

A Comparative Approach to Identifying the Irish in Long Eighteenth Century London

Author: Adam Crymble (adam.crymble@gmail.com), University of Hertfordshire

Abstract: Historians seeking to identify the Irish have overwhelmingly relied upon nominal record linkage, thus limiting studies to periods and contexts in which corroborating records exist. Surname analysis provides an alternative; a subset of 283 Irish surnames was able to correctly isolate 40% of known Irish individuals across thousands of entries – sufficient for sampling the Irish in demographic studies. This conclusion was based on an analysis of 278,949 names from the London area in the 1841 census, and was tested and refined against 42,248 historical records pertaining to the poor in London between 1777 and 1820.

Keywords: data mining, demographic history, surname, London, quantitative history

ANN ADAMS, alias RILEY, was indicted for a Misdemeanor. Not Guilty.

London Jury, before Mr. Recorder.ⁱ

Was Ann Adams Irish? The above is a London criminal trial transcript from the *Old Bailey Proceedings (OBP)*ⁱⁱ and is a typical example of the type of record that long eighteenth century historians are faced with. For historians interested in a data mining approach to understanding how the Irish in Britain experienced life differently than did the locals, determining the national identity of defendants like Adams is the first step in that process. As digital records prove much easier to analyze than their paper predecessors, the potential for this type of research has come of age – that is, if one can accurately identify the Irish in the records. This study will not answer historical questions about trends amongst the Irish diaspora. Instead it will ask how researchers interested in conducting such studies can identify, with reasonable accuracy, the Irish population within a set of historical textual records.

In too many historical studies, Irish means ‘Irish born’; if a researcher wants to claim Ann Adams was Irish, he or she must find another historical source related to Adams that says she was born in Ireland. This ‘nominal record

linkage' approach was used in Peter King's 2013 study on ethnicity and the Old Bailey. To limit our understanding of the Irish to those with an Irish birthplace is an oversimplification, and is built upon modern ideas about passport eligibility and birth certificates. Yet, according to Ian McBride (1998), this very narrow approach is anachronistic to an eighteenth century worldview, not least because *Irishness* is a social construct rather than a definitive attribute.

In the long eighteenth century, Ireland was home to several different ethnic and religious groups including the descendants of English and Scottish settlers, and who were distinct from the poor Catholics majority, which considered itself the island's original inhabitants. In the 1790s, revolutionary Wolfe Tone (1827, 429) estimated this Catholic majority at seven in ten. Though not an entirely uncontroversial definition of Irishness, it is predominantly this Catholic group that both historians and contemporaries referred to as the 'Irish', and this study will use the same definition, while acknowledging the selection bias that creates.

To identify this group in textual records, historians currently have a number of modes. The present study compares three such methods:

1. Nominal Record Linkage (evidence of an Irish birth in another record)
2. Geographic Keywords (e.g., 'Dublin', 'Connaught', 'Irish')
3. Surname Analysis (has an Irish surname)

This paper looks at the effectiveness of each of those methods for identifying Irish defendants in the *OBP* between the 1801 union of parliaments and 1820. The *OBP* is a text-rich set of records, containing details of all 25,267 defendants tried for felonious crimes at London's Old Bailey courthouse in the period. The *OBP* are typical of the type of text-rich resources historians use and contain similar information about a large number of defendants, making them useful for data mining. The quality of the digitization means that keyword searching should not be impaired by transcription quality. While this case study focuses on the Irish, it could also be applied to other groups of geographically distinct migrants by following the methods outlined in this paper.

1) Nominal Record Linkage

Most studies to date have used only nominal record linkage, which has been revised very little since it was developed on a large scale in the 1960s by

demographers and social historians. It usually involves manually piecing together sets of corroborating archival records, and depends upon the existence of relevant records pertaining to the same individual. The vast majority of research on the Irish in Britain uses this method, focusing on evidence that an individual was born in Ireland in order to classify his or her national identity (see Lees 1979; Boulton 2000; Busted 2001; Linebaugh 2003; King 2013). The *Middlesex Criminal Registers (MCR)*, ledgers that were created to record demographic details about defendants as they entered the gaol system, can be 'linked' to details about the same defendants appearing in the *OBP*.ⁱⁱⁱ Because the *MCR* contain information about the defendant's birthplace, we know that Ann Adams, mentioned above, was Irish-born.^{iv} However, an Irish birthplace proves problematic, both because of the ethnically English people who had Irish birthplaces and controlled the political landscape of Ireland but did not necessarily consider themselves Irish – but also because, as W. J. Lowe (1989) pointed out, it under-represents the Irish population of Britain by roughly half by overlooking the English-born children of Irish immigrants.

Nominal record linkage is popular with historians because the researcher can limit the number of assumptions he or she must make. Peter King's (2013)

article on the Irish at the Old Bailey, 1750-1825 is a perfect example of nominal record linkage in practice. In that study of ethnicity and prejudice in the justice system, King used the *MCR* to identify Irish-born defendants in the *OBP*. He was able to identify 1,188 Irish-born accused between 1791-1805, just over ten percent of the defendants in the *MCR*. Also within that sample were 539 individuals of unknown birthplace and 4,544 who were born in London – some of whom were presumably the children of Irish-born parents.

King has not chosen his date range at random. The *MCR* do not provide birthplace data before 1791 or after 1805. Neither can he turn to the records of the General Register Office to find more Irish defendants, as it did not begin collecting standardized birth, marriage and death records until 1837. The first census to record birthplace was produced in 1841. That means King is without relevant records outside this fifteen-year window. By relying exclusively on nominal record linkage, he is limited in the types of claims he can make. The extent of these limits can be demonstrated clearly by looking at defendants found in the *OBP* who faced trial at the Old Bailey between 1801 and 1820. The *MCR* used by King only provide details about Irish-born defendants in the first five years of this period, as can be seen in Figure 1.^v

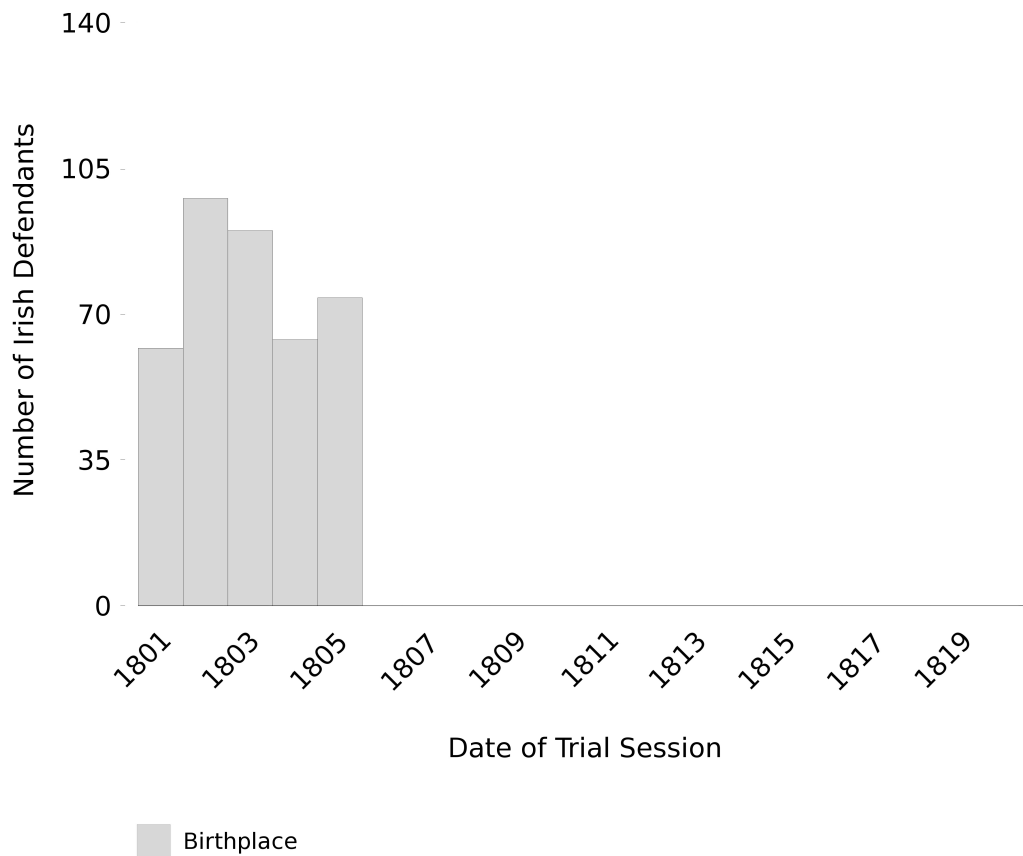


Figure 1: Defendants identified by birthplace.

Irish defendants tried at the Old Bailey by year, identified by birthplace using the *MCR*.

2) Searching for Geographic Keywords

In the past decade, digitization has allowed keyword searching of vast online archives. Words such as the names of principle Irish towns, counties, cities, and 'Irish' or 'Ireland' are obvious choices for finding sources pertaining to the Irish. The historiography of keyword searching as applied to historical texts is slight, and what exists is not promising. A study by Lancaster et al.

(1994; see also Stierman 1997; MaKinster 2002; Kuiper, Volman, and Terwel 2005) showed that basic keyword searching by advanced subject specialists only returned about one-third of the important items in a test database. The result of geographic keyword searching in the present study was much worse.

The most obvious keyword, 'Irish', was incredibly ineffective because national adjectives appear infrequently in the *OBP*. Less than seven percent of transcripts contain national adjectives because the nationality of a defendant was rarely relevant to the trial.^{vi} Even when 'Irish' does appear it very often refers to Irish linen, which was produced in Ireland and heavily imported into England at the time (see McKernan 1994).^{vii} As a portable and relatively nondescript item that could be sold onwards, it was frequently stolen and therefore the term finds its way into the description of charges against the defendant. 'Irish' identified only 56 out of 25,267 defendants as Irish – 8 of whom had already been identified using nominal record linkage and birthplace data.

Other geographic keywords are potentially useful. 'Dublin Poll' is the alias of Mary Ann Caffray, who was convicted in 1816 for pick pocketing.^{viii} The

keyword 'Dublin' in Caffray's alias puts forth a strong case for her Irishness, even without direct birthplace evidence. Yet, as with Irish cloth, Irish place names are problematic keywords. The Irish counties, "Down", "King's", and "Queen's" cause obvious problems. 'Cork' is both a city and county in the south of Ireland and a type of tree used in a range of products. 'Limerick', is both a city and a county in Ireland, but is also a well-known form of poetry and a not infrequent surname.

One might expect these ambiguities to feature rarely in London criminal trials, but this is not the case. 'Cork' returned fourteen useful hits, but four times as many references to corkscrews, wine bottle corks, sheets of cork, and cork cutters.^{ix} Even Westminster's 'Cork Street' appears in the record in 1805.^x Though no references to poetic limericks appear, the term finds useful hits in only 4 of 26 matches.^{xi} By adding the remaining thirty-seven Irish counties, principle towns and cities, and 'Ireland', only 225 trials that used the word in a context related to Ireland were found.^{xii}

Of those 225 trials, less than half (only 98) provided details that could be used to classify the defendant as Irish. Another 29 entries were ambiguous, but probably refer to an Irish defendant. The remaining entries referred to

Irish victims, Irish witnesses, or people who happened to mention Ireland while giving their testimony. These keywords were less than 50% effective at identifying trials containing Irish defendants.

The 225 trials required individual interpretations, and took several days of full time research to unpick, which would not be practical across thousands of potential matches. Despite this, the keywords were only able to classify 98 out of 25,267 defendants. If combined with the defendants already classified as Irish (minus duplicates), there are now 480 Irish defendants, or 1.9% of all defendants, which are still primarily clustered in the years 1801-1805 (Figure 2).

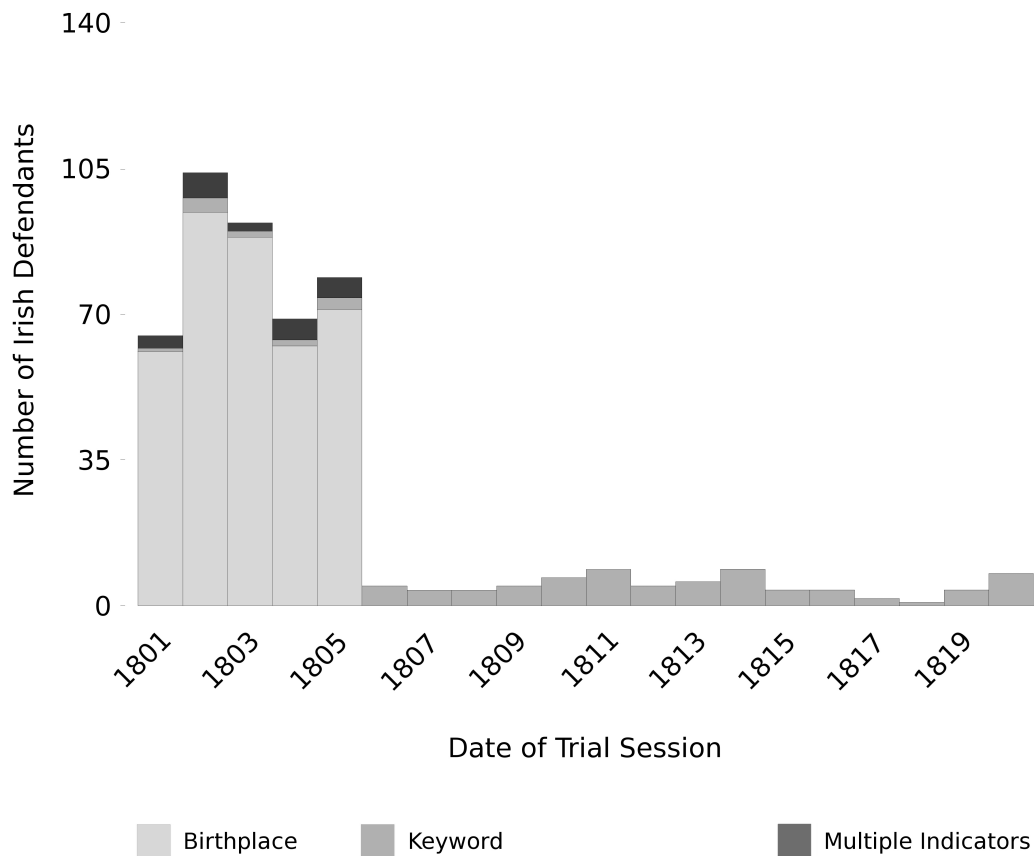


Figure 2: Defendants identified by birthplace and geographic keywords.

Irish defendants tried at the Old Bailey by year, identified by birthplace and thirty-seven geographic keywords, using the *MCR* and *OBP*.

3a) Surname Analysis

The third technique, surname analysis, has long roots amongst onomasticians. A group of epidemiologists led by Huey-Huey Hage touted a surname analysis of the phonebook as an effective way to identify the ethnic Chinese population in Australia (Hage et al 1990; also Choi et al 1993; Hsu-Hage and Wahlqvist 1993). In 1992, Rory Williams applied a similar idea to

Irish surnames, using the Dublin and Aberdeen phonebooks to identify a set of surnames that could reliably identify Irish Catholic descendants living in Scotland (Williams 1992). Both of these studies used modern data for modern ends. In a historical context most discussion of surname analysis has occurred in America. The *American Historical Association* (AHA) used it in the 1930s to calculate the ethnic makeup of the United States at the time of the American Revolution, and has been used in several studies to identify the American-Hispanic population (see Barker and Hansen 1932; Willie 1960; Buechley 1961; Matute-Bianchi 1986). By the 1980s, that original report from the AHA sparked a heated exchange in the pages of *William and Mary Quarterly* (McDonald and McDonald 1980; Akenson 1984; Purvis 1984). The original authors were accused of conducting their analysis with inappropriate data and making poor methodological decisions. After this exchange, surname analysis has since seen little development amongst academic historians, with the exception of work by Smith and MacRaild who continue to use it for identifying the likely county of origin of known Irish individuals, and Longley, Cheshire and Mateos, who work on the regional geography of surnames in Britain (Smith and MacRaild 2009, 2011; Longley, Cheshire and Mateos 2011).

While the vast majority of trial transcripts do not contain national adjectives, every single defendant has a known name, suggesting a surname analysis may be productive. In cases such as that of John Driscoll, a defendant convicted of uttering false bank notes in 1818, a name is virtually all that could point to the defendant's Irishness or lack thereof.^{xiii} A reliable solution that focuses on surnames makes it possible to simplify the problem from one in which a historian must decide if John Driscoll is Irish, into an easier problem: is someone named Driscoll likely Irish? By classifying surnames known to appear in the London area as either Irish or not, it becomes possible to drastically increase the number of probable Irish defendants to the levels seen in Figure 3.

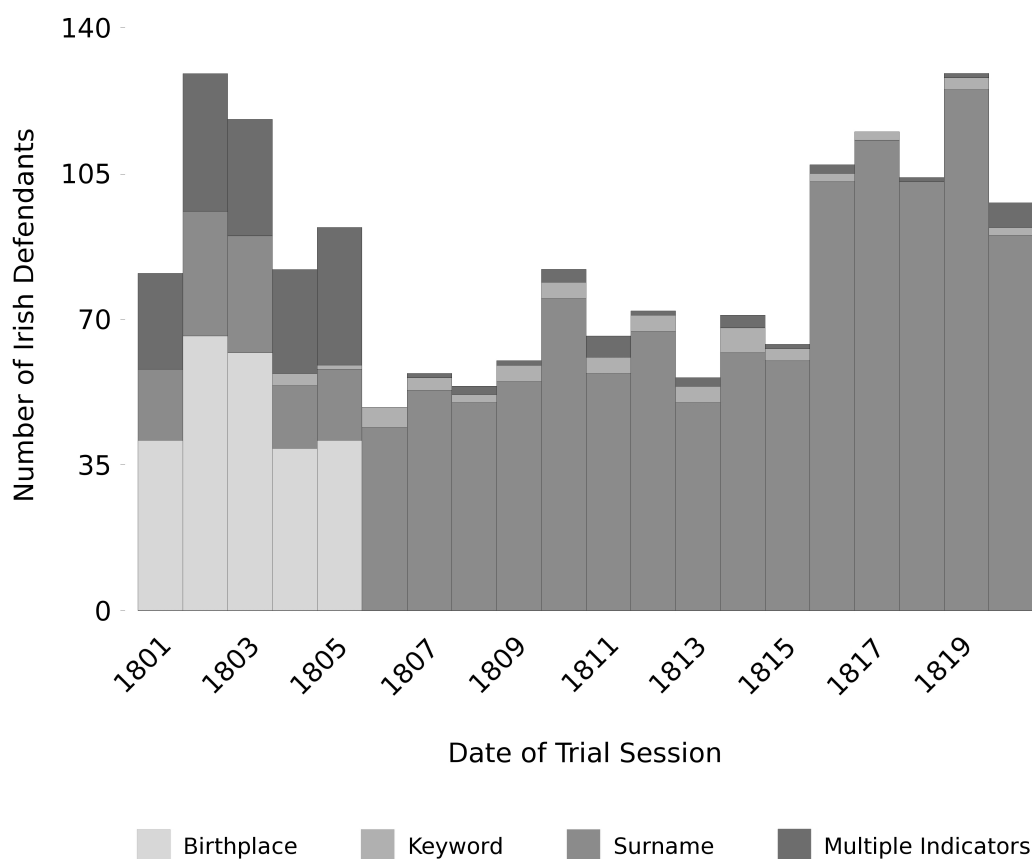


Figure 3: Defendants identified by birthplace, geographic keywords, and surname.

Irish defendants tried at the Old Bailey by year, identified by birthplace, thirty-seven geographic keywords, and surname analysis, using the *MCR*, *OBP*, and the 1841 census.

Criticisms about the accuracy and reliability of surname analysis are important to acknowledge; however, this author believes the works of Willie (1960), Buechley (1961), Matute-Bianchi (1986), Hage et al (1990), Williams (1992), MacRaild and Smith (2009, 2011), and Longley, Cheshire and Mateos (2011), show that surnames can and should be used to identify a subset of

the population when no other alternative is practical and an appropriate set of data can be acquired to determine relevant names.

Compared to keyword searching (98 Irish defendants) and nominal record linkage (388 Irish defendants), surname analysis (1,242 Irish defendants) was by far the most effective of the three methods, bringing the number of probable Irish defendants to 6.7% of the entire defendant population.

Contemporary wisdom and historical evidence suggest a little more than a tenth of criminals in London during this period were Irish, meaning that these combined approaches missed some people. Nevertheless the number of records in the years after 1805 is now at the level where it would be possible for studies like King's (2013) work on ethnicity and prejudice in the justice system to be re-analysed using a far larger proportion of relevant records. Using surname analysis, this author is confident that John Driscoll, mentioned at the opening of this section, was probably Irish.

3b) Surname Analysis – Deriving an Appropriate Method

One of the biggest criticisms of surname analysis came from McDonald and McDonald (1980, 184, 191), who complained that the AHA 1931 study had associated inappropriate names with specific ethnicities. In particular, the

pair argued that the names used to represent the English were overwhelmingly Welsh. They also noted that the names used to identify the Irish were based on Robert Edwin Matheson's study from 1894, using data from 1881, nearly a century later than Barker and Hansen's 1790 study date, and were therefore of questionable reliability. This complaint highlights the importance of using an appropriate list of names for identifying the Irish. Matheson's *Special Report on the Surnames in Ireland* (1894), Patrick Woulfe's *Irish Names and Surnames* (1922) or Edward MacLysaght's *The Surnames of Ireland* (1969) include thousands of Irish surnames, and are seemingly obvious places to turn for a list of names. However, it is important to acknowledge that these works were not created for the purposes of identifying the Irish in long eighteenth century sources. Both Matheson and MacLysaght's lists included the name Smith, which is of course English, perhaps quintessentially the most English name there is. Lists of Irish names often include Smith because the name was the fifth most prevalent surname in Ireland – representing an Anglicized form of MacGowan, or son of the smith (see Matheson 1894; Woulfe 1922; MacLysaght 1969). Matheson's top 100 Irish surnames includes a number of problematic names besides Smith, including Wilson, Clarke, Brown, Martin, White, and Robinson, all of which are common British names and obviously problematic for this purpose. Any

list of names used to identify the Irish in London in the long eighteenth century therefore must be bespoke, based on the realities of that place and time.

Malcolm Smith and Donald MacRaild's 2009 paper, 'Paddy and Bidy no more', built upon earlier work by Williams (1992) and used the Index of Griffiths' Valuation and the 1881 census of England and Wales in an attempt to generate a 'Irish Surnames Corpus'. However, the vast waves of Irish migrants that entered Britain during the Famine era of the 1840s and 50s dramatically changed the demographics of both countries, and to use a national surnames corpus based on data from the 1880s may skew the results of a local study set in the long eighteenth century. To avoid this, the present study focuses on the 1841 census of England and Wales, which offers a solution. Like modern censuses, everyone living in England or Wales on a single night in June 1841 was recorded along with, for the first time, the person's place of birth. Like any primary source, the census has its limits as a historical record, including its underreporting of young children (Wrigley and Schofield 1989, 589-591), and the not infrequent listing of strangers as 'NK' – shorthand for 'not known'. Despite these and other minor shortfalls, for the vast majority of people in the country, the census records delineate

who was born in England and Wales (classified together), Ireland, Scotland, Foreign Parts, or the local area. While 1841 is beyond most historians' definition of the end of the long eighteenth century, the distance in time does not impact the census' usefulness for determining if a surname can accurately predict the Irish-born in the preceding century, as discussed below.

Using all 278,949 records of adult males in the hundred of Ossulstone from the 1841 census, each unique surname was analyzed.^{xiv} The hundred is a former administrative district in Middlesex County that was home to approximately one million people by the mid-nineteenth century. It excluded Westminster, the City of London proper, as well as Southwark, immediately south of the river opposite the city. Nevertheless, it included large parts of the metropolis' newly developing urban areas such as St. Marylebone, St. Pancras, St. Giles-in-the-Fields, and the Docklands, as well as fairly vast swaths of countryside and rural villages within striking distance of the capital. The combined urban and rural nature of the hundred makes it particularly useful, by capturing a wider variety of people than could be found in purely urban parishes without going beyond the practical reach of the metropolis.

These more than quarter-million records were sorted by surname, revealing 58,964 unique surnames. The vast majority (89%) of names have no Irish connection at all, while 5% have *only* Irish-born members.^{xv} There is clearly a core of names that may be useful as a means for identifying the Irish.

To determine the best names, each name was then given an 'Irishness' score based on the following formula to calculate the percentage of adult males with that name born in Ireland:

$$\text{Irishness} = \text{Irish-Born Adult Males} / \text{Total Adult Males} * 100$$

Women were not included because of the added complexity of maiden names, which may have undermined confidence in surnames as an appropriate method of identifying the Irish. Further study is needed to determine the efficacy of surname analysis on women. Names were also given an 'English/Welsh' score using the same methods – noting that the census does not distinguish English from Welsh names, which in this case effectively means not Irish. Figure 4 shows that there are a small subset of

surnames that have strong connections to the Irish in London, and which could be used to classify defendants as probably Irish or probably not.

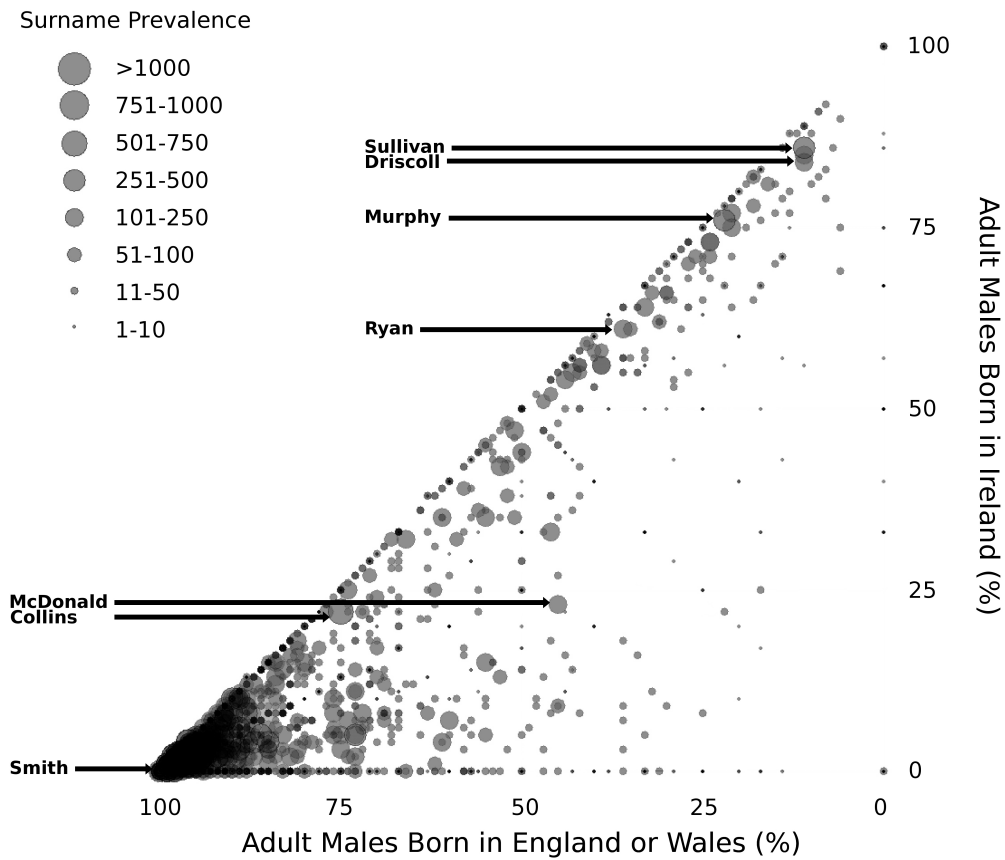


Figure 4: Surnames of the Hundred of Ossulstone.

The 58,964 unique surnames found in the Hundred of Ossulstone in the 1841 census, with each name plotted as a percentage of adult males born in England/Wales versus born in Ireland.^{xvi}

Figure 4 shows each surname that appeared in the census represented as a circle. More common surnames are larger. The names were plotted using the

formulae described above. The further up the y-axis a circle appears the more Irish it is. The further to the left a circle appears the more English or Welsh it is. Most names have no, or at best a very weak Irish connection, seen in Figure 4 as a dark cluster in the bottom left corner. Without evidence to suggest otherwise, defendants with these very English and Welsh names are probably not Irish, while those that show strong Irish connections are the most likely matches. Amongst this set of common Irish names are Murphy, Sullivan, McCarty, Mahoney, Donovan, and Driscoll, which allows us to conclude that John Driscoll mentioned above was likely Irish despite an absence of any direct evidence in his trial transcript.

3c) Surname Analysis – Who is Irish Enough?

6,632 names have at least one Irish-born member, and 3,026 names contain only Irish-born members. Which are Irish enough? Should the list of reliable names be limited to those with scores greater than 80%? 50%? 10%? To determine this, the results of the initial census analysis were compared against the 5,969 criminals recorded in the *MCR* from 1801 to 1805 who had a known place of birth. These are the same records used by King in his nominal record linkage paper, and were selected for this study because they represent a way to validate the surnames analysis using records that have

already been used by other historians for a similar task. They included a known subset of 572 Irish individuals that could be used to determine what Irishness score accurately identified the largest proportion of the Irish, and at what point adding more names became counter-productive.^{xvii}

Surnames were grouped by their Irishness score starting with scores greater than 95%, adding the surnames with scores of 5% less for each subsequent test. The first group contained only names with an Irishness score of 95 or greater; the second group contained names with a score of 90 or greater, and so on, until all names from the 1841 census sample were included. Each of these cohorts was then tested against the *MCR* to see what proportion of Irish-born defendants each captured (see Figure 5). This means Murphy, which has an Irishness score of 75.8 was added in the fifth test, at which point it becomes possible to see how many more entries names like Murphy impacted the number of Irish-born defendants identified.

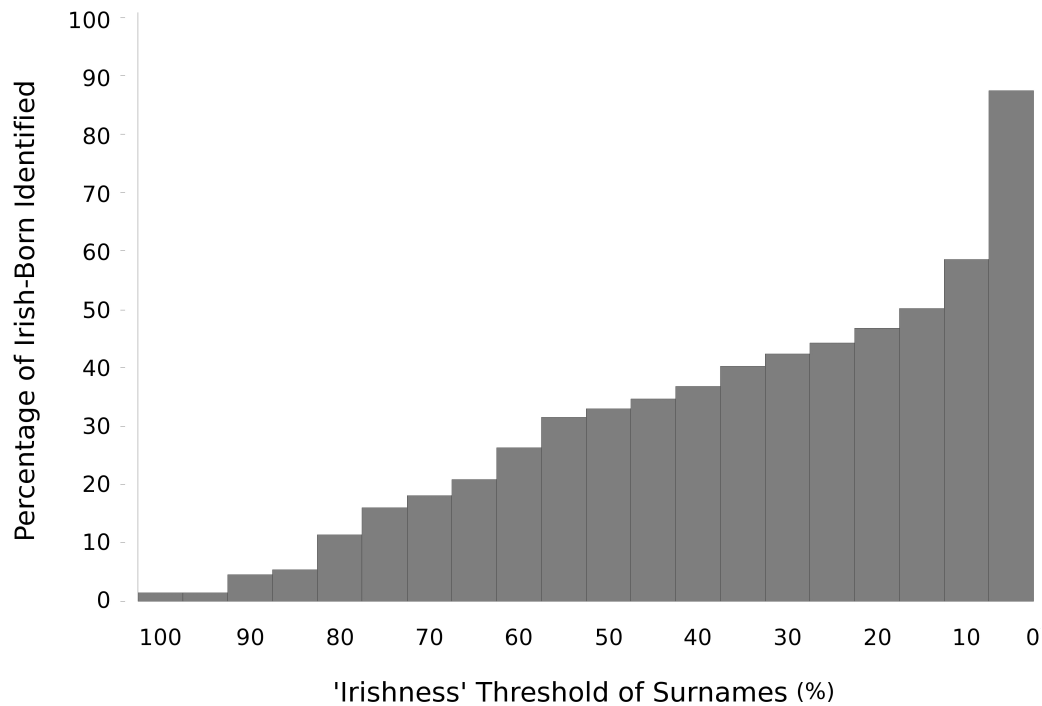


Figure 5: Efficacy of surname analysis.

The percentage of Irish-born defendants identified using Irish surname lists with decreasing 'Irishness scores'.

Figure 5 shows that the names with incredibly high Irishness scores (left-hand columns) identified few Irish-born defendants. At the other extreme, every surname in the 1841 census sample captured slightly fewer than 90% of the Irish-born in the *MCR*. One-off spelling of surnames mean complete accuracy using an onomastic approach is unlikely. Nevertheless, this graph shows that it is possible to capture a reasonably large proportion of the Irish

defendants – probably between ten and fifty percent – using surnames alone. That is a considerable step forward from the results of using keywords and nominal record linkage.

Unfortunately the best lower confidence threshold is not clear from Figure 5 alone. An Irishness score of 5% would capture 60% of the Irish-born population; but such a low score would also increase the number of false positives. The goal is to capture as many Irish-born defendants listed in the *MCR* as possible, while ensuring that any gains in the Irish-born group are not offset by false positives.

If we apply these surname cohorts to the *MCR*, it may seem as if there are only two outcomes, a good result or a false positive. However, the complex nature of Irish identity means that there are in fact four outcomes.

1. **Good result:** A surname correctly classifies an Irish-born individual
2. **Ambiguous outcome:** It is unclear if the person is Irish
3. **False positive:** A surname incorrectly classifies an individual as Irish
4. **Missed result:** An Irish-born individual is not identified.

Good Result

A good result accurately identifies an Irish-born individual. If the *MCR* confirm that livestock thief Charles O'Connor was born in Ireland and the list of surnames includes *O'Connor*, then that is a good result.^{xviii} There are 572 individuals in the *MCR* that were born in Ireland and that could be classified as a good result (of whom 388 were also tried in the Old Bailey). These are the entries we would want our surname cohorts to find.

Ambiguous Outcome

An ambiguous outcome is someone born in the London area (London, Middlesex, or Surrey). There were 2,335 individuals in the *MCR* in this category (of whom 74 will ultimately be classified as 'Irish' based on their surname, meaning that these Londoners overwhelmingly represent the non-Irish population). Individuals such as 18 year-old John Murphy who was born in Middlesex County and arrested in 1802,^{xix} were classified as ambiguous because they were probably the children of Irish immigrants, and as W.J. Lowe wrote in 1989, were an important part of culturally Irish communities in Britain. This means we would not want to classify John Murphy as a false positive (without further evidence to support that conclusion), but we equally do not have blatant evidence that he is definitively Irish. In order to

compare surname analysis as closely as possible to the standards set out by nominal record linkage advocates, entries like John Murphy have been classified as ambiguous so that they do not over- or underemphasize the number of false positives.

False Positive

A false positive is one that identifies a person incorrectly or probably incorrectly as Irish. These are the leftovers from the first two groups, and 2,462 individuals in the *MCR* could potentially be a false positive. Most of these people are clearly English (or foreigners) and if the surname cohorts are effective, then these individuals will not be identified as Irish. Charles Donovan, a black man born in Africa was convicted of stealing handkerchiefs in 1805 is an example of someone we would hope not to include. Though *Donovan* is an extremely Irish name, in this case it is most likely a slave name.^{xx} London had a large black community in the eighteenth century, many of whom were shipped to the Sierra Leone colony in 1787; however, some did remain in London, though few if any had uniquely African surnames (see Hitchcock 2007, ch5 p30). Similarly, the very Irish surname *Murphy* would identify Norfolk-born Rebecca Murphy as Irish.^{xxi} Based on her place of birth and what is known about Irish immigrants, it is quite probable that Rebecca's

maiden name is English. She may have married an Irish man, and though it is entirely possible that Rebecca lived a very Irish life in London, for the sake of prudence while determining the accuracy of names on the list, she was deemed mis-classified.

Missed Result

Missed results are known Irish-born individuals that were not identified as Irish with a surname analysis. These missed results will be discussed in the next section.

*

The best set of surnames identifies the greatest number of good results while minimizing false positives. Using the individuals in the *MCR*, this was determined by testing each of the Irish surname cohorts described above to find the point at which adding more names starts generating more false positives than good results (see Figure 6).

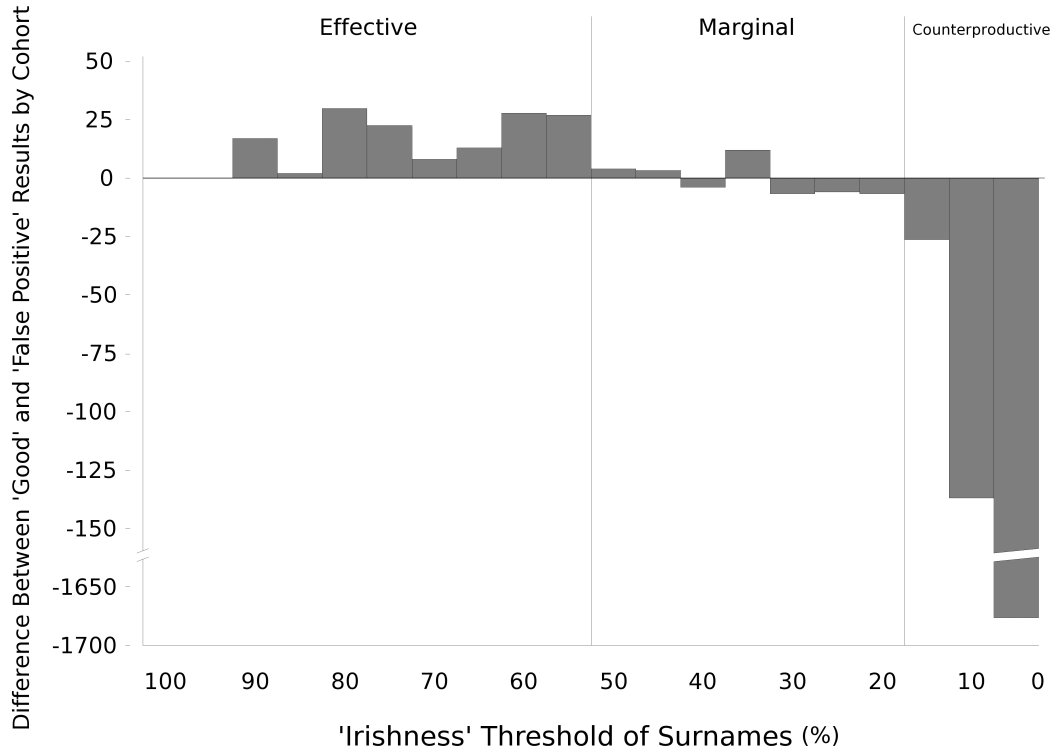


Figure 6: Identifying the threshold of effective surnames.

The difference between 'good' results and 'false positive' results for each surname listed in five percentage point cohorts, compared against known birthplace data from the *MCR*.

Entries above the zero line show more 'good' results than 'false positives', and the reverse is true below the line. This means that for every bar above the line, adding the additional surnames in that cohort was productive. With this in mind, Figure 6 shows clear positive results down to an Irishness score of 50. Between 50 and 35, the results hover around zero, being neither good

nor bad. At the 35-30 range there is a slight improvement – a product of names that are one-third Irish that prove better than not – before falling off dramatically. While names with an Irishness score of 5 or greater would accurately identify 60% of the Irish-born defendants as seen in Figure 5, using such a low cutoff threshold would introduce far more false positives than would be desirable (see the second last column of Figure 6).

The ‘best’ Irishness score threshold to use as a lower limit is subjective, as a case could be made for either 50 or 30. To err on the side of caution, this study decided to use 50 as the threshold of ‘Irish enough’. At this cutoff, the Irish surnames list included 4,094 uniquely spelled surnames, making it possible to identify 1,686 *probable* Irish defendants in the *OBP* out of 25,267 defendants (6.6% of defendants), as seen in Figure 3 above, and with some further refining, a ‘false positive’ rate of only 0.5%, across more than 5,000 entries in the *MCR*.

3d) Surname Analysis – Refining the Results

Because the classification of so many of these names was based on just a few entries pertaining to rare names, there is room to refine the list. *Zagorski* has an Irishness score of 100, as do *Moncko*, *Tinling*, and *Wimphey*. 2,556 Irish

surnames (62%) appear only once in the 1841 census. This is in sharp contrast with names such as *Murphy* that appears 348 times, and in which one can be much more confident. These 2,556 one-off names are also incredibly ineffective, identifying only three Irish-born individuals in the *MCR*. This is less effective than ‘Limerick’, which pointed to four individuals. As can be seen in Figure 7, these uncommon names do not pull their weight.

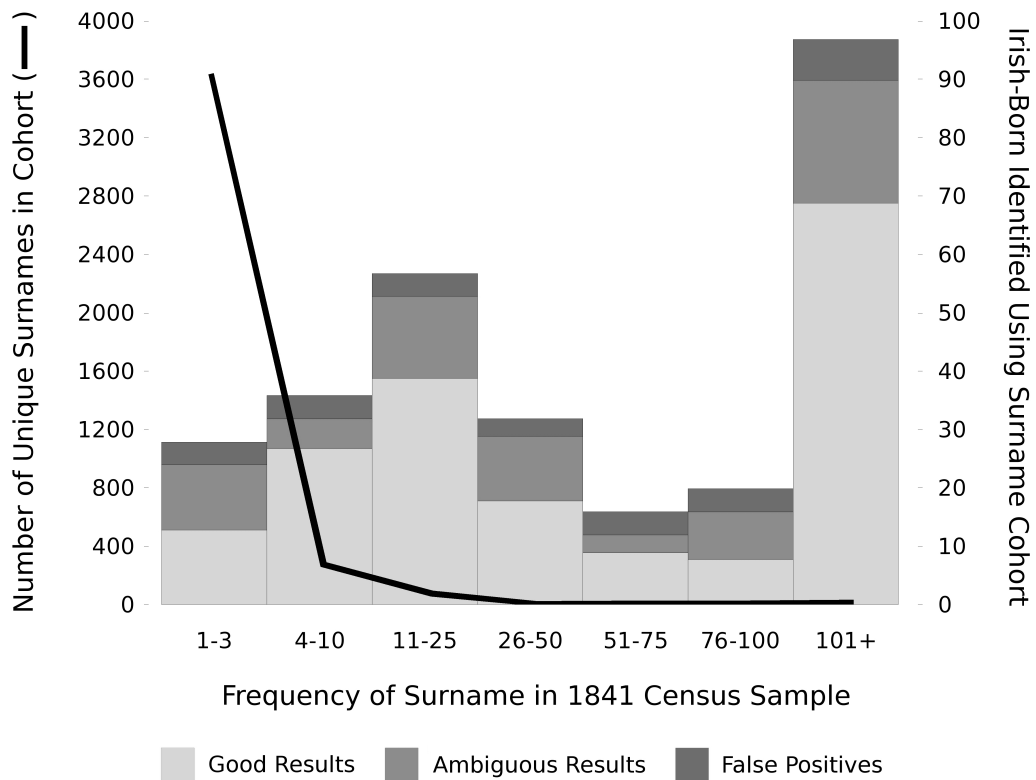


Figure 7: Efficacy of common versus uncommon names.

The rate at which common versus uncommon names produce Good, Ambiguous and False Positive results, when used to identify Irish-born individuals in the *MCR* from 1801-1820, excluding entries for which birthplace is unknown or illegible.

To generate Figure 7, surnames were split into one of seven groups by frequency in the 1841 census. The bars show how many Irish-born individuals were identified using each group of surnames, splitting the results into good, ambiguous, and false positives as above. The most effective names are those that appeared in the census at least 101 times. The black line on the graph shows the number of surnames in each group, indicating that a very large number of uncommon names (the left column) are almost useless, while a handful of very common names (the right column) are tremendously effective.

By eliminating all names from the list that appear less than four times in the census, only 13 good matches from the set of known Irish-born defendants are lost, while reducing the number of surnames from 4,094 to 457. This is worthwhile because entries with the lowest confidence can be removed from the list of surnames without dramatically reducing its efficacy. This more reliable list of 457 surnames, can accurately identify 30% of the Irish-born defendants in the *MCR* (down from 32%). This in itself is not particularly

helpful, since the *MCR* tell the reader which individuals from within its own pages are Irish-born; however, 30% is a useful measure of how effective one might expect the surname analysis to be when seeking to identify the Irish in other similar record sets.

The remaining 457 names can be further refined to adjust for phonetic variations: *Sullivan*, *Sulevan*, *Sulivan*, *Sulivon*, and *O'Sullivan* all appear as unique names on the list. The Soundex algorithm, which helps identify homonyms, patented by Robert C. Russell in 1918, can further refine the list.^{xxii} The algorithm is not perfect and will incorrectly classify both 'McCaffrey' and 'McIvor' as homonyms; however, despite these occasional failings, using Soundex with some close reading can refine the original list of 457 names down to 283 Irish root surnames and their variations.

*

Despite these successes, many entries were missed. The list missed 402 Irish-born accused in the *MCR* because seemingly English or Welsh names appear to be far more common than one might have expected, affecting 172 of the missed entries – 30% of the entire Irish-born group in the *MCR*. These names

include Smith, Jones, Brown, and Davis, which are British names that may be tied to the descendants of the ethnically British Protestant population in Ireland. The surnames on this list are more useful for identifying Irish Catholics and anyone using a surname analysis should keep this in mind. Another factor is Anglicization. Writing in 1853, James MacGrady suggested many Irish surnames had undergone transformations as they moved from oral Irish to written English (McGrady 1853). According to Nollaig Ó Muraíle a large proportion of Irish words had never been written down prior to the Ordnance surveys of the 1830s and in many cases the first time an Irish term found its way onto paper was at the hand of an Englishman (Ó Muraíle 2001). This introduced a level of Anglicization that causes problems for the present task (MacNeill 1915). Mac Giolla Padraig (son of Patrick) became *Kilpatrick*, *Kirkpatrick*, *MacFeteridge*, *MacFadzean*, and *MacFadden*, and may even have become *Pattison* or *Patterson* amongst settlers who ended up in England (see Lawlor 1925; Ryan 1917).

It was also not uncommon to find two different spellings of a single individual's name. This is because the person writing the name often heard it spoken rather than saw it on paper. This explains why some Irish-born defendants such as Edmund Nowlan, who was tried for theft in 1802, had an

Irish name in one record, and a more geographically ambiguous name (Noulan) in another record.^{xxiii} What name he actually used is anyone's guess – he may have been unaware of the difference – but it is easy to imagine how his Irish accent could be to blame for the misinterpretation.

Unfortunately, very little can be done to classify individuals with Anglo/Welsh names using surname analysis, because to do so would involve wrongly classifying dozens or perhaps hundreds of English people as Irish. It may be possible to improve results by considering distinctly Irish forenames as an added measure of confidence. For example, names such as Patrick, Michael, Daniel, Timothy, Dennis, Jeremiah, Cornelius, and Owen appear to be Irish according to a similar analysis of the 1841 census. However, with only eight such names to draw upon, forenames were not incorporated into this study. Based on the limits of the current methodology, readers can expect that roughly one third of relevant individuals will be classifiable and one third of relevant individuals will be missed.

The remaining one third of outstanding individuals have names that fall into one of three categories:

1. Non-Irish surnames
2. Names that do not appear in the 1841 census
3. Names that are borderline Irish – those with Irishness scores between 30 and 50, and those uncommon names removed in the previous section.

The remaining names include obvious homonyms of common Irish names, which can increase the number of matches without straying from names that already show strong Irish scores. Based on this new evidence, *Bryan* can be classified as a variant of *Brian*, as can *Callahan* of *Callaghan* and *Donahough*, *Donohew*, *Donohugh* of *Donoghue*. As a result, an additional 58 Irish-born entrants can be identified that had previously been missed.

Homonyms did not appear equally across all three groups, scarcely affecting the group of non-Irish names, revealing only four additional entries. It worked very well for names that did not appear in the census, classifying 28% of those entries. Likewise, the group of borderline Irish surnames saw even better results, with 57% now classifiable.

Homonym matching using Soundex identified an additional 37 variant spellings of common Irish surnames and made it possible to increase the effectiveness of the surnames list in this case by another 58 good results, 8 ambiguous results and only 10 false positives, meaning adding these names to the list was worthwhile. The final result of this expanded list of Irish surnames identified 228 out of 572 Irish-born individuals in the *MCR*, representing a good result rate of 39.8%. The ambiguous results accounted for 74 individuals (13%), and the false positives were only 32 (5%).

3e) Surname Analysis – Effectiveness on Earlier Records

Despite the promising results of surname analysis using the 1841 census on the *MCR* (1801-1805), the criticisms of the McDonalds', as they relate to the distance in time between the data collection (1841) and the study period (1801-1820) remain. Can a surname analysis based upon London in 1841 be used to identify the Irish decades earlier? With high mortality rates and low life expectancy most of the people who appeared in the *OBP* between 1801 and 1820 had died or moved elsewhere before census was taken.

The *Middlesex Vagrancy Removal Records* contain the details of 11,016 individuals or family groups removed from Middlesex County between 1777

and 1786.^{xxiv} They provide a unique opportunity to test the efficacy of a surname analysis of 1841 on individuals living approximately six decades earlier. Like the *MCR*, the vagrancy records are ledgers of individuals who had been arrested – in this case for ‘wandering and begging’ without a legal settlement in the area. The parish had the right to expel these individuals to prevent undue economic strain on local resources (Hitchcock, et al, 2014). These vagrancy removal records generally included the same classes of character that were typically found on trial at the Old Bailey – namely the poor and labouring classes. Of the 11,000 people in these records, 2,559 (23%) were sent back to Ireland.^{xxv} If the surnames identified as Irish via surname analysis are able to reliably identify these Irish vagrants from sixty years earlier, then one can be confident in the connection between Irishness and surnames throughout the long eighteenth century.

Even though the sample size is much smaller – c.11,000 versus c.250,000 – the surnames on the list were able to identify one quarter of the Irish-born with 83% accuracy.^{xxvi} Recall, the same list had been able to identify 30% of the Irish-born entries in the *Middlesex Criminal Records*, so the results are encouraging. Not all names had the same or even similar Irishness scores across the two sets of records. Some names such as Ryan increased in their

reliability, rising from 65% to 93% further back in time. The top fifteen names improved on average 20% when used on the earlier set of records. This is almost certainly because fewer English-born descendants of Irish immigrants lived in the eighteenth century than in the nineteenth, and shows that the distance in time between the 1841 census and the eighteenth century does not impact the usefulness of an onomastic approach in identifying the Irish in London.

4) Conclusion

This paper started with a challenge: how best to identify Irish defendants in a set of historical records to end up with the most comprehensive set of results possible. The solution to this problem in this case was threefold, using nominal record linkage, keyword searching, and surname analysis. The results can be seen in Figure 8, which includes the various refinements discussed above. Using these combined indicators of those defendants who appeared before an Old Bailey magistrate in the early decades of the nineteenth century, 1,712 (6.7%) were identified as Irish.

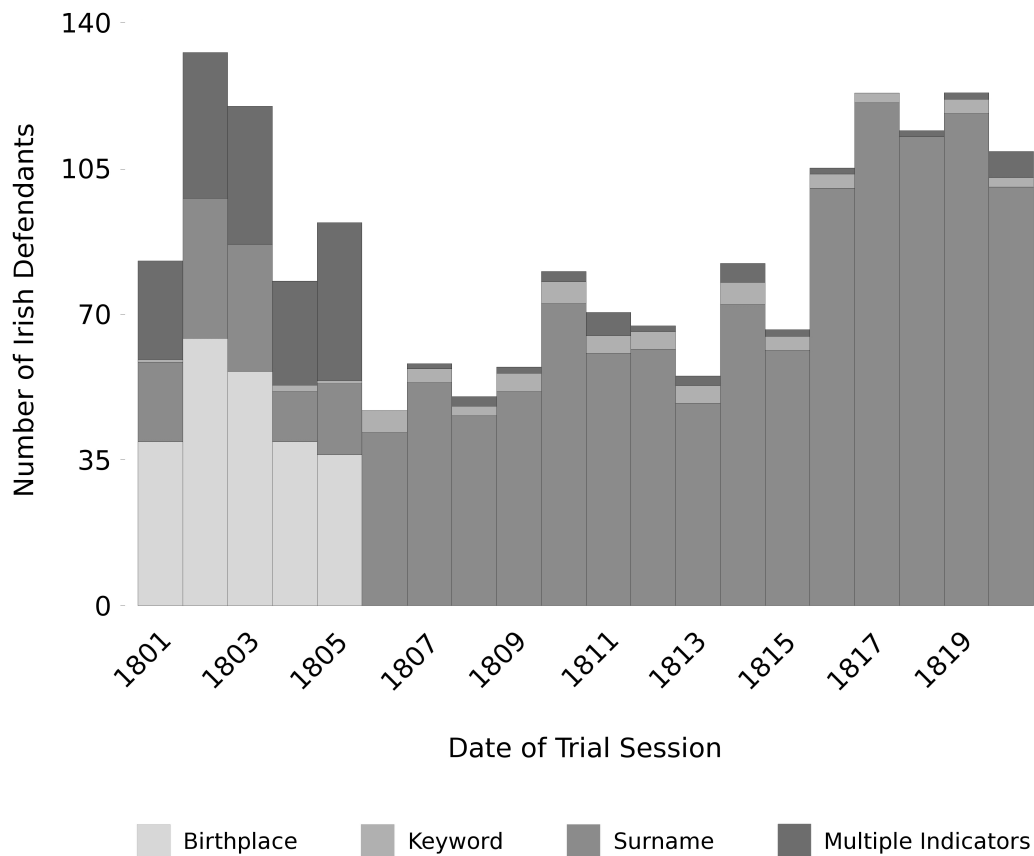


Figure 6: The final set of identified defendants.

The final 1,712 probable Irish defendants tried at the Old Bailey, graphed by year and identifiable by birthplace, relevant geographical keywords, and surname analysis using the *MCR*, *OBP*, and the 1841 census.

When it is possible to use a set of corroborating records that can identify the Irish with a high degree of confidence, nominal record linkage is an academically rigorous approach. Using the *MCR*, 387 probable Irish defendants were linked to their *OBP* trial transcript, providing a reliable set of records. Nominal record linkage's primary weakness comes from a lack of

corroborating records for the years 1806 to 1820. The task of linking records is also incredibly laborious, requiring thousands of individual judgments, making this type of research expensive. The output is also inflexible, tying together two sets of documents, but *only* two sets of documents. No generalized results can be taken away and applied to another set of sources.

Keyword searching identified a handful of relevant entries, but was unfruitful when applied to this specific set of records, identifying only 98 probable Irish defendants. Almost half of these were identifiable using other means, making this too an expensive approach considering the results. Each keyword match had to be manually interpreted and often required reading several pages of text to make an informed decision. The high number of false positives made this very time-consuming, taking almost thirty minutes of work for each good result. Unlike with nominal record linkage, these results are reusable as these keywords could be applied to another set of historical records such as a newspaper database, in which the keywords may prove more useful because of the different nature of the content.

Finally, surname analysis made it possible to identify a further 1,242 individuals. These extra results span all twenty years of the study period and

may represent as much as 40% of the actual Irish defendants in London.

Through validation, the set of Irish surnames created has been shown reliable, producing far more good results than false positives.

The surname analysis of this particular group of Irish migrants is also entirely reusable. Though it too took a tremendous amount of labour – several times more in fact than the nominal record linkage or keyword searching approaches – for those working on London in the eighteenth century, that work is now complete and can be repurposed by any researcher (see Appendix I for the full set of Irish root surnames). I would suggest that anyone interested in this group of predominantly Irish Catholics in London or the southeast between c. 1700 and 1850 now has a tool at hand with which to efficiently identify a relevant sample of materials for further study. That sample will naturally make it easier to identify the Irish Catholic population because of that group's common history and historic naming patterns that were distinct from the English and Scottish immigrants to Ireland. This methodology is particularly useful for anyone studying historical demography, focusing on the Irish, but could conceivably also be applied to individuals of other nationalities. It is most appropriate for large-scale analyses that look for trends across thousands of similar sources, such

as evidence of Irish involvement in crime, poverty, or in generating migration statistics. It would be inappropriate for studies dependent upon close readings of a small number of individuals, in which a few anomalies could drastically reduce the accuracy level. It is also difficult to say how useful names are for identifying Irish married women, who traditionally took their husband's name; however, evidence suggests these women tended to marry fellow-nationals so the concern may be moot (Busteed and Hodgson 1996, 139-153, 148). There will always be scope for improving the results; nevertheless, the present study has shown that these three methods of searching together were able to identify roughly 40% of the Irish in the records under review, which many statisticians and demographers would consider a generous sample size. Though there were thousands of records under investigation, only 32 false positive matches were found. With the option for quickly repurposing the names on the list, new doors are open for studying the Irish in London, and the connection between surnames and Irishness in the eighteenth century is firmly established.

References:

Akenson, Donald H. 1984. "Why the accepted estimates of ethnicity of the American people, 1790, are unacceptable," *The William and Mary Quarterly* 41:102-119.

Barker, H.F. and M.L. Hansen. 1932. "Report of the committee on linguistic and national stocks in the population of the United States," *Annual Report for the Year 1931* 1. Washington.

Boulton, Jeremy. 2000. "Irish migration to England in the late middle ages: the evidence of 1394 and 1440," *Irish Historical Studies* 32:1-21.

Buechley, Robert W. 1961. "A reproducible method of counting persons of Spanish surname," *Journal of the American Statistical Association* 56:88-97.

Busteed, M.A., and R.I. Hodgson. 1996. "Irish migrant responses to urban life in early nineteenth-century Manchester" *The Geographical Journal* 162:139-153.

Busteed, M.A. 2001. *Patterns of Irishness in nineteenth century Manchester*.
Manchester.

Choi, B.C.K., A.J.G. Hanley, E.J. Holowaty, and D. Dale. 1993. "Letters to the editor," *Epidemiology* 4:86.

Hage, Bridge Huey-Huey, R. Graeme Oliver, John W. Powles, and Mark L. Wahlqvist. 1990. "Telephone directory listings of presumptive Chinese surnames: an appropriate sampling frame for a dispersed population with characteristic surnames," *Epidemiology* 1:405-408.

Hitchcock, Tim. 2007. *Down and out in eighteenth-century London*. London.

Hitchcock, Tim, Adam Crymble, and Louise Falcini. 2014. 'Loose, idle and disorderly: Vagrant removal in late eighteenth-century Middlesex', *Social History* 39:509-527.

Hsu-Hage, B. Huey-Huey, and M.L. Wahlqvist. 1993. "Letters to the editor," *Epidemiology* 4:87.

King, Peter. 2013. "Ethnicity, prejudice and justice. The treatment of the Irish at the Old Bailey 1750-1825," *Journal of British Studies* 52:390-414.

Kuiper, E., M. Volman, and J. Terwel. 2005. "The web as an information resource in K-12 education: strategies for supporting students in searching and processing information," *Review of Educational Research* 75:285-328.

Lancaster, F.W., C. Elzy, M.J. Zeter, L. Metzler, and Y. Low. 1994. "Searching databases on CD-ROM: comparison of the results of end-user searching with results from two modes of searching by skilled intermediaries," *RQ* 33:370-386.

Lawlor, H.C. November 1925. "The origins of some common surnames (concluded)," *The Irish Naturalists' Journal* 1:35-36.

Lees, Lynn Hollen. 1979. *Exiles of Erin: Irish migrants in Victorian London*. Manchester.

Linebaugh, Peter. 2003. *The London hanged: crime and civil society in the eighteenth century*. London.

Longley, Paul A., James A. Cheshire, and Pablo Mateos, 2011 "Creating a regional geography of Britain through the spatial analysis of surnames" *Geoforum* 42:506-516.

Lowe, W.J. 1989. *The Irish in mid-Victorian Lancashire: the shaping of a working-class community*. New York.

MacGrady, James. 1853. "Irish surnames: their past and present forms" *Ulster Journal of Archaeology* 1:117-119.

MacLysaght, E. 1969. *The surnames of Ireland*.

MacNeill, John. 1915 [1911]. "Part 3. place-names and family names" *Proceedings of the Royal Irish Academy. Section B: Biological, Geological, and Chemical Science* 31:42.

MaKinster, J.G., R.A. Beghetto, and J.A. Plucker. 2002. "Why can't I find Newton's third law? case studies of students' use of the web as a science resource," *Journal of Education and Technology* 11:155-172.

Matheson, R.E. 1894. "Special report on surnames in Ireland with notes as to numerical strength, derivation, ethnology, and distribution" in *Twenty-ninth detailed annual report of the registrar-general of marriages, births, and deaths in Ireland*. Dublin.

Matute-Bianchi, Maria . 1986. "Ethnic identities and patterns of social success and failure among Mexican-descent and Japanese-American students in a California high school: an ethnographic analysis," *American Journal of Education* 95:233-255.

McBride, Ian. 1998. "'The common name of Irishman': Protestantism and patriotism in eighteenth-century Ireland" in *Protestantism and national identity: Britain and Ireland, c. 1650-c.1850*, edited by Tony Claydon and Ian McBride, 236-261. Cambridge.

McDonald, Forrest, and Ellen Shapiro McDonald. 1980. "The ethnic origins of the American people, 1790," *The William and Mary Quarterly* 37:179-199.

McKernan, Anne. 1994. "War, gender, and industrial innovation: recruiting women weavers in early nineteenth century Ireland," *Journal of Social History* 28:109-124.

Ó Muraíle, Nollaig. 2001. "Some thoughts on matters onomastic" *Journal of the Galway Archaeological and Historical Society* 53:25.

Purvis, Thomas L. 1984. "Why the accepted estimates of ethnicity of the American people, 1790, are unacceptable: commentary," *The William and Mary Quarterly* 41:119-135.

Ryan, John. May 1917. "Irish family names (continued)" *The Irish Monthly* 45:331-336.

Smith, Malcolm, and Donald M. MacRaid. 2009. "Paddy and Bidy no more: An evolutionary analysis of the decline in Irish Catholic forenames among descendants of 19th century Irish migrants to Britain," *Annals of Human Biology* 36: 595-608.

Smith, Malcolm, and Donald M. MacRaild. 2011. "The origins of the Irish in northern England: an isonymic analysis of data from the 1881 census," in *Irish Identities in Victorian Britain*. edited by R. Swift and S. Gilley, 24-49. London.

Snell, K.D.M. December 1991. "Pauper settlement and the right to poor relief in England and Wales," *Continuity and Change* 6:375-415.

Stierman, J. 1997. "Efficient strategies for searching historical databases," *OAH Magazine of History* 11: 45-50.

Tone, Theobald Wolfe . 1827. "On the present state of Ireland, delivered to the French Government, in February 1796" in *Memoirs of T.W. Tone*. London.

Williams, Rory. 1992. 'Can data on Scottish Catholics tell us about descendants of the Irish in Scotland?' *New Community* 18: 296-309.

Willie, William W. Jr. 1960. "The Spanish surname criterion for identifying Hispanos in the southwestern United States: a preliminary evaluation," *Social Forces* 38:363-366.

Woulfe, P. 1922. *Irish names and surnames*.

Wrigley, E.A., and R.S. Schofield. 1989. *The Population History of England 1541-1871: A Reconstruction*. Cambridge.

Appendix 1

283 Root surnames that have been shown to be reliable for identifying probable Irish individuals in eighteenth century London

Defined as a name with an Irishness score of at least 50 and appearing in the 1841 census sample at least four times, shown in alphabetical order, with punctuation removed and all names in lower case. If you use this appendix in your research, please cite this article:

ahern, alton, anglin, ashe, baggot, barry, bermingham, bernick, boran, boyan, breen, brenan, brian, broderick, broman, brosmann, bryden, bucke, buckly, bulger, burke, byrne, cabe, caffrey, cahill, calden, callaghan, calnon, canty, carby, carmody, carney, carrigan, carroll, casey, cassidy, cathey, caulfield, cavanagh, caveney, clancy, cleary, cline, cochran, cocklin, coffee, cokely, conard, condon, connell, connor, conroy, cooney, corey, corley, costello, cotter, coulter, coyle, crawley, creagh, crone, cronin, curley, curran, cusack, cushen, daby, dacey, daly, darcy, darey, darley, daun, delany, desmond, dillon, dimsey, divine, doherty, dolan, dolphin, doming, donnelly, donerson, donoghue, donovan, doody, doran, dowd, dowman, downey, doyle, driscoll,

duby, duffy, duggan, duggins, dundon, dunne, dunworth, durkin, dwyer, early, egan, enright, fagan, farrell, feeny, finn, finnigan, fitton, fitzgerald, fitzgibbon, fitzmaurice, flaherty, flanagan, fling, flynn, foley, gafney, galey, gallagher, gannon, garry, garvey, gathers, gaugh, gavesk, gavin, gilligan, glynn, gogin, gorman, grady, grogan, gulliford, guman, hackley, hagerty, hallahan, halman, hanigan, harnett, harrigan, harrinton, haye, hefferman, helton, helvy, hely, hennesy, hickey, hogan, holey, hollan, hollifield, horan, hormel, hormley, hosty, hurley, hynes, jaffe, joice, kean, kearney, keefe, keegan, keenan, kellard, keller, kelly, kenedy, kennelly, kenny, keogh, kilsey, lally, lamp, lawler, Leahy, leary, lewman, linehan, loury, loves, lovet, lowney, lynch, lynot, maabie, macguire, madden, magee, magner, maher, mahoney, maker, maloney, mara, mardock, marney, mcauliff, mccabe, mccauffrey, mccarthy, mccathy, mccormack, mccooy, mcdermott, mcdonnell, mcdonough, mcgrath, mchugh, mcivor, mcjames, mckenna, mckensie, mcloughlin, mcmahon, mcmanus, mcnamara, mcquin, megan, mellett, molley, mooney, moran, moriarty, morrice, moylan, mulcahy, mulligan, mullin, murphy, myan, myler, neagle, neil, nowlan, nowland, ohara, ormsby, pursell, quill, quin, quinlan, rainy, readen, rearden, reddy, regan, reily, roach, ronan, roney, rourke, ryan, sayce, scanlan, scannell, scully, seary, shannon, shaughnessy, shea, sheen, sheridan,

spelling, sulison, sullivan, sweeney, taff, tierney, tobin, tool, toomey, tracy,
troy, walsh, warling, whelan

Notes

Acknowledgements: The author would like to thank Angela Kedgley, Janice Turner, Amanda Goodrich, Ian McBride, and Willard McCarty for commenting on earlier versions of this manuscript, and to the attendees of the British History in the Long Eighteenth Century seminar at the Institute of Historical Research (2012), and the participants at the London Irish conference at the University of Warwick (2012) for their questions and comments on the paper.

ⁱ *Old Bailey Proceedings Online* (www.oldbaileyonline.org, version 7.1, 24 September 2013), October 1804, trial of Ann Adams, alias Riley (t18041024-74).

ⁱⁱ The *Old Bailey Proceedings* can be consulted at the *Old Bailey Online*. Tim Hitchcock, Robert Shoemaker, Clive Emsley, Sharon Howard and Jamie McLaughlin, et al., *OBO, 1674-1913* (www.oldbaileyonline.org, version 7.1, April 2013).

ⁱⁱⁱ The *Middlesex Criminal Registers* (National Archives, HO 26) is available electronically through Ancestry.com.

^{iv} National Archives, HO 26, Middlesex; 1804. Piece: 10; Page: 3.

^v The *MCR* contain 3,858 individuals who also appear in the *OBP*. Of these, 388 (1.5 percent) are Irish born out of a total 25,267 defendants.

^{vi} French (2,988 matches), English (2,760), Irish (1,438), German (1,104), British (872), American (608), Scotch (584), Spanish (550), Welsh (519), Dutch (499), Italian (492), Indian (308), Russian (275), Scot (234), African (132) Portuguese (129), Prussian (122), Australian (121), Belgian (100), Turkish (94), Austrian (81), Swedish (78), Chinese (75), Scottish (55), Canadian (51), Japanese (42), Danish (41), Norwegian (36), Hungarian (24), Flemish (19), Finnish (2), Total: 14,433 matches out of a possible 197,745 criminal trials. This calculation is via the *OBP* web-search interface.

^{vii} In the twenty years between 1801 and 1820, *Irish* appears just 187 times in the *OBO*, 118 of which are national adjectives describing a person – though not necessarily describing the defendant.

^{viii} *OBP*, September 1816, trial of Mary Ann Caffray (t18160918-82).

^{ix} *OBP*, Searched for all offences where the transcription matches 'cork', between 1801 and 1820.

^x *OBP*, April 1805, trial of Elizabeth Otway (t18050424-125).

^{xi} *OBP*, Searched for all offences where the transcription matches 'limerick', between 1801 and 1820.

-
- ^{xii} The following keywords related to Irish place names were searched in the *OBP*: Irish, Ireland, Dublin, Belfast, Waterford, Antrim, Armagh, Carlow, Cavan, Clare, Cork, 'County Down', Donegal, Fermanagh, Galway, Kerry, Kildare, Kilkenny, Leitrim, Limerick, Londonderry, Longford, Louth, Mayo, Meath, Monaghan, Roscommon, Sligo, Tipperary, Tyrone, Westmeath, Wexford, Wicklow, Ulster, Munster, Connacht, Leinster.
- ^{xiii} *OBP*, September 1818, trial of John Driscoll (t18180909-48).
- ^{xiv} An adult male was defined as a male of at least 20 of age. Children were omitted from the analysis because of the high numbers of London-born children with Irish parents..
- ^{xv} 3,026 of the 58,964 names in the sample have *only* Irish-born members. 6,632 names have at least one Irish-born member.
- ^{xvi} Derived from "1841 UK Census", 6 June 1841, *Ancestry.co.uk* <http://search.ancestry.co.uk/search/grouplist.aspx?group-1841uki> (Accessed 2010-2011).
- ^{xvii} HO 26, the National Archives.
- ^{xviii} National Archives, HO 26, Middlesex; 1801-1805. Piece: 8; Page: 99.
- ^{xix} National Archives, HO 26, Middlesex; 1801-1805. Piece: 8; Page: 90.
- ^{xx} *OBP*, October 1805, trial of Charles Donovan (t18051030-42).
- ^{xxi} National Archives, HO 26, Middlesex; 1801-1805. Piece: 8; Page: 80.
- ^{xxii} R.C. Russell, US Patent 1,261,167, April 2, 1918.

^{xxiii} *OBP*, April 1802, trial of Edmund Nowlan (t18020428-74); National Archives, HO 26, Middlesex; 1801-1805. Piece: 8; Page: 95.

^{xxiv} Data available from: Adam Crymble, Tim Hitchcock, Louise Falcini 2014, 'Vagrant lives: 14,789 vagrants processed by Middlesex county 1777-1786', (2014); DOI: 10.5281/zenodo.13103. The *Middlesex Vagrancy Removal Records* were originally part of *Sessions Papers – Justices' Working Documents*; 'Home Office: Criminal Registers, Middlesex', 1805-1791, HO 26, The National Archives, Kew. Names were only counted once per family group.

^{xxv} Settlement was not necessarily a marker of origin or birthplace. One could gain settlement in a new parish through a number of channels, including renting a property worth more than £10 per annum, possessing a freehold estate, through an apprenticeship, yearly service to a master, by holding certain offices, paying parish poor rates, through parentage, and in the case of women, through marriage. However, since vagrancy records refer primarily to the very poor, being sent to Ireland in this instance is a useful indicator of Irishness. For more, see K.D.M. Snell, "Pauper settlement and the right to poor relief in England and Wales," *Continuity and Change* vol. 6, no. 3 (December 1991), 375-415.

^{xxvi} The surnames list correctly identified the Irish returnees in 642 out of 776 matching individuals. There were 2,559 Irish vagrants in the whole sample.