

REVIEWS

Watts, <i>The Cambridge Dictionary of English Place-Names</i> (Hough)	133
Kristendommens indflydelse på nordisk navngivning, ed. Sigmundsson (Fellows-Jensen)	142
Coates and Breeze, <i>Celtic Voices English Places</i> (James Room, <i>The Penguin Dictionary of British Place Names</i> (Nicolaisen)	147
Randall, <i>Place-Names: How they define the World—and More</i> (Nicolaisen)	150
Sims-Williams, <i>The Celtic Inscriptions of Britain. Phonology and Chronology, c.400–1200</i> (Okasha)	156
Mills, <i>A Dictionary of London Place Names</i> (Cullen)	158
Pierce, <i>Place-Names in Glamorgan</i> (Cane)	160
Also received (Smart)	164
Bibliography for 2003 Carole Hough	165
Work in Progress	167
Notices:	181
English Place-Name Society	90
Essay Prize	126
Scottish Place-Name Society	180

The Forenames and Surnames from the GB 1998 Electoral Roll Compared with those from the UK 1881 Census

Ken Tucker

Carleton University, Ottawa

1. Abstract

The results of the analysis of the Electoral Rolls for the GB (hereafter ER) are presented in the form of tables of the most popular forenames and surnames, and graphs of population against number of names, and population against rank, for forenames and surnames. All results show marked power law relationships.¹

The results are consistent with those of previous studies of contemporary USA, and Canada, and are directly compared with those of the UK 1881 Census for England and Wales (hereafter Census).² A drifting breakpoint in the distribution of forenames which separates the fashionable from the rare forenames is identified for both females and males. The impact of the new immigrants is discussed. The apparent loss of surnames since 1881 is identified, as is the large growth of numbers of different surnames (hereafter types of surname) since then.

¹ A Note on Power Laws, Logarithms, and Names: If all the words in a book were ordered by descending frequency and the frequency plotted against order on logarithmic-logarithmic paper, the graph would be a descending straight line. The relationship of the words, one with another, is known as Zipf's Law. An excellent description of this law and the plethora of cases to which it applies may be found at: <http://linkage.rockefeller.edu/wli/zipf/>. The mathematical expression for such a line is of the form: $y=Cx^r$ which should be read as: 'y equals C times x raised to the power r', where y is the dependant variable, x the independent variable, and C a constant equal to the y value when $x=1$. Such relationships, whether between words or names, are thus known as a 'power law relationship'. If the names are in descending frequency then r will be negative. The choice of linear or logarithmic scales is made to ensure clarity of the resulting curve.

² D. K. Tucker, 'An analysis of the forenames and surnames of England and Wales listed in the UK 1881 Census data', *Onoma* (2004) (in press).

2. The Data

The data set used in this study was collected in 1997 for the 1998 ER, and was graciously provided in 2003 by Experian Ltd through the good offices of Richard Webber. Experian collects and consolidates the data from the 463 local authorities who compile the electoral register for their particular region. The data set is one of the increasingly more available country-wide databases of surnames and forenames, others being telephone listings and census data for example.

A word on classes, forms, types, and tokens: there are three classes in this article: surnames, female forenames, and male forenames. A name form, such as *Christie*, may appear in one or more classes. The surname type *Christie* has 21,428 entries or tokens, the female forename type *Christie* has one token, and the male forename type *Christie* has 14,163 tokens in the 1998 ER data set.

There were 47,054,569 entries in the data supplied, one entry per registered voter. It is axiomatically also the number of tokens (that is, entries) for both surnames and forenames. The field structure for the data is: (first) *forename*, *initial* (of second forename), *surname*, *postal code*, and *gender*. The *initial* and *postal code* data have not been used in this analysis. *Gender* is not from the local authorities but is allocated by the Experian Voting Roll team. It is allowed three states: male; female; and unknown. Surprisingly, some forenames make use of all three categories.

No attempt has been made to correct misspellings to avoid 'correcting' something that was correct in the first place. This data set, in common with those of the USA, Canada and the UK (1881) suffer from type inflation due to typographical errors.³ All forenames and surnames in the received data set were in uppercase only. The initial processing yielded 805,168 surname types and 368,864 forename types for the 47 million tokens.

³ D. K. Tucker, 'Distribution of forenames, surnames, and forename-surname pairs in the United States', *Names*, 49 (2001), 69-96; *idem*, 'Distribution of forenames, surnames, and forename-surname pairs in Canada', *Names*, 50 (2002), 105-32; *idem*, 'An analysis of the forenames and surnames of England and Wales'.

2.1. Forenames

The forename types include short forms, diminutives, pet forms and suchlike including names like *Jacey* but exclude initials like *JC*. The forename types were split into gender categories: male, female and unknown; the type counts were 126,726; 154,408, and 206,355 respectively. As some forenames are used by both sexes we would expect that the sum of the types by gender would be greater than the count of forename types: 368,864 but not as great as the total of 487,489 so there appears to be about 25% inflation of the number of forename types due to the gender issue alone. Examination of the data shows that both well established forenames, and forenames from newer immigrants from Africa and Asia, have high *unknown* scores as well as scores for both sexes as shown in Table 1.

Table 1: A Selection of Forenames Illustrating the Forename Gender Inflation Issue

Forename	Male	Female	Unknown
Annalisa	16	263	138
Armarjh	997	897	963
Bronwyn	15	648	266
Christoper (sic)	395	17	202
Geofrey (sic)	292	12	172
Imtiaz	596	259	514
Olatokunbo	26	26	52
Yuksel	28	27	28

Table 1 shows twenty-four name types whereas there are really far fewer types, ignoring any typographical errors. The forename types are not only overstated by the gender issue, they are also overstated by the number of typographical errors. Using *Christopher* as an example we see that there are at least 279 typographical errors (or are they typos? e.g. *Cristofer*) for the forename *Christopher* as shown in Table 2. (The list is not claimed to be exhaustive.)

choice, but our parents are free, in most places, to name us whatever, but there are over twice as many surname types than there are fore-name types.

2.3. Data Summary

The number of types, and tokens (albeit inflated), by class, used in the analysis are given in Table 3.

Table 3: The Types and Tokens by Name Class for the 1998 Electoral Roll Data

Class	Sex	Types	Tokens
Forenames	Female	154,408	23,971,370
Forenames	Male	126,726	21,835,108
Forenames	Unknown	206,355	1,248,091
Surnames	Not App	781,782	47,054,569

3. The Results

The results are presented in graphical and tabular form with supporting text explanation where necessary. Two types of graphs are each presented for female forenames, male forenames, and surnames. All graphs have logarithmic scales as it is the nature of names to have power law relationships. Care must be taken in interpreting the results as the Census Data is for England and Wales only, whereas the ER data is for the whole of GB. The voting age in 1997 was eighteen years so in order to prepare for the following year's ER, data was collected for seventeen year olds and older.

The lack of Scottish data in the Census data may bias some results such as the under-representation of Scottish forenames and surnames. Nevertheless many valuable points are yielded in the comparisons.

3.1. The Tables

3.1.1. Measuring Popularity

Some government agencies provide information about the most

popular male and female names given to new-borns in the past year. In the UK this information may be found at <http://www.statistics.gov.uk/>. Since there is a difference between being 'born and being alive, it is essential to differentiate between these surveys, and popularity as it is being measured in this article. The difference is of degree but the resulting names can be very different and misleading.

The new-borns' names are measured initially over a year, for not every new-born whose name is captured survives. What is measured in a census are the names of everybody who, at that time, is alive. For the sake of argument we can assume the oldest people to be a hundred years old. We thus have a hundred years of naming decisions minus all those who have died. If death rates are not equal across all segments of society, and if naming choices are tied to societal segments, as some suspect, then we have another problem. Ideally, when dealing with long-term naming trends one would like the numbers for the century. The ER data was generated in 1997 for the 1998 Electoral Roll and would contain the names of all people still living who were born between about 1898 and 1980. Table 4, extracted from <http://www.statistics.gov.uk/> gives a snapshot of the popular names for new-borns for some years during this period. It should not be surprising that the most popular names shown in the ER for 1998 are not those for the new-borns of 1980.

Table 4: The Most Popular Forenames in the UK by Year

Gender	Rank	1904	1934	1964	1994	1999	2003
Males	1	William	John	David	Thomas	Jack	Jack
	2	John	Peter	Paul	James	Thomas	Joshua
	3	George	William	Andrew	Jack	James	Thomas
	4	Thomas	Brian	Mark	Daniel	Daniel	James
	5	Arthur	David	John	Matthew	Joshua	Daniel
Females	1	Mary	Margaret	Susan	Rebecca	Chloe	Emily
	2	Florence	Jean	Julie	Lauren	Emily	Ellie
	3	Doris	Mary	Karen	Jessica	Megan	Chloe
	4	Edith	Joan	Jacqueline	Charlotte	Jessica	Jessica
	5	Dorothy	Patricia	Deborah	Hannah	Sophie	Sophie

There are two principal reasons for this. One is the declining absolute popularity of the most popular name, and the other is the increasing volatility of the most popular name. In Galbi Table 1, Mary was the most popular female forename for 1925 at 3.7% of the population, and Margaret for 1944 at 4.5%.⁴ For males it was John in both cases at 7.3% and 8.3% respectively. (These rises are blips in the overall declining rate for females from 23.9% (Mary) for 1800 to 3.4% (Emily) for 1994, a sevenfold reduction, and for males from 21.5% (John) for 1800 to 4.2% (James) in 1994, a fivefold reduction.) Here again we see that the most popular forename is male with the transition occurring at the new-borns level about 1880. The increasing volatility of the most popular forename for new-borns is shown in Table 4 for example.

Thus the most popular name in the nation, as measured below, is more heavily influenced by naming decisions of the parents of the older members of the population and the current fragmentation is not easily seen except in documents like *The Baby Name Countdown*, which lists the 116,308 forename types given to babies in Canada and the USA mainly in 1994 and 1995.⁵

3.1.2. Forenames — Female

The top hundred forenames for females are listed in Table 5 which includes thirty-two of the top hundred from the Census. *Margaret* is the clear favourite. No new immigrant forename appears in the top hundred. *Michelle* made the most significant improvement going from a count of two in the Census to a count of 138,156, and rank 44. The population coverage of these hundred female forenames is 15,468,883, or just over 64.5% of the female names. This is lower than the Census result of just under 91.5% and reflects on the broadening of the popular forename group.

⁴ D. A. Galbi, 'Long term trends in the frequencies of given names', *Names*, 50 (2002), 275-88 (p. 277).

⁵ J. Schwegel, *The Baby Name Countdown*, 5th edn (New York, 2001).

Table 5: The Top 100 Ranked Female Forenames

Forename	ER Count	Census Count	ER Rank	Cen Rank	Count Ratio ER/Cen
Margaret	727598	355004	1	11	2.04955
Mary	528980	1870086	2	1	0.28286
Susan	514472	93261	3	31	5.51648
Elizabeth	435892	1193437	4	2	0.36524
Patricia	379695	91	5	1278	4172.47253
Jean	316610	1059	6	305	298.97073
Sarah	305202	918410	7	4	0.33232
Christine	292528	1636	8	238	178.80685
Joan	283513	1108	9	295	255.87816
Helen	267774	32447	10	57	8.25266
Linda	257738	898	11	328	287.01336
Janet	256529	15129	12	83	16.95611
Julie	249753	869	13	336	287.40276
Karen	246748	65	14	1623	3796.12308
Kathleen	242857	3776	15	160	64.31594
Barbara	239095	10477	16	97	22.82094
Ann	231661	946365	17	3	0.24479
Catherine	223414	168607	18	17	1.32506
Anne	219371	131590	19	25	1.66708
Dorothy	217375	16499	20	75	13.17504
Jane	215815	672773	21	5	0.32078
Jacqueline	205212	9	22	6815	22801.3333
Carol	193427	82	23	1377	2358.86585
Jennifer	188805	26	24	3055	7261.73077
Angela	185604	444	25	504	418.02703
Sheila	180178	8	26	7590	22522.2500
Maureen	175175	18	27	4018	9731.94444
Sandra	171786	12	28	5403	14315.5000
Joyce	170325	996	29	309	171.00904
Pamela	166376	1417	30	253	117.41426
Gillian	160801	43	31	2159	3739.55814
Pauline	160021	1629	32	239	98.23266
Sharon	158968	19	33	3892	8366.73684
Nicola	158796	12	34	5531	13233.0000

Forename	ER Count	Census Count	ER Rank	Cen Rank	Count Ratio ER/Cen
Alison	156355	362	35	576	431.91989
Joanne	154491	96	36	1237	1609.28125
Lisa	152611	122	37	1052	1250.90984
Irene	151448	1017	38	306	148.91642
Eileen	150372	157	39	907	957.78344
Claire	147468	108	40	1142	1365.44444
Deborah	145159	4828	41	141	30.06607
Emma	140282	383302	42	10	0.36598
Valerie	139120	48	43	2012	2898.33333
Michelle	138156	2	44	26224	69078.0000
Caroline	130594	134393	45	24	0.97173
Wendy	130075	3	46	18084	43358.3333
June	127350	1392	47	257	91.48707
Amanda	127102	1704	48	229	74.59038
Doreen	126521	15	49	4463	8434.73333
Brenda	122959	138	50	977	891.00725
Louise	116205	14155	51	86	8.20947
Doris	115315	133	52	1000	867.03008
Elaine	114437	105	53	1161	1089.87619
Sylvia	112272	1266	54	270	88.68246
Diane	111277	44	55	2134	2529.02273
Shirley	108556	67	56	1566	1620.23881
Rachel	103422	41314	57	49	2.50332
Tracey	99048	49	58	1992	2021.38776
Audrey	96268	144	59	957	668.52778
Betty	94905	16452	60	76	5.76860
Victoria	93564	3318	61	169	28.19892
Marjorie	93534	150	62	940	623.56000
Ruth	93202	39785	63	51	2.34264
Marie	90987	9299	64	105	9.78460
Judith	90939	3065	65	175	29.67015
Maria	89946	174115	66	16	0.51659
Rosemary	88452	57	67	1789	1551.78947
Rebecca	87379	62067	68	37	1.40782
Marion	87373	12434	69	89	7.02694
Yvonne	86630	13	70	5250	6663.84615

Forename	ER Count	Census Count	ER Rank	Cen Rank	Count Ratio ER/Cen
Frances	86229	99650	71	29	0.86532
Denise	84597	21	72	3565	4028.42857
Edith	83993	148742	73	19	0.56469
Tracy	83708	26	74	3017	3219.53846
Dawn	83318	10	75	6106	8331.80000
Fiona	83219	11	76	5595	7565.36364
Florence	83090	144193	77	21	0.57624
Laura	83085	39725	78	52	2.09150
Sally	82631	2886	79	184	28.63167
Lorraine	82303	5	80	9916	16460.6000
Lesley	80988	18	81	4017	4499.33333
Elsie	80354	5331	82	137	15.07297
Winifred	79796	11190	83	94	7.13101
Heather	79768	65	84	1617	1227.20000
Gladys	79471	706	85	383	112.56516
Samantha	79200	8	86	7170	9900.00000
Alice	78089	401638	87	8	0.19443
Beryl	75612	82	88	1376	922.09756
Donna	75057	59	89	1735	1272.15254
Phyllis	75032	379	90	558	197.97361
Annie	73754	385600	91	9	0.19127
Janice	72273	51	92	1934	1417.11765
Clare	71987	2057	93	213	34.99611
Hazel	71592	17	94	4192	4211.29412
Carole	71134	35	95	2463	2032.40000
Julia	70956	56621	96	40	1.25317
Lilian	70454	16429	97	77	4.28839
Vera	69756	213	98	773	327.49296
Edna	67839	2962	99	180	22.90311
Agnes	67730	99524	100	30	0.68054

3.2.2. Forenames — Male

The top hundred forenames for males are listed in Table 6. *John* is the clear favourite, but *David*, number two, has a substantial lead over the rest of the field; half as much again as *Michael*, the third male forename. The top twenty, and forty-eight other forenames from the Census top hundred appear in the list: over twice the number for females. All names in the list appear in the Census, as with the females. There is one new immigrant forename in the list: *Mohammed* at entry 93 which has a Census rank of 26,311. *Mohammed* went from a count of two to 45,111. However, the greatest count advance was made by *Derek* at entry 35, going from a count of five to 145,634.

John, the most popular male forename, is significantly more popular than *Margaret* 1,305,357 to 727,598 (1.79) but at the Rank=100 level the count for the male forename *Shaun*, 39,145 is significantly less than the count for the female forename *Agnes* 67,730 (0.58). In the Census *Mary* with a count of 1,870,086 was the most popular forename followed by *William* at 1,783,264 so we have seen the most popular forename pass to the males, albeit with a diminution of level of popularity, from *Mary* at 1,870,086 in the Census to *John* at 1,305,357 in ER.

The population coverage of these hundred male forenames was 15,468,883, or just over 82% of the male names. This is lower than the Census result of just over 93.1% and again reflects on the broadening of the popular forename group, although less than that for female forenames.

It appears that male fore-naming, because of the fewer forenames employed, is more ordered, uniform, than the more entropic (volatile) female fore-naming.

Tables 5 and 6 also show the Census Count and Rank for the subject forenames so that the changes in count and rank can be seen. For example in Table 6 *John* which has Rank 1 in the ER data was Rank 2 in the Census data but its ER count is only 78.9% of the Census count, another indication of the flattening-out of the counts.

Table 6: The Top 100 Ranked Male Forenames

Forename	ER Count	Census Count	ER Rank	Cen Rank	Count Ratio ER/Cen
John	1305357	1654418	1	2	0.78901
David	1082574	149300	2	18	7.25100
Michael	684021	45554	3	29	15.01561
James	601207	881040	4	5	0.68238
Robert	597235	364949	5	9	1.63649
Peter	577572	53895	6	27	10.71662
Paul	574356	5078	7	109	113.10673
William	535811	1783264	8	1	0.30047
Andrew	495528	27493	9	35	18.02379
Stephen	429173	40623	10	31	10.56478
Richard	403267	257241	11	13	1.56766
Mark	381868	22991	12	41	16.60946
Christopher	365613	22543	13	43	16.21847
Alan	344991	1089	14	239	316.79614
Ian	322046	182	15	795	1769.48352
Thomas	315309	1051542	16	3	0.29985
Brian	310285	448	17	447	692.60045
Anthony	307974	11470	18	61	26.85039
George	277438	924765	19	4	0.30001
Kenneth	231783	907	20	271	255.54906
Philip	218929	22709	21	42	9.64063
Colin	207220	1274	22	220	162.65306
Steven	196475	991	23	249	198.25933
Ronald	195738	767	24	297	255.19948
Martin	193897	15907	25	53	12.18941
Kevin	193726	14	26	5065	13837.5714
Simon	188245	5180	27	108	36.34073
Graham	175378	815	28	288	215.18773
Keith	172165	143	29	938	1203.95105
Edward	168324	347770	30	10	0.48401
Raymond	163863	1640	31	196	99.91646
Gary	159712	21	32	3694	7605.33333
Charles	151415	542170	33	7	0.27928

Forename	ER Count	Census Count	ER Rank	Cen Rank	Count Ratio ER/Cen
Neil	149257	489	34	420	305.22904
Derek	145634	5	35	10764	29126.8000
Matthew	139814	26947	36	37	5.18848
Joseph	139074	453465	37	8	0.30669
Nicholas	131332	8917	38	76	14.72827
Daniel	130861	65205	39	26	2.00692
Stuart	130260	907	40	270	143.61632
Geoffrey	121254	761	41	299	159.33509
Jonathan	118466	17047	42	49	6.94938
Patrick	116827	45384	43	30	2.57419
Terence	115544	507	44	409	227.89744
Arthur	112629	249210	45	14	0.45194
Barry	112398	186	46	774	604.29032
Alexander	109743	33123	47	32	3.31320
Frederick	109223	319514	48	11	0.34184
Roy	107246	144	49	934	744.76389
Gordon	106542	1233	50	224	86.40876
Malcolm	103188	1189	51	230	86.78553
Eric	102148	495	52	415	206.35960
Roger	100581	4468	53	115	22.51141
Timothy	99150	9456	54	74	10.48541
Darren	98911	4	55	12579	24727.7500
Donald	95128	3036	56	138	31.33333
Dennis	95103	5984	57	96	15.89288
Frank	95080	108262	58	20	0.87824
Nigel	94400	127	59	1020	743.30709
Norman	88963	3601	60	129	24.70508
Trevor	86937	514	61	404	169.13813
Albert	84978	172359	62	17	0.49303
Leslie	80425	1076	63	240	74.74442
Jason	80076	471	64	429	170.01274
Douglas	79449	3006	65	139	26.43014
Adrian	77024	342	66	523	225.21637
Francis	73575	85141	67	25	0.86415
Stanley	70950	6035	68	94	11.75642
Henry	70313	619337	69	6	0.11353

Forename	ER Count	Census Count	ER Rank	Cen Rank	Count Ratio ER/Cen
Lee	69177	1038	70	243	66.64451
Craig	68931	45	71	2130	1531.80000
Leonard	65097	16000	72	52	4.06856
Bernard	62600	8308	73	80	7.53491
Ernest	59854	100955	74	22	0.59288
Gerald	59684	2031	75	170	29.38651
Harry	59283	142369	76	19	0.41640
Adam	58415	7961	77	83	7.33765
Alfred	55888	271490	78	12	0.20586
Harold	55767	14914	79	56	3.73924
Allan	55397	3547	80	131	15.61799
Reginald	53213	8746	81	77	6.08427
Clive	52902	148	82	911	357.44595
Scott	51974	762	83	298	68.20735
Gareth	50915	7	84	8489	7273.57143
Samuel	50269	247004	85	15	0.20351
Jeffrey	50103	694	86	319	72.19452
Wayne	50012	6	87	9631	8335.33333
Walter	48769	185802	88	16	0.26248
Carl	47769	1677	89	191	28.48479
Jack	46923	1438	90	210	32.63074
Sean	45347	26	91	3198	1744.11538
Hugh	45225	29662	92	33	1.52468
Mohammed	45111	2	93	26311	22555.5000
Benjamin	44829	88723	94	24	0.50527
Maurice	43374	5362	95	106	8.08915
Phillip	43190	10035	96	69	4.30394
Dean	42157	293	97	583	143.88055
Victor	40588	3987	98	120	10.18009
Russell	39896	866	99	278	46.06928
Shaun	39145	6	100	9488	6524.16667

3.2.3. Surnames

The top hundred surnames are listed in Table 7. These hundred surnames account for just under 22% of the population and 0.0129% of the surname types. *Smith* is clearly number one by a margin of 121,608 over *Jones* which has an even larger margin of 143,368 over third ranking *Williams*.

Table 7: The Top 100 Ranked Surnames

Surname	ER Count	Census Count	ER Rank	Census Rank	Count Ratio Census:ER
Smith	569914	370701	1	1	1.53740
Jones	448306	337129	2	2	1.32978
Williams	304938	213651	3	3	1.42727
Brown	274679	156769	4	5	1.75213
Taylor	264905	172737	5	4	1.53357
Davies	232247	151712	6	6	1.53084
Wilson	204388	100239	7	10	2.03901
Evans	185582	130040	8	7	1.42711
Thomas	165636	122727	9	8	1.34963
Roberts	154923	110923	10	9	1.39667
Johnson	153133	98948	11	11	1.54761
Walker	141877	82981	12	18	1.70975
Thompson	139276	85302	13	16	1.63274
Wright	139142	88078	14	13	1.57976
Robinson	139090	94223	15	12	1.47618
White	132828	86754	16	15	1.53109
Hughes	132455	80474	17	21	1.64594
Edwards	124664	81671	18	20	1.52642
Hall	122968	83831	19	17	1.46686
Green	121124	81845	20	19	1.47992
Martin	119394	61962	21	34	1.92689
Lewis	118304	77895	22	24	1.51876
Wood	117789	87418	23	14	1.34742
Harris	116659	74489	24	25	1.56612
Clarke	115551	59313	25	36	1.94816
Jackson	115468	79408	26	22	1.45411

Surname	ER Count	Cen Count	ER	Census	Count Ratio
Clark	112474	71316	27	27	1.57712
Turner	110432	79391	28	23	1.39099
Scott	109230	48250	29	49	2.26383
Hill	106738	71368	30	26	1.49560
Moore	105999	57361	31	38	1.84793
Cooper	104525	69272	32	28	1.50891
Morris	100920	63612	33	32	1.58649
Ward	100302	63262	34	33	1.58550
Watson	97742	51425	35	45	1.90067
King	97280	59862	36	35	1.62507
Morgan	94385	56882	37	39	1.65931
Harrison	94228	64940	38	31	1.45100
Baker	93258	65203	39	29	1.43027
Young	92606	44441	40	57	2.08380
Anderson	90232	25365	41	109	3.55734
Allen	89641	55862	42	40	1.60469
Patel	88110	Nil	43	Nil	NM
Mitchell	87884	42416	44	59	2.07195
James	87530	58125	45	37	1.50589
Phillips	87027	52021	46	43	1.67292
Campbell	86669	13571	47	253	6.38634
Bell	86041	43652	48	58	1.97107
Lee	82363	47803	49	51	1.72297
Kelly	82271	25057	50	110	3.28335
Parker	79682	55171	51	41	1.44427
Davis	78540	65135	52	30	1.20580
Bennett	78265	48047	53	50	1.62893
Miller	77591	37106	54	74	2.09106
Price	76242	51945	55	44	1.46774
Shaw	75633	48553	56	48	1.55774
Cook	75623	54932	57	42	1.37667
Griffiths	75596	48843	58	47	1.54773
Simpson	75250	39153	59	67	1.92195
Stewart	74342	12357	60	290	6.01619
Richardson	73789	47129	61	52	1.56568
Marshall	72199	41231	62	61	1.75109
Collins	71719	40712	63	63	1.76162

Surname	ER Count	Cen Count	ER	Census	Count Ratio
Carter	71497	49787	64	46	1.43606
Bailey	70349	44919	65	54	1.56613
Gray	69368	27504	66	97	2.52211
Murray	68890	15238	67	217	4.52093
Murphy	68539	19428	68	159	3.52785
Cox	66122	44519	69	56	1.48525
Adams	65472	39845	70	65	1.64317
Graham	64491	21338	71	136	3.02235
Richards	63747	45053	72	53	1.41493
Ellis	61661	42307	73	60	1.45747
Robertson	61072	9311	74	394	6.55912
Wilkinson	60970	44883	75	55	1.35842
Foster	60227	38722	76	68	1.55537
Chapman	58220	40742	77	62	1.42899
Russell	58019	26667	78	103	2.17569
Mason	57303	37723	79	71	1.51905
Powell	56356	37127	80	73	1.51792
Rogers	56257	38372	81	70	1.46610
Webb	56240	40383	82	64	1.39267
Owen	55284	35317	83	78	1.56537
Gibson	54332	25784	84	107	2.10720
Hunt	54130	39485	85	66	1.37090
Holmes	53728	35404	86	77	1.51757
Mills	53461	37190	87	72	1.43751
Palmer	52994	34487	88	80	1.53664
Matthews	52702	25674	89	108	2.05274
Reid	52645	7747	90	463	6.79553
Thomson	52266	5240	91	731	9.97443
Fisher	51077	32040	92	84	1.59416
Lloyd	50949	34798	93	79	1.46414
Barnes	50462	35630	94	76	1.41628
Knight	50200	35933	95	75	1.39704
Harvey	50003	29058	96	89	1.72080
Jenkins	49911	33886	97	81	1.47291
Barker	48826	38661	98	69	1.26293
Butler	48764	28541	99	93	1.70856
Dixon	47983	31966	100	87	1.50106

All the ER to Census ratios are over 1.0. All surnames increased their count from the Census counts in contrast with the forename counts which fluctuated widely. In particular *Patel* increased its count from zero to 88,110 to become the forty-third most popular surname in the UK: a phenomenal increase. There are eighty-six of the top hundred Census surnames in the group and the No.1 to No.81 inclusive are present. These eighty-six surnames seem to have been reasonably stable and growing in population with growth factors of between 1.2 and 2.1.

The new entries in the top hundred include *Patel* and thirteen others: *Thomson, Reid, Robertson, Campbell, Stewart, Murray, Anderson, Murphy, Kelly, Graham, Scott, Russell and Gibson*. These thirteen surnames are all common in either Scotland, or Ireland, or both, and their absence in the top hundred of the 1881 Census probably reflects the fact that the data used for that study was for England and Wales only.

A review of the top ranking five thousand surnames identified a further 114 *new immigrant* surnames which are listed in Table 8. (The list is not claimed to be exhaustive. The Culture-Ethnic-Language Group was taken from Hanks (2003) as supplemented by Hanks.) Since immigration in the UK has been a continuous process from at least the time the Celts held sway, a *new immigrant* surname is loosely defined as a surname from Africa, India, Asia or the Caribbean and would include Muslim, Indian, Sikh, Chinese, Arabic, Korean, Vietnamese and other Cultural-ethnic-language group (CELG) surnames. Some of the surnames listed are not strictly speaking, surnames. Examples would be *Begum* and *Bibi*. Both are terms of respect for Muslim women, honorifics, but they appear in the data very clearly as surnames. Whether this is an error of understanding or whether the family has adopted it as the family name is unknown. A casual review of the forenames associated with the surname *Bibi* shows a majority to be female forenames but there are enough male forenames such as *Mohammed* to suggest that it is being used by some as a family name. However, forename-surname entries like *Akhtar Bibi*, which is strictly speaking a surname-honorific combination, suggest that the transition from the old naming traditions to the current UK conventions is not complete, and, or, the issues are not well understood. There is an excellent discussion on immigrant naming and the need for sensitivity at www.jsboard.co.uk/etad/index.html.

Table 8: New Immigrant Surnames in Top 5,000 Surnames by Count

Surname	Culture-Ethnic-Lang	Count	Rank
Patel	Indian	88110	43
Begum	Muslim	43635	112
Khan	Muslim	43460	113
Singh	Indian	40119	129
Hussain	Muslim	35833	146
Ali	Muslim	34599	156
Kaur	Indian	33892	160
Ahmed	Muslim	29358	197
Shah	Muslim	23197	263
Akhtar	Muslim	16910	387
Bibi	Muslim	16348	404
Miah	Muslim	14161	464
Mistry	Indian	11407	587
Rahman	Muslim	9926	675
Wong	Chinese	9052	753
Iqbal	Muslim	9015	755
Chan	Chinese/Vietnamese	8804	784
Mohammed	Muslim	8555	806
Mahmood	Muslim	8494	812
Malik	Muslim Hindu	8328	830
Sharma	Indian	7447	949
Bi	Chinese	7416	992
Uddin	Muslim	6682	1055
Ahmad	Muslim	5852	1201
Hassan	Muslim	5764	1222
Parmar	Indian	5549	1271
Rashid	Muslim	5391	1307
Choudhury	Indian	5300	1328
Cheung	Chinese	5010	1406
Islam	Muslim	4843	1449
Kumar	Indian	4648	1512
Chauhan	Indian	4571	1532
Chowdhury	Indian	4232	1654
Aslam	Muslim	3865	1813
Parveen	Indian	3573	1967
Bashir	Muslim	3525	1982

Surname	Culture-Ethnic-Lang	Count	Rank
Sheikh	Muslim	3521	1986
Ullah	Muslim	3499	1996
Ho	Korean	3474	2007
Johal	Indian Sikh	3451	2024
Sidhu	Indian Sikh	3436	2031
Aziz	Muslim	3387	2055
Tang	Chinese	3378	2062
Li	Chinese	3371	2067
Lau	Chinese	3329	2090
Zaman	Muslim	3299	2106
Qureshi	Arabic	3292	2111
Lam	Chinese	3217	2160
Joshi	Indian	3161	2194
Bhatti	Indian	3129	2221
Rai	Indian	3063	2266
Mohamed	Muslim	3052	2274
Ibrahim	Muslim	3039	2287
Desai	Indian	2937	2366
Jan	Muslim	2776	2474
Chung	Chinese/Korean	2729	2511
Karim	Muslim	2695	2537
Ng	Chinese/Vietnamese	2672	2561
Akram	Muslim	2599	2612
Tsang	Chinese	2503	2700
Sahota	Indian Sikh	2485	2720
Mustafa	Muslim	2466	2739
Raja	Indian	2464	2740
Latif	Muslim	2267	2934
Hussein	Muslim	2253	2954
Rafiq	Muslim	2244	2964
Shaikh	Muslim	2241	2968
Sharif	Muslim	2214	2995
Leung	Chinese	2178	3044
Liu	Chinese	2080	3157
Tong	Chinese	2008	3244
Lal	Indian	2008	3245
Hanif	Muslim	1986	3274
Wan	Chinese	1854	3463

Surname	Culture-Ethnic-Lang	Count	Rank
Majid	Muslim	1845	3479
Tan	Chinese	1809	3539
Parekh	Indian	1764	3608
Hamid	Muslim	1738	3669
Riaz	Muslim	1735	3673
Dhaliwal	Indian	1709	3713
Bhogal	Indian	1682	3763
Syed	Muslim	1652	3830
Azam	Muslim	1645	3845
Mohammad	Muslim	1596	3958
Arshad	Muslim	1573	4001
Farooq	Muslim	1514	4112
Sarwar	Muslim	1509	4123
Pandya	Indian	1491	4171
Akhter	Muslim	1446	4276
Dsouza	Indian Goan	1426	4324
Asghar	Arabic	1418	4343
Sagar	Indian Sikh	1414	4350
Abdul	Muslim	1411	4359
Yousaf	Muslim	1403	4380
Yusuf	Muslim	1397	4391
Sultana	Muslim	1392	4406
Lai	Chinese	1392	4407
Atwal	Indian Sikh	1351	4537
Kausar	Muslim	1350	4540
Khaliq	Muslim	1340	4567
Randhawa	Indian Sikh	1305	4661
Hasan	Muslim	1273	4780
Chand	Indian	1263	4809
Javed	Muslim	1254	4837
Yasin	Muslim	1251	4847
Haq	Muslim	1246	4869
Popat	Indian	1243	4881
Sohal	Indian	1236	4899
Ayub	Muslim	1230	4918
Mir	Muslim	1228	4924
Meah	Muslim & Irish	1224	4943
Bassi	Indian Sikh	1214	4977

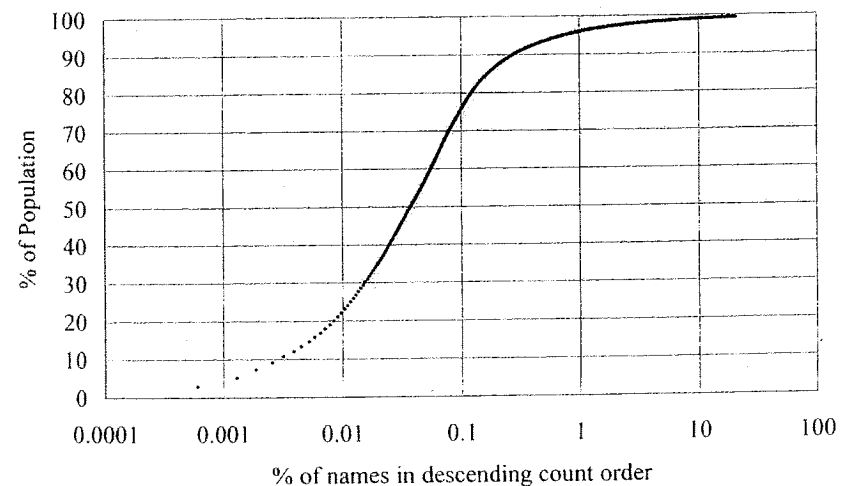
3.3. The Graphs

Two types of graph are presented: coverage and ranking. Distribution coverage graphs plot the linear cumulative percentage population (linear scale) against cumulative percentage of name types, in descending count order (logarithmic scale). This may be called 'the publisher's curve' since it shows population coverage (customers) against number of names (size of dictionary). Ranking graphs plot percentage of population (logarithmic scale) against rank in descending rank order (logarithmic scale).

3.3.1. Distribution Coverage — Female forenames

Graph 1 shows the plot for female forenames. It is a typical forename curve similar to those for USA, Canada, and 1881 UK Census. Compared with the Census curve it originates at about the same point but rises more slowly than that of the Census curve as the most popular forenames command less of a percentage of the population in the ER. The resulting curve is thus to the right of the Census equivalent for both female and male forenames. Over 95% of the population is covered by just 1% of name types; this is slightly less than that for the Census.

Graph 1 - ER Female Forename Coverage

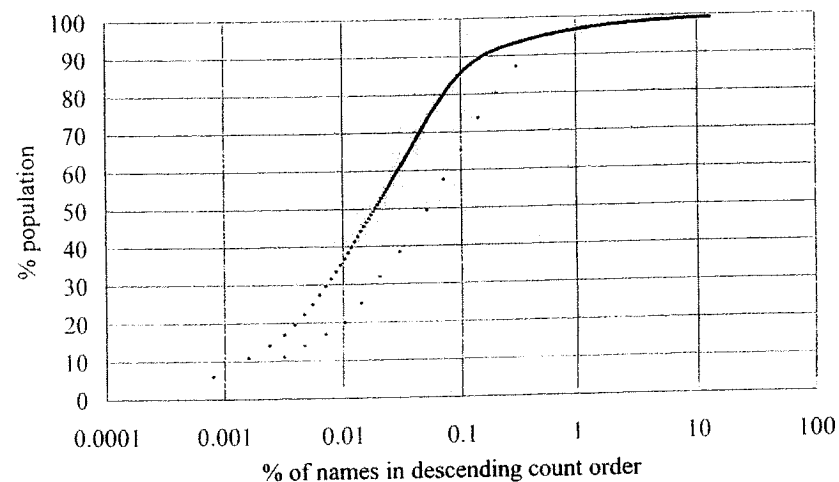


3.3.2. Distribution Coverage — Male forenames

Graph 2 shows the plot for male forenames. Ignoring for the moment the second set of plots to the right of the main curve, the main curve is similar to the plot for female forenames but rises earlier and more sharply as expected of a less entropic group, but it is still to the right of the Census curve.

The problem caused by typographical errors was discussed earlier. In order to see the impact of a large number of errors it was assumed that all male forenames with a count of one—80,855 entries of the 126,726 entries—were typographical errors. As each typographical error represents a real member of the population, the 80,855 typos were uniformly distributed across the remaining forenames according to count, and the graph re-plotted for the modified data. This, the Test Case, is shown as a series of plot points on Graph 2 to the right of the original male forename data. For simplicity the values for $0.5 < x \leq 100$ have been omitted but they tail off to 100,100 as does the original curve. Both curves represent the same population but the new curve has far fewer forename types. The envelope formed by the two curves, is the area where the true curve for the male forenames would likely be if the data were without typographical errors.

Graph 2 - ER Male Forename Coverage



The introduction of typographical errors forces the data curve to the left of the correct curve. Typographical errors do not invent people so the population is constant and all that has happened is that the type counts have been inflated. Thus when reading the Distribution Coverage Graphs the true curve is slightly to the right of the curve shown. If we look at the male forename coverage on Graph 2 for 0.1% of the names we see that the original data shows about 86% cumulative population but the test case is under 66%: a significant difference. However at 1% of male forenames the measures are 96% and 94% respectively: a very small difference.

Forename type measurement is further compromised by gender confusion; many forenames have non-trivial counts for *female*, *male*, and *unknown*. It is thus not possible to establish which forenames are used by both sexes, the unisex forenames. The problems caused by typographical errors can be bypassed by looking at the top 90% of the data at the cost of severely reduced numbers of types.

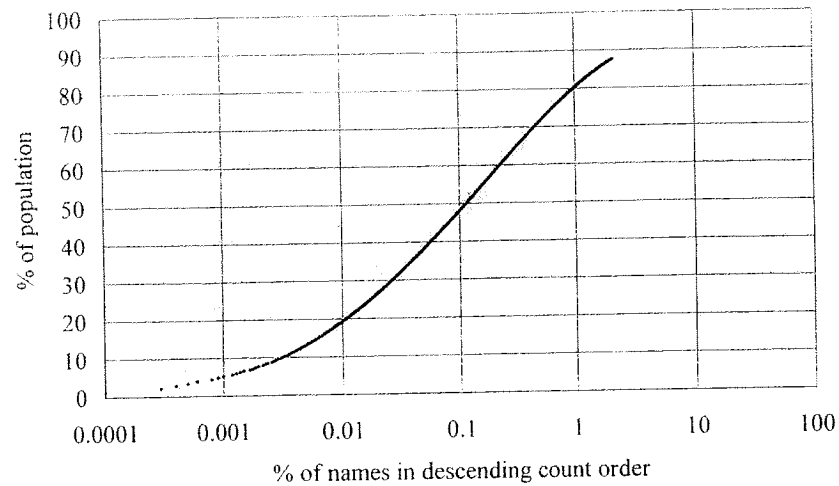
3.3.3. Distribution Coverage — Surnames

Graph 3 shows the plot for surnames. It is a typical surname curve similar to those for USA, Canada, and 1881 UK. It is slightly to the right of the Census curve as a result of the increase in the number of surnames. It is of a gentler slope than the forename curves and about 80% of the population is covered by just 1% of name types. The curve increases smoothly from the last point shown to the 100,100 point.

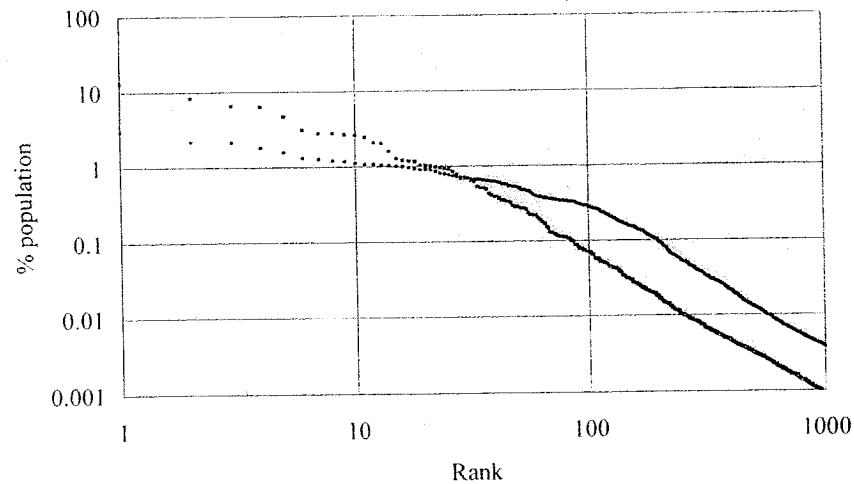
3.3.4. The Ranked Graphs — The Female and Male Forenames

The percentage of population for the top thousand female forenames is shown in Graph 4 which is a log-log graph where a power law relationship would result in a straight line. The ER plots are circles and the curve starts lower than the Census curve plotted in squares; notice that both curves are flattened in the initial range with the ER curve flattened more, and longer, than the Census curve. These two segments for the ER curve are from one to about a hundred, and from a hundred on. The first slope is gentle with a trend line in the order of $Y=3x^{(-0.4)}$, but the second is quite steep with a trend line in the order of $Y=3600x^{(-2)}$. The Census curve transition is at about rank 15.

Graph 3 - ER Surname Coverage

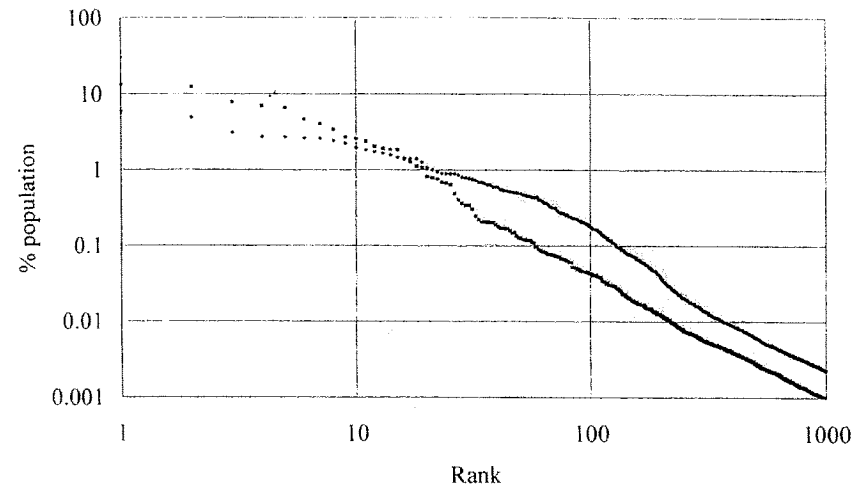


Graph 4 - ER & Cen Female Population by Rank



The percentage of population for the top thousand male forenames is shown in Graph 5. Again, the ER plots are circles and the curve starts lower than the Census curve plotted in squares. Again both curves are flattened in the initial range with the ER curve flattened more, and longer, than the Census curve. These two segments for the ER curve are from one to about sixty, and from sixty on. The first slope is moderate with a trend line in the order of $Y=8x^{(-0.7)}$, but the second is quite steep with a trend line in the order of $y=900x^{(-1.9)}$. The Census curve transition is at about rank 15 as it was for the female names.

Graph 5 - ER & Cen Male Population by Rank



The transition occurs about the same population value, between 1% and 0.1% (see Y axis on Graphs 4 and 5), in both the female and male cases. To make the transition clearer, the ER plots from Graphs 4 and 5 have been simplified and both presented on Graph 6; the solid line is the Female plot (from Graph 4) and the dotted line is the Male plot (from Graph 5).

Graph 6 - Male & Female Breakpoints

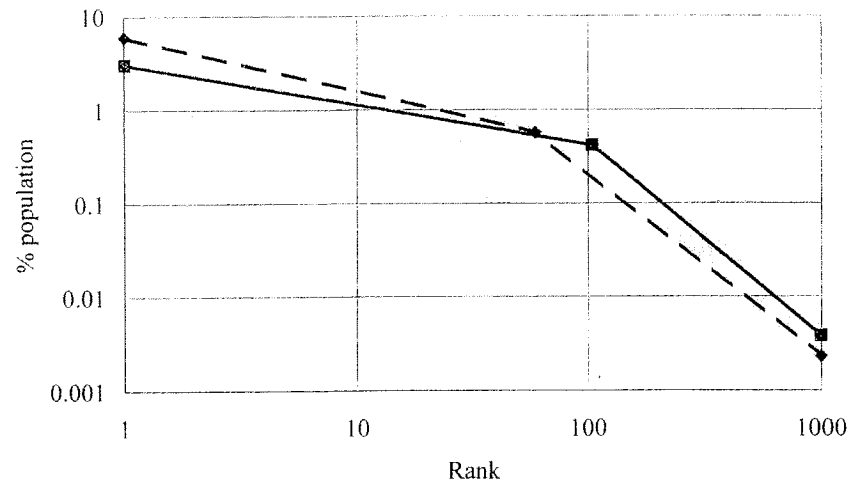


Table 9 Female Forenames, and Table 10 Male Forenames, below, are the cumulative percentage population represented by all ranks up to and including the rank shown for both the Census data and the ER data. For example the entry (10, 64.7) in Table 8 means that the cumulative percentage population at rank 10 is 64.7%. The actual break point is difficult to determine graphically; for the male forenames it appears to be about fifteen for the Census Data and about sixty for the ER Data. Although the rank ranges are different for the Census data and the ER Data the cumulative population percentages are in the same order and there is an exact match with Census Rank=13 with ER Rank=59 for percentage population of 71%.

Table 9: The ER and Census Breakpoint Values for Male Forenames

Cen Male		ER Male	
Rank	%age	Rank	%age
10	64.7	56	69.7
11	67.0	57	70.2
12	69.0	58	70.6
13	71.0	59	71.0
14	72.9	60	71.4
15	74.3	61	71.8
16	76.1	62	72.2
17	77.4	63	72.6
18	78.5	64	73.0
19	79.6	65	73.3
20	80.4	66	73.6

Table 10: The ER and Census Breakpoint Values for Female Forenames

Cen Female		ER Female	
Rank	%age	Rank	%age
10	53.0	96	63.4
11	55.5	97	63.7
12	57.5	98	64.0
13	59.6	99	64.3
14	61.2	100	64.6
15	62.4	101	64.9
16	63.6	102	65.4
17	64.8	103	65.7
18	66.0	104	66.0
19	67.0	105	66.2
20	68.0	106	66.5

Galbi compares graphically, in the manner of Graphs 4 and 5, forenaming for both males and females, for the period 1819–30 with 1994 to show the flattening-out over time of the popularity against rank curves.⁶ It is interesting to see a breakpoint in the 1819–30 curve for both males and females. No breakpoints are shown for the 1994 curves but the data only extends to rank 60 and we would not expect from the information presented in this article, to see a breakpoint in that range.

Galbi argues that, 'Naming is seen to be representative of more general patterns of behavior in the information economy'.⁷ With respect to the flattening, he argues that: 'This change can be interpreted as a reduction in the magnitude of information encoded in the name distribution and an increase in the extent of personalization in naming'.⁸ A male example of that personalization would be *Au-Kanai'l*, and a female *Jy'wuanseia*, both drawn from Schwegel.⁹

It is hypothesized that the break occurs at a given percentage of the population and that the number of names to achieve this percentage is driven by fashion. We may test this hypothesis with the female forenames.

The break point for the female forenames is about fifteen for the Census Data and about a hundred for the ER Data: see Graph 6. Comparing the percentages in the Female Forenames, Table 9, we see that there is an exact match with Census Rank=18 with ER Rank=104 for percentage population of 66%.

It is fair to say that the results for the female forenames support those for the male forenames. What is the significance, if any, of the break point? It would appear that the break point defines the current popular group of forenames: the fashionable names. By fashionable I mean popular choice. I decided on the term *Fashionable Group* before I heard Professor Stanley Lieberman's Keynote Address at the joint American Name Society and Linguistic Society of America Meeting 8–11 January 2004, or obtained his book, *A Matter of Taste*.

⁶ Galbi, 'Long term trends in the frequencies of given names'.

⁷ *Ibid.*, p. 275.

⁸ *Ibid.*, p. 282.

⁹ Schwegel, *The Baby Name Countdown*.

For females this group has about 104 members, whereas for males the size is smaller at fifty-nine members. The population represented by fashionable female forenames is 15,746,552 and that by the fashionable male forenames is 15,508,288: a ratio of 1 to 1.02 which is quite close. Whether the closeness of population counts for the 'most popular sets' means anything is, as yet, unknown. Table 11 summarizes this information.

Table 11: Summary of Forename Break Point Characteristics

Measurement	Female	Male	Ratio F/M
Census breakpoint	18	13	NA
ER breakpoint	104	59	NA
Percentage Population*	66	71	NA
Census Population*	9,490,823	9,470,965	1.002
ER Population*	15,746,552	15,508,288	1.020

*at breakpoint

There appears to be a relationship between the set of fashionable forenames, which can expand to include more names but the total percentage of the population represented by these forenames remains constant at about 66% for females and 71% for males. Eleven of the thirteen fashionable male forenames from the Census appear in the fifty-nine fashionable male forenames of the ER. Twelve of the eighteen fashionable female forenames appear in the 104 fashionable forenames of the ER.

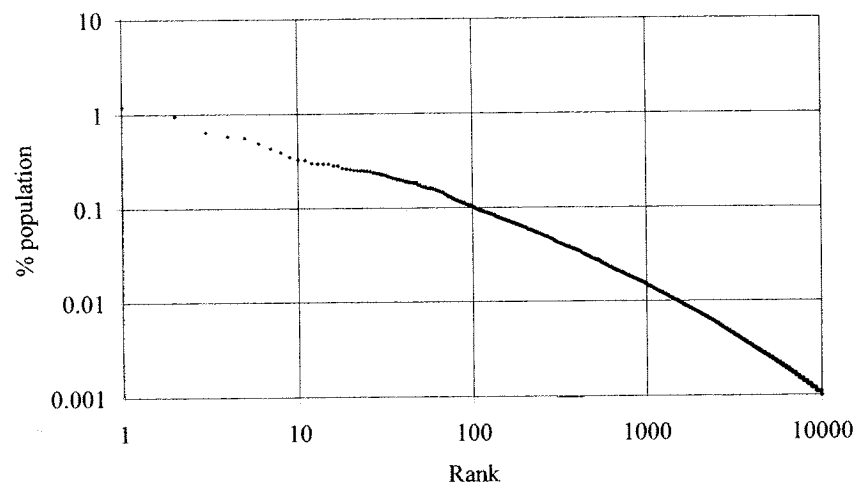
Fashionable forenames have an ordered relationship one with another; their Population-Rank plots are described by the curve $Y = Cx^{(-p)}$, but the unfashionable forenames plots are described by the curve $Y = Kx^{(-q)}$ where C and K are constants with $C < 10$ and $K > 1000$, and p and q are also constants where $p < 1$ and $q > 1$.

3.3.5. The Ranked Graph — Surnames

Graph 7 shows the ranked graph for surnames. This graph is of a different shape than that of the forenames graphs; it does not have a

distinctive 'break' as in the case of the forenames. In Graph 7 a smooth descending convex (viewed from above) results, originating at Rank 1 plot, passing through Rank 2 plot and Rank 30 and subsequent plots, whereas extending the forenames graphs to Rank 10,000 does not change the curves. Clearly the surnames curve is driven by different influences from the forename curves. It is not a simple power law.

Graph 7 - ER Surnames Population by Rank



3.4. Growth and Shrinkage in the Number of Surname Types

The 1881 Census results listed a surname count of 401,197 unique surnames, plus a pseudo name of *Unknown*, for a total population count of 26,124,584 people. The 1998 ER results listed a surname count of 781,728 for a total population count of 47,054,569. There is no reason to suspect that either of these population counts is overstated, but there is evidence that the number of types is exaggerated as discussed earlier. The *Smith* to *Smiht* corruption increases the number of false types by one and reduces the number of tokens for *Smith* by one, but the total number of people recorded stays constant; it is just that we have miscalled one person *Smiht*.

In 1881 there was no suitable technology to tackle type inflation. However, the technology has been available for years to allow the capture of surnames and forenames almost without error; it is just that the need has not been recognized generally and the technology is not being deployed.

The surname lists for the Census and the ER were compared to see whether any of 1881 surnames had become extinct in the more than a century between the events. The results were not expected. 272,327 surname types from 1881, 68% of the total, were not found in the 1998 listings, which means that only 128,870 types, 32% of the 1881 surname types, from England and Wales survived to form part of the 783,507 surname types of GB in 1998. The missing surname types, however, accounted for only 3.2% of the population, and hence had generally low counts which would make them more likely candidates for extinction.

The high apparent extinction rate raises the question of the absolute growth of the number of types of surnames in the period between the Census and the ER. The growth factor in surname types is a vigorous six times: from 128,870 to 783,507 types. Where did all these surname types come from?

There is an inviting solution both to the shrinkage of the Census types and the growth of the ER types and that is the typographical errors. This may be part of the solution but I suggest not the whole. Many of the Census surnames that do not appear in the ER are *well-formed*. By that I mean that they look like surnames from the UK; typographical errors are as seen in Table 2; they often have arrangements of consonants not known in English. Three surnames that had counts of over a hundred in the Census but which do not appear in the ER are: *Seeker*, *Shatton* and *Shirlock*. These are well-formed, English looking names. Both *Seeker* and *Shirlock* appear in the US telephone directory for 1997; *Seeker* also appears in a UK telephone directory for the same time; *Shatton* and *Shirlock* are in the Soundex Guide. Reaney and Wilson cite a *Ralph Shirloc* (1159).¹⁰ An explanation of the missing types

¹⁰ P. H. Reaney and R. M. Wilson, *A Dictionary of English Surnames*, 3rd edn (Oxford, 1997), p. 405, s.n. *Sherlock*, *Shurlock*.

is the subject of another article.

3.5. Hyphenated Surnames

The first thing that one notices is that there is a large increase in the number of hyphenated names of the form *A-B* where *A* and *B* are unhyphenated surnames. In today's society Ms *Walker* marries Mr *Jolly* and they call themselves Mr and Mrs *Jolly-Walker* (or *Walker-Jolly*). These hyphenated types are legitimate surname types but it is useful to measure the impact of this form of surname growth, since no new surname components are introduced. There are 149,695 hyphenated types in the ER data. To qualify as a hyphenated type, both parts must be present in their own right in the ER. In our example *Jolly-Walker*, both *Jolly* and *Walker* are listed in the Roll.

3.6. Growth in Surname Types

There is 120% growth in surname types over the Census even after removing the hyphenated surnames as shown in Table 12. This growth comes from new immigrant surnames and, unfortunately, our old problem of typographical errors. At this time, it is impossible to give any reasonable estimate of the real size of the growth. Only much better recording will allow that.

Table 12: The Electoral Roll Growth of Surname Types

#	Data Set-Surnames	Event	Types
1	1881 Census	Initial State	401,197
2	1881 Census	Disappeared	272,327
3	1881 Census	Survivors in 2000: #1 - #2	128,870
4	1998 Electoral Roll	Initial State	783,507
5	1998 Electoral Roll	Hyphenated	149,695
6	1998 Electoral Roll	Non-Hyphenated: #4 - #3	633,812
7	1998 Electoral Roll	Growth: #6 - #3	504,942

3.7. Fashionable Forenames

One swallow does not make a summer, but it appears possible that there is a fashionable group defined by the break point in the respective Population against Rank graphs for females and males. Current evidence suggests that the female "Fashionable Group" is larger than the Male group, 104 cf 59, having grown from 18 cf 13 for the Census, but the Female group covers less of the population, 66% cf 71% levels which are unchanged from the Census results. In absolute numbers of people, however, the females and males were approximately the same both for the Census and for the ER. The hypothesis is that for both male and females there is a fashionable group of forenames. Members of this fashionable group come and go with more joining the group than leaving with the result that the group expands. However the percentage of all forenames represented by this fashionable group remains constant with the result that the percentage share of the most popular member declines with the increase in group members.

Parents' choices of naming suggest that they collectively seek a fashionable name 71% of the time for males but only 66% for females. The balance of parental choices of naming is for the rarer and more unusual forenames: 34% of the time for females and 29% for males. In making a popular naming decision for a male a namer chooses one of fifty-nine forenames; in selecting from the other group he or she chooses one from over 100 thousand. In the female case the numbers are 104, and, again, over 100 thousand. The wonder is the stability of these percentages over time when there is no direct linkage between the namers, an interesting social phenomenon.

4. Summary

It is not possible to answer in 2004 exactly how many female forenames, male forenames, or surnames there are in the UK, or the USA, or Canada, because of type inflation. However, it is possible to state the number of types that constitute 90% of the population and this is recommended for surnames: 22,211 surname types, all with a count of over 150 representing over 42.3 million people. With forenames this level is also good but perhaps a better level is the percentage of population that constitutes the fashionable group.

The top hundred surnames are remarkably stable with the exception of the new immigrant surname of *Patel* at Rank 43. There is only one such name amongst forenames and that is *Mohammed* at Rank 93. However, this new population is clearly visible in the data. Naming and naming conventions within these communities need to be studied to ensure that names are being accurately recorded. For example *Begum* is the second most popular new immigrant surname at Rank 112, but it is a female honorific. Is it being used as such, or is it being used as a surname?

Looking at the growth in the number of surnames from the Census to the ER we find that 68% of the 401,197 Census surnames do not appear amongst the 783,507 ER surname types. An explanation for the apparent disappearance of 272,327 surname types will be the subject of a future paper. However, 149,695 of the ER surname types are hyphenated surnames, another growing social phenomenon.

Negotiating Bynames

David Postles

University of Leicester

Agnes *Daythef* took sanctuary in the church of St Mary le Bow London, confessed that she had stolen a surcoat and committed many other thefts and abjured the realm before the chamberlain and sheriffs. Nothing is known of chattels because she was a vagabond from Oxfordshire. Afterwards it is testified that a woman of this name is living in the City, so let her be arrested. Later a woman called Agnes comes, and asked by the justices what her name is and how long she has lived in the City, says she is called Agnes *de Leic'* and does not know by what surname others call her. Because the justices agree that at the time the event occurred she was not yet born, she is quit.¹

Contained within this presentment before the justices in eyre in London exists the whole conundrum of the attribution of bynames in the twelfth and thirteenth centuries, before such *cognomina* developed into hereditary, family surnames. How were these *cognomina* acquired and received by their bearers? To what extent was there a process of negotiation between the bearer and the rest of a local society?

It is known, of course, that identification by a *cognomen* might not fully represent the colloquial identification and, indeed, representation of the individual. At the end of the thirteenth century, in 1297, the abbey of St Mary, York, leased a capital messuage and carucate in Moor Monkton for a term of fifteen years for a rent of four and a half marks to Stephen *de Spaunton—dictus Judas*. Retrospectively another lease referred back to him as Stephen *dictus Judas*.²

Numerous issues are thus involved in the attribution of bynames: the first concerns the flexibility of use of bynames associated with individuals; and the second relates to alternative colloquial descriptions

¹ *The London Eyre of 1244*, edited by H. M. Chew and M. Weinbaum (London Record Society 6, 1970), 12 (no. 38).

² *Yorkshire Deeds*, edited by W. Brown, C. T. Clay and M. J. Hebditch (Yorkshire Archaeological Society, 10 vols, 39, 50, 63, 65, 69, 76, 83, 102, 111 and 120, 1909–55), II, 120 (327) and 122 (331).