

The Welsh diaspora : Analysis of the geography of Welsh names

Project undertaken for the Welsh Assembly

Richard Webber, Visiting Professor,
Department of Geography, University College London

1 : Objective of the project	1
2 : The use of names to infer identity	1
3 : Data sources	2
4 : Family names	2
5 : Personal names	3
6 : Concentrations of Welsh family names around the world	4
7 : Concentrations of Welsh family names in Great Britain	5
8 : Concentrations of Welsh family names within the United States, Australia and New Zealand	5

1 : Objective of the project

The purpose of the project summarised in this report is to improve the ability of the Welsh Assembly to identify the location of the overseas Welsh community within the English speaking world.

In particular the project seeks to identify the level of concentration of people of Welsh origins in England, Scotland, the United States, Australia and New Zealand. It also seeks to enable the Welsh Assembly to identify where within these countries people of Welsh origin are most likely to be found.

The project also supplies the Assembly with a list of personal and family names most strongly associated with the Welsh community, together with a measure of the strength of that association. This list will enable the Assembly to select from population registers those individuals whose names are most likely to be of Welsh origin.

2 : The use of names to infer identity.

This project is based on the assumption that it is both possible and effective to use the elements of a person's name as a proxy for their origin.

To this end the project makes use of the 'Origins' database. This contains information on some 280,000 different family names and some 120,000 different personal names, each of which have been coded according to the cultural / ethnic / linguistic group to which it is believed that they are most likely to belong.

Clearly at the level of the individual it is not possible to infer their ethnicity or region of origin with a guaranteed 100% certainty just from his or particularly her name. In an area as ethnically diverse as the British Isles the ancestry of many individuals incorporates more than one community of origin. This is true to an even greater extent in the English speaking diaspora.

Nevertheless, notwithstanding the increased level of mobility in recent centuries, there are certain personal and family names which continue to be disproportionately found in Wales, and more particularly in the least Anglicised regions of Wales. People bearing Welsh family names, we have found, are disproportionately likely also to bear Welsh personal names. From measurements of this sort it is possible, in addition to allocating a name to an ethnic group, to identify the extent to which that name is to a greater or lesser extent associated with 'Welshness'.

Given that the geographical concentration of such names varies between different parts of England, Scotland, the United States, Australia and New Zealand, and that these variations are systematic rather than random, we believe that it is not unreasonable to assume that this is the result of people of Welsh origin having settled disproportionately more in certain of these regions than they have in others. These higher than average values will reflect both a higher proportion of the population with some Welsh ancestry

and a higher proportion of an individual's ancestry originating in Wales. Combined this shows the level of Welsh ancestral blood in the local population.

3 : Data sources

The statistical analysis on which this report is based has used a file created and maintained by the information company Experian International, whose headquarters are in Nottingham.

The name of the file is 'Consumerfile'. Consumerfile contains names and addresses, together with postcodes, for all UK electors who have not 'opted out' of their names being released for marketing purposes. The file also includes a large number of records from other sources, including Companies House, which publishes the names and addresses of company directors. Lifestyle questionnaires, shareholder files and lists of County Court Judgments also contribute to the 46 million records held on the file.

This count is slightly larger than the national adult population on account of some persons having associated records at more than one address.

It should be noted that the versions of the electoral register on which this particular study has been based were captured in 1996 and 1998. This preceded the opt out legislation.

The data for the United States originates from telephone files. These cover only 50% of the adult population. The Australian and New Zealand files are based on registers of electors.

It is also important to appreciate that whereas in Great Britain we have access to information on all names, in the United States our data pertains only to family names with more than 200 occurrences. In New Zealand we only hold data for family names which have more than 100 occurrences in Great Britain.

4 : Family names

Of the 280,000 family names recorded in the Origins database there are 6,461 that are identified as Welsh.

A number of different criteria may have been used to determine that these are Welsh. For family names with fifty or more occurrences on the Great Britain electoral register the determining criterion is likely to be that there are relatively more occurrences of that family name in Wales than in England or Scotland relative to the total number of named individuals on Consumerfile.

Less frequently occurring names may be designated Welsh on the basis on their structure. For example the family name 'Ab-Owain', of which there are three occurrences only in Great Britain, would be designated Welsh whether or not it had occurrences in Wales. This is because its form or structure. Likewise most double barrelled names, such as

Griffiths Jones, where one or more of the name elements are Welsh, are also designated as Welsh even if there are few or no Welsh occurrences extant.

These 6,461 names account for 3,020,224 out of the 46,366,087 entries in the Experian database (6.52%). However in Wales these names account for 718,040 out of the 2,057,463 entries (34.90%). In other words family names which have been classified as Welsh are over five times more common in Wales, pro rata to population, than in Britain as a whole

By measuring the proportion of the holders of each individual name that are resident in Wales it is therefore possible to measure variations between family names in the extent to which they are indicative of Welsh origins.

Likewise when we measure the proportion of the holders of each individual name that also have a Welsh personal name we also able to measure variations in the extent to which they have retained a Welsh approach to personal naming.

Taking the more common Welsh family names, those with 5,000 or more occurrences in Great Britain, we find that Lloyd, Llewellyn, Griffith, Harries and Rees are the ones which are most likely to be associated with Welsh personal names. By contrast Howell, Ellis, Preece, Johns and Griffin are Welsh names which are least strongly identified with Welsh personal names.

These variations in the Welshness of Welsh names reflects the fact that some names, such as Richards, may have originated in other parts of Britain as well as in Wales whilst others may have originated in the more Anglicised areas of Wales where the sense of Welsh identity is weakest. Being able to discriminate the level of Welshness of Welsh names is useful when targeting communications campaigns.

The results of this analysis are found in file one.

5 : Personal names

Of the 120,000 personal names recorded in the 'Origins' system 3,136 are identified as being Welsh.

The criteria for identifying these as Welsh are similar to those used to allocate ethnicity to the family names. These names tend either to be disproportionately found in Wales or to be associated with records on the Experian database with Welsh family names.

Just as the Welsh family names can be scored according to the extent to which they are associated with Welsh personal names, so too the Welsh personal names can be scored according to the extent to which they are associated with Welsh family names.

Geraint, Huw, Gwyn, Alun and Evan are examples of the popular Welsh personal names that are disproportionately associated with Welsh family names. By contrast Melvyn,

Edgar, Trevor, Rowland and Phillip are personal names which, though relatively more common among people of Welsh origin than not, nonetheless only marginally meet the criteria for being considered Welsh.

Although the proportion of family names that are Welsh is broadly similar to the proportion of personal names that are Welsh, Welsh family names tend to have many more occurrences on average than Welsh personal names. Thus whilst 3,020,224 (6.50%) of records in Great Britain have a Welsh family name, only 811,300 (1.75%) have a Welsh personal name. However the extent to which these personal names are geographically concentrated in Wales is very similar to that found among family names. People in Wales are just over five times more likely than in Britain as a whole to bear a Welsh personal name.

The results of this analysis are found in file two.

6 : Concentrations of Welsh family names around the world

We have seen that 34.9% of people in Wales bear a Welsh family name. This compares with an average of 5.44% in England and of only 1.64% in Scotland. In the United States 3.77% of population bear a Welsh family name. Australia contains a slightly higher proportion of people with a Welsh family name, 4.12%, than the United States. In New Zealand the proportion of people with a Welsh name is slightly higher still, 4.75%.

The difference in the concentration of the population of Welsh origin in these countries is broadly similar to the difference in the concentration of the population with British names. In the United States a much higher proportion of the population have names which originate from elsewhere than Britain than is the case in Australia. New Zealand has the lowest proportion of names of non British origin.

Based on these figures we can make the following estimates of the size and distribution of Welsh ancestry as follows:

Estimated size of the Welsh diaspora			
Region	Population 2006 (millions)	% Welsh ancestry %	Total Welsh ancestry (millions)
Wales	2.903	34.9	1.013
Rest of GB	57.315	5.3	3.038
United States	296.498	3.8	11.267
Australia	20.456	4.1	0.839
New Zealand	2.903	4.7	0.136
			16.293

The term 'Welsh Ancestry' is used not to indicate the proportion of people who are Welsh. Rather the term should be considered as indicating the average share of Welsh genetic ancestry among individual people. For very few will this be either 100% or 0%. The majority of people will have shares of Welsh genetic ancestry around the mean for the region in which they live

On this basis we might reasonably suppose that, disregarding countries such as South Africa and Argentina not covered by this analysis, two thirds of all Welsh genetic ancestry is to be found among citizens of the United States and just under 7% still residing in Wales. These differences are largely accounted for by the very much larger size of the US population, not by the high proportion of Welsh ancestry in its genetic make up.

7 : Concentrations of Welsh family names in Great Britain

File three contains information for each Great Britain postcode sector (eg N6 4) showing the total number of family names on the Consumerfile database and the proportion of these names that are of Welsh origin.

The mapping of this file shows quite distinct variations within Wales in the proportion of the population of Welsh origin. Typically the lowest rates are found in South Pembrokeshire, in the Vale of Glamorgan and along the Clwyd coast.

Proportions of people of Welsh origin in England vary by linear distance from the Welsh border with highest concentrations west of the M5 and M6 motorways. High concentrations are evident in Oxford, Bristol and Liverpool. Generally there are few people of Welsh origin in Scotland or the North East of England.

8 : Concentrations of Welsh family names within the United States, Australia and New Zealand

For both the United States and Australia we are able to establish how the proportions of the population with Welsh names varies across the country using, in the case of the United States, the US states, in the case of Australia, the Australian states and Capital Territory.

These patterns are shown in maps one and two. These are based on data held in files four and five.

In both maps the concentrations of Welsh names are shown in the form of what are known as 'index' values, a convention whereby a concentration equal to the national average is given a value of 100, a value of 50 indicating a concentration of Welsh names one half the national average and a value of 200 indicating a concentration of Welsh names two times the national average.

It is evident from the two maps that the population of Welsh origin is very much more unequally distributed in the United States than it is in Australia. In the United States the proportion of the population of Welsh origin ranges from 9.5% in South Carolina to 1.1% in North Dakota. Typically people of Welsh origin are concentrated in the mid Atlantic states, the Carolinas, Georgia and Alabama and in Appalachia, West Virginia and Tennessee. By contrast there are relatively fewer people of Welsh origin in New England, which is mostly Irish, the northern mid West, settled mostly by continental Europeans, and the South West, populated disproportionately by Hispanics and Asians.

The geographical pattern of the Welsh in the United States is a reflection in part of the era of maximum Welsh emigration, which was earlier than that of the Irish, continental Europeans, Hispanics and Asians. It may well be that the concentration of Welsh names in Appalachia is explained by the importance of coal mining to that region's economy, just as Cornish names are disproportionately found in metal mining regions of the United States and Australia.

It is worth noting that in the United Kingdom there appears to be a disproportionate concentration of Welsh and Scottish family names in neighbourhoods with high proportions of residents of Caribbean descent. A likely reason for this pattern is the tendency for slaves to have adopted the family names of slave masters, a large proportion of whom appear to have been Welsh or Scottish. It is possible that a similar practice applied in the American South and that many of the residents of the mid Atlantic states who bear Welsh names are Black.

Variations in the concentration of Welsh names in Australia are not as evident as they are in the United States. Here the highest concentration, which is Tasmania, is only 17% above the Australian national average. The lowest, which is in Victoria, is just 8% below the national average.

In Australia as in the United States Welsh family names are relatively more common in the states which accepted earlier settlers.

Map three (which is based on data in file six) shows the distribution of people of Welsh origin in New Zealand. In the case of New Zealand we have not been able to obtain counts of family names for the 15 provinces into which the country is divided. However for each name we do hold the name of the province with the highest concentration together with the extent to which that concentration exceeds the New Zealand average.

We have therefore counted for each province the number of Welsh names for which it has a higher concentration than does any other province. We have also measured the average 'index' values for the names associated with that province. Multiplying these two values we have been able to come up with a value of the 'Welshness' of each province.

Variations in the concentration of the Welsh are much greater in New Zealand than they are in Australia with two provinces, Marlborough and Taranaki, having twice the national

average proportion of people of Welsh origin and three, Southland, Canterbury and Manawatu-Wanganui, having concentrations less than half the national average. I can offer no explanation for this pattern.



