others. The West Riding of Yorkshire contains the greatest proportion of all names recorded in the PT, being 13.46%, while Buckinghamshire contains the lowest proportion of 0.16%. However, so long as comparisons of name frequency are carried out in terms of a name's proportional representation in each county, distribution analysis is still possible. Of course, medieval name samples from counties which have very few surviving PT returns are unlikely to be truly representative, but if this is borne in mind, and the possible implications are discussed where relevant, meaningful comparison of medieval name distribution for different counties can still be carried out using the PT returns.

There are, then, many difficulties in using the PT returns for a name distribution study, as is also the case for anthroponomastic research with other medieval records. It has been argued, however, that with suitable caution, such a study is possible. This being the case, the remainder of this paper will consider the national and regional distribution of some names in the PT, selected for discussion because they exhibit different patterns of distribution in the fourteenth century and 1881, in order to show that the recent distribution of a name is not necessarily a reliable indication of its medieval pattern. It is hoped that this will lead to further, more detailed research using the PT returns, an important source of name data which could greatly improve our knowledge of surname distribution from the fourteenth century up to the present day.

The most frequent toponymic name (derived from a place-name) in the PT is *Burton*, in that exact form,<sup>6</sup> with its 135 bearers making up 0.07% of all names recorded. It has been selected for analysis because there are a number of places so named in England from which any instance of the surname could possibly be derived, but as most medieval bearers are likely to have been found close to the place of origin of their name, its distribution might suggest that the majority of *Burton* surnames are likely to have originated in only a small number of the numerous corresponding place-names. If, for example, bearers of the name *Burton* were most heavily concentrated in Northamptonshire, then it is likely that the majority of people with that surname today, and all others since the establishment of surname heredity, originally took their surname from Burton Latimer in that county, rather than from another of the numerous major and minor similarly named places in England.

It is known that some surnames which were once presumed to be polygenetic are in fact monogenetic, apparently originating from a single ancestor (see Sykes and Irven's (2000) study of the surname *Sykes*), and so, the fact that there are many places named Burton in England is no reason to assume that the surname *Burton* is polygenetic. In other words, it is possible that the surname *Burton* came from a single place, or a certain number of places, so named, rather than having been coined separately in each English place called Burton.

It is important to note that the by-name or surname *Burton* is not necessarily from a place-name which shares the exact same form today. As spelling was not standardized at the time of by-name and surname formation, there was much variation in the spelling of by-names, surnames, and place-names. The most likely origin of the surname *Burton* is any place named with Old English *burh-tūn* ‘fortified enclosure, fortified settlement’, most of which have the modern form *Burton*, but some have *Bourton*. There are also other places with the name *Burton*, with various other etymological origins, which could have given rise to the surname *Burton*, and there are also place-names with other modern forms which could have given rise to the surname, such as Barton in Cambridgeshire (etymologically from Old English *beretūn* ‘barley enclosure, barley farm; outlying farm, demesne farm’) which was recorded as *Burton* in 1202, 1219, and 1342 (Reaney 1943, 72–3).

The most frequent form of the surname in 1881 was *Burton*, and its proportion, out of all names in the census, is not too dissimilar from that of *Burton*’ in the fourteenth century, making up 0.1% of names in England.<sup>7</sup> In terms of relative frequency, then, the name shows continuity between the two periods. It is worth mentioning here that the variable nature of name forms between the two periods means that the diachronic comparison of identical forms is often not worthwhile, and is sometimes impossible as some fourteenth-century forms had died out by 1881, and some 1881 forms did not exist in the fourteenth century. The approach suggested in this article, where



the most frequent forms of names which appear to share the same origin are compared, assumes that the majority of fourteenth-century name forms would have been changed, at some point, for what became their most frequent equivalent form by 1881. The comparison of these different forms therefore represents the distributional development of what is essentially the same name.

If, for example, only names of the exact form *Burton*, without an apostrophe, were compared, it would not give a general picture of the distributional change of people named after any of the relevant place-names. Instead, it would be a meaningless comparison of an unusual fourteenth-century scribal variant, which had a highest concentration of three bearers in the East Riding of Yorkshire, with the 129th most frequent name in the United Kingdom in 1881. The adoption of a policy where only names of identical form are compared would be to disregard changes in orthographical convention and written representation of phonology between the two periods. Of course, it could be argued that all names which appear to share the same origin, regardless of form, should be considered in comparison, but this might risk including forms of low frequency which are therefore more likely to be a result of scribal error or the normalisation of an unfamiliar name, and so do not represent the true national distribution of the name being investigated.

Of course, there are instances where apparently related names of different form, when they are the most frequent in either period, cannot be

usefully compared, such as *filia/filio/filius/relicta* (etc.) *Johannis* in the PT and *Johnson* in the 1881 census. While they are both the most frequent names from their respective periods which have been formed from the name known today as *John*, the uncertain development of *Johannis*, a Latin genitive form of *Johannes*, means it cannot be considered simply as an early form of the surname *Johnson*. *Johannis* will have been applied as a non-hereditary patronymic description, meaning ‘of John’, used to refer to the bearer’s relationship to someone called *John*. *Johannis* could then have, hypothetically, developed into a number of patronymic by-names and subsequent surnames, such as *John* and *Jones*, as well as *Johnson*, whose differences are not simply orthographical. However, on the whole, so long as only orthographical and phonological variants are considered for comparison, the safest approach for discovering differences and changes in name distribution between the fourteenth century and 1881 seems to be the comparison of the most frequent forms of a name in either period, as has been carried out for *Burton*’ and *Burton* to which we will now return.

In 1881, the proportional distribution of the surname *Burton* shows it was most common to the counties of Rutland and Nottinghamshire, which were the only two to contain over 300 instances of the name per 100,000 people; Rutland had 509 and Nottinghamshire had 359. Overall, the name appears to have been most common to the Midlands in 1881, though also with relatively high concentrations in the north, as shown in Figure 1. According to

the 1881 distribution of the name, the surname researcher might assume that the most likely origin of the name *Burton* is a place so named in Rutland or Nottinghamshire, the only candidates being Burton Joyce and West Burton in Nottinghamshire. However, the PT distribution is slightly different, and is shown in Figure 2.

<FIGURE 1 ABOUT HERE>

<FIGURE 2 ABOUT HERE>

While there is still a fairly high proportion of the name in the Midlands, it was most heavily concentrated in York and the East Riding of Yorkshire, and was not present at all in Rutland, where it had the highest concentration in 1881. Perhaps, then, according to the fourteenth-century distribution of the name *Burton*', most instances of the name actually derive from a toponym near York or in the East Riding of Yorkshire, such as Cherry Burton (though there are 6 other parishes in the East Riding with the place-name *Burton*), whose bearers might then have migrated to the Midlands some time before 1881. This, of course, cannot be certain, and the many different places named Burton throughout England make it likely that the surname is polygenetic, even though most of the names may have originated in the north and the Midlands. The important point is that the name *Burton*'/*Burton* had noticeably different national distributions in the fourteenth century and in 1881, showing that more recent distributions of English

surnames are not necessarily an accurate reflection of their geographical spread throughout their history.

While it is possible to recognize some slight continuity in the distribution of the name *Burton*/*Burton* between the time of the PT and 1881, being present in north England and the Midlands, this is not necessarily the case for all names. The surname *Adam*, derived from the Hebrew given name which first became popular in England in the twelfth century (see Withycombe 1977, 3), was rare in England in 1881, and mostly found in Scotland as shown in Figure 3. The fourteenth-century distribution is, however, very different, with greater relative proportions of the name in most English counties, shown in Figure 4. Therefore, it seems that sometime between the late fourteenth century and 1881 most bearers of the name *Adam*, either as a by-name or surname, who lived in England, had lost the name, perhaps due to an ancestor who, for example, died and failed to pass on the surname to male offspring, moved to Scotland, or began to use some sort of altered form of the surname. The latter suggestion is certainly a possibility, as the post-medieval addition of an excrescent *-s* to surnames, while relatively rare, is well evidenced. See, for example, the name *Bartons*, which had 13 bearers in 1881; the final *-s* can only be an unetymological addition as the surname is derived from the toponym Barton, which occurs frequently in England but never with final *-s*. Alternatively, the addition of a genitive *-s* to

*Adam* before it became hereditary could also have contributed to its apparent reduction in numbers.

<FIGURE 3 ABOUT HERE>

<FIGURE 4 ABOUT HERE>

As with the *Burton/Burton*' example, the change in distribution of the name *Adam* should serve as a warning against the use of recent surname distributions as evidence of a name's geographical origin or medieval concentration. Such an approach can be appropriate, as the recent distribution of a name is sometimes almost indistinguishable from its medieval alternative, particularly when studying widespread polygenetic names, such as *Smith*. However, if the 1881 distribution of *Adam* was taken as being indicative of the name's continual distribution since its formation, then the genealogist, historian or anthroponomastician would incorrectly assume that it was rarely ever found as a by-name or surname in England, and could have originated in Scotland.

In this paper, discrepancies between medieval and more recent surname distributions have been highlighted to show that post-medieval distribution is not necessarily a suitable indication of a surname's geographical origin. In addition to this method, their comparison can also contribute to an investigation of past population movements. If a surname is monogenetic, in other words having a single progenitor from whom all bearers of that name

descend, then differences between its distribution close to the time of hereditary surname adoption and a more recent period indicate that migration has occurred at some time within the date range covered by the datasets used for analysis.

It is not possible to be certain that a name is monogenetic without DNA testing, and so this type of analysis can be methodologically problematic, though there is a relatively high possibility that surnames from unique toponyms are monogenetic, and can therefore be usefully analyzed in an investigation of population movements. Even if multiple unrelated families took their surname from the same toponym, making the surname polygenetic, any bearer of a surname from a unique toponym must have an ancestor who once resided at the place denoted by the name, and so the occurrence of such a surname outside of the place which it denotes is an indication of past migration.

As an example, let us consider the surname *London*. The PT distribution of the name (see Figure 5) shows its highest proportional concentration was in Staffordshire. Whether or not the family or families responsible for this concentration were first bestowed with the name *London* on arriving in Staffordshire, or already bore the name beforehand, it is reasonable to assume that they or their ancestors once lived in or near London. It is possible to imagine that some people may have been bestowed with *London* as a by-name for some other reason, such as expressing a desire to live

in the city, but it seems most likely that the majority of these bearers were so named because they or their ancestors lived in London. The county with the next highest proportion of people named *London* in the PT was Kent, unsurprisingly given the county's proximity to London.

In 1881, the distribution of the name is different, with the epicentre in Norfolk and the next highest concentration in the neighbouring county of Suffolk (see Figure 6); in Staffordshire, where the name was proportionately most common in the 14<sup>th</sup> century, the name had become relatively uncommon, with only 3 bearers of *London* per 100,000 people, as opposed to 119 in the 14<sup>th</sup> century. There are a number of possible reasons for such changes in the surname's concentration. Perhaps the family responsible for the medieval frequency of *London* in Staffordshire died out; perhaps another unrelated family with the name *London* migrated to Norfolk from London after the time of the PT; perhaps a family adopted the name *London* after the time of the PT and settled in Norfolk; or, perhaps the majority of people with the name *London* in Staffordshire in the 14<sup>th</sup> century migrated to Norfolk before 1881.

<FIGURE 5 ABOUT HERE>

<FIGURE 6 ABOUT HERE>

Whatever the reason, the discrepancies between the two maps show some form of change in the population between the two periods they represent, be it migration, surname death, or new surname adoption. If the

name *London* was borne by only a single family, then the differences must be due to migration, and in such cases this kind of comparison is a powerful tool for the investigation of historical migration across England. However, even in the case of a surname from a unique toponym, such as *London*, it is not possible to be certain that the surname is and has always been monogenetic; a number of unrelated families may have taken their surnames from the city independently.

With increasing Y chromosomal DNA testing in academic research (see, for example, King and Jobling 2009), as well as by those interested in their family history using affordable and widely available testing kits, unambiguously monogenetic surnames can sometimes be identified, and their medieval and more recent distributions compared in order to discover past migration patterns of the general population, without the possibility that there may be other reasons for the apparent change in their distribution.

While only a small number of examples have been given, it is hoped it has been made clear that the PT is a valuable data source with the potential to greatly improve our knowledge of by-name and surname distribution and development, as well as historical patterns of migration, even though there are many methodological difficulties to overcome in their use. While in some ways distribution has remained fairly similar over roughly 500 years, there are also noticeable changes which could affect the choices made by researchers in the investigation of a name's origin. Additionally, comparison of medieval and



1881 distributions of the surname *London* shows that the present-day distribution of surnames from unique toponyms does not necessarily provide a complete picture of population movements since the time of surname formation. The PT, then, provides new evidence on the history of English surname distribution, and through a comparison of its medieval names with those from more recent periods, a wide range of new and previously unrecognized information on surname origins and history can be uncovered.

The PT returns do not only have the potential to greatly improve our knowledge of certain by-name and surname origins through an appreciation of their medieval distributions, but to provide a more accurate picture of changes in surname frequency since the fourteenth century, close to the time of their formation. Mathematical estimates have been made in the past (see Sturges and Haggett 1987), but no real data with national coverage have yet been analyzed for this purpose.

As a starting point to such research, the most frequent names in the PT and the 1881 census have been compared. Out of the fifty most frequent names in the PT, excluding those which could have developed into multiple forms, such as *Johannis*, and only counting once those names which have more than one particularly frequent variant, such as *Smyth'* and *Smyth*, only twenty-two were also included in the fifty most frequent names in 1881. Further comparison of all names in the PT and 1881 census could be carried out in order to determine the extent to which surname frequency has changed,

as well as to identify typical patterns of national population mobility over an extended period.

Clearly, then, there was noticeable change in the relative frequency and the distribution of England's names between the end of the fourteenth century and 1881. The mechanisms and reasons behind such change are not clear, but further analysis of the PT will allow us to reach greater precision on the level of change in the English name stock from the fourteenth century to the present day, which was not possible with the sporadic coverage provided by previously published record collections. Now that the PT returns have been transcribed, and the national distribution of medieval by-names and surnames is much clearer than it once was, it appears that there is much about surname development between the time of the PT and 1881 that we do not understand. Future anthroponomastic research will benefit from studying the names of the PT in detail to improve our knowledge of their use and stability at this crucial point in the history of their development.<sup>8</sup>

## NOTES

Acknowledgements: The author wishes to thank Richard Coates, for the many fruitful discussions related to this research; Carolyn Fenwick, for providing her electronic database of the PT data; and the editors and anonymous reviewers, for their valuable insights and constructive comments.

1. The term *by-name* is used to refer a non-hereditary name which was descriptive of the bearer, while *surname* is used to refer to an hereditary family name.
2. The three Ridings of the county of Yorkshire are, however, considered separately in analysis, effectively increasing this number to 37.
3. The vill was an administrative “district or group of houses that bore a name. A parish might contain several vills” (FitzHugh 1988, 293).
4. This was done by duplicating the “Surname” column in Fenwick’s database, and selecting these two identical columns, as well as the “County” column. With this data selected, a pivot table was created using the option under the “Insert” toolbar in Microsoft Excel. Within the resulting “PivotTable Field List”, the “County” field was added to the “Report Filter” box, one of the “Surname” fields was added to the “Row Labels” box, and the other was added to the “Values” box.
5. This PT surname count of 189,220 represents 7.5688% of the estimated total population of the time (based on Broadberry, Campbell, and

van Leeuwen's (2011, 26) estimate that the population of England was 2,500,000 in 1377). This is lower than the 60% suggested as the proportion of the population covered by the PT (Rogers 1995, 149), probably partly because many wives, children, other relatives, and servants were not recorded with surnames. Furthermore, any person under the age of 14 was not recorded in the 1377 PT, any person under 16 was not recorded in the 1379 PT, and those under 15 were not recorded in the 1381 PT; paupers were also excluded from the tax. Damage to a number of the PT returns is also a likely contributor to this discrepancy in the representation of the population.

6. Variants have not been grouped together in this study, allowing for more direct comparison with Archer's (2011) atlas, which distinguishes surnames by form only. Names including apostrophes, like in the case of *Burton'*, have been treated as distinct from their corresponding names which do not end with an apostrophe, as this punctuation represents an abbreviation, and therefore a different form to an apostrophe-less equivalent.

7. Note that, in this paper, the names from Scotland, Wales and the Channel Islands have been excluded from the data, giving a total sample of 24,454,028.

8. In order to do this, the PT returns must be widely available in machine-readable form, allowing them to be accessed and analyzed by any person who wishes to do so. Dr Carolyn C. Fenwick, whose transcription of the PT returns is currently available in printed form, has, over the last five

years, attempted to have her electronic database of the material made freely available to all by an academic institution, but has had no success. In personal correspondence, Dr Fenwick has advised me that she is open to ideas on how such a database could be hosted and made freely available, and that she would be very grateful for any suggestions that the readers of *Historical Methods* may have. Please send any suggestions to me, the author, by the contact email address provided in this paper.

## REFERENCES

Archer, S. 2007. *GenMap UK*. Version 2.2. CD Rom. Archer Software.

\_\_\_\_\_. 2011. *British 19<sup>th</sup> century surname atlas*. Version 1.1. CD Rom. Archer Software.

Barker, S., et al. 2007. *An atlas of English surnames*. University of Bamberg Studies in English Linguistics 52. Oxford: Peter Lang.

Broadberry, S., B. M. S. Campbell and B. van Leeuwen. 2011. English medieval population: Reconciling time series and cross sectional evidence. Working paper [pdf] available at: <http://www.lse.ac.uk/economicichistory/pdf/broadberry/medievalpopulation.pdf>

Clark, C. 2002. Socio-economic status and individual identity: Essential factors in the analysis of Middle English personal-naming. In *Naming, Society*

*and Regional Identity*, edited by D. Postles, 101–121. Oxford: Leopard's Head Press.

FamilySearch. 2014. *International Genealogical Index*. [online] Available at: <https://familysearch.org/search/collection/igi> [accessed 28th June 2014].

Fenwick, C. C. 1998. *The poll taxes of 1377, 1379 and 1381. Part 1 Bedfordshire–Leicestershire*. Oxford: Oxford University Press.

\_\_\_\_\_. 2001. *The poll taxes of 1377, 1379 and 1381. Part 2 Lincolnshire–Westmorland*. Oxford: Oxford University Press.

\_\_\_\_\_. 2005. *The poll taxes of 1377, 1379 and 1381. Part 3 Wiltshire–Yorkshire*. Oxford: Oxford University Press.

FitzHugh, T. V. H. 1988. *The dictionary of genealogy*. Revised ed. Sherborne: Alphabooks.

Guppy, H. B. 1890. *Homes of family names in Great Britain*. London: Harrison and Sons.

Hanks, P. 1992. The present-day distribution of surnames in the British Isles. *Nomina* 16: 79–98.

Hanks, P., R. Coates and P. McClure. 2012. Methods for studying the origins and history of family names in Britain: Philology meets statistics in a multicultural context. In *Facts and findings on personal names: Some European examples. Proceedings of an international symposium in Uppsala*,

*October 20–21, 2011*, edited by L. Larsson and S. Nyström, 37–58. Uppsala: Acta Acadamiæ Regiæ Scientiarum Upsaliensis (Kungl. Vetenskapssamhällets Handlingar).

Hey, D. 2000. *Family names and family history*. London: Hambledon and London.

King, T. E. and M. A. Jobling. 2009. What's in a name? Y chromosomes, surnames and the genetic genealogy revolution. *Trends in Genetics* 25 (8): 351–360.

Leeson, F. 1989. The study of single surnames and their distribution. *The Journal of One-Name Studies* 3 (6): 174–181.

McKinley, R. A. 1990. *A history of British surnames*. London: Longman.

Reaney, P. H. 1943. *The place-names of Cambridgeshire and the Isle of Ely*. Cambridge: Cambridge University Press, for the English Place-Name Society.

Reaney, P. H. 1967. *The origin of English surnames*. London: Routledge.

Redmonds, G. 1997. *Surnames and genealogy: A new approach*. Boston, MA: New England Historic Genealogical Society.

Rogers, C. D. 1995. *The surname detective: Investigating surname distribution in England, 1086–present day*. Manchester: Manchester University Press.

Schürer, K. 2004. Surnames and the search for regions. *Local Population Studies* 72: 50–76.

Sturges, C. M. and B. C. Haggett. 1987. *Inheritance of English surnames*. London: Hawgood Computing.

Sykes, B. and C. Irven. 2000. Surnames and the Y chromosome, *American Journal of Human Genetics* 66: 1417–19.

Viereck, W. 2005. Towards an atlas of English family names. *Romanian Journal of English Studies* 2: 129–158.

\_\_\_\_\_. 2008a. English family names. *Studia Anglica Posnaniensia* 44: 155–191.

\_\_\_\_\_. 2008b. Language on the map: On the historical and geographic diffusion of English family names. *Revue Roumaine de Linguistique* 52 (1-2): 47–77.

\_\_\_\_\_. 2009. On the geography of English family names. *Dialectologia* 3: 73–108.

Withycombe, E. G. 1977. *The Oxford dictionary of English Christian names*. 3rd ed. Oxford: Oxford University Press.

Wrigley, E. A. 1967. A simple model of London's importance in changing English society and economy 1650–1750. *Past and Present* 37: 44–70.