

ELECTRONIC RESOURCES FOR LOCAL POPULATION STUDIES

DEMOGRAPHIC PROCESSES IN ENGLAND AND WALES, 1851–1911: DATA AND MODEL ESTIMATES

Dov Friedlander and Barbara S. Okun¹

Dov Friedlander is Professor Emeritus in the Department of Statistics, Faculty of Social Sciences, The Hebrew University of Jerusalem, Israel. His research interests include historical population processes, particularly of England and Wales, and recent social and demographic studies of the population of Israel. Barbara S. Okun is Associate Professor in the Department of Sociology and Anthropology, The Hebrew University of Jerusalem, Israel. Her research interests include historical demography, demographic methods, family and social demography, and social stratification

Introduction

Recently, we have deposited a newly available data source entitled 'Demographic Processes in England and Wales, 1851-1911: Data and Model Estimates' with AHDS History at the UK Data Archive (Study Number 5587). The relevant link, with full documentation, is <http://www.data-archive.ac.uk/findingData/snDescription.asp?sn=5587>

This source includes data for 578 registration districts of England and Wales which were compiled and digitised from two primary sources, both published by H.M.S.O: (1) printed volumes of the Censuses of England and Wales, which were administered every decade during the period from 1841 to 1911; and (2) printed reports of the Registrar General of Births, Deaths and Marriages in England and Wales, various volumes between 1851 and 1911. The data contains a wide variety of demographic and socio-economic information for the period in question, which we believe would be of value and interest to local population historians. In particular, the data is presented in a way which facilitates the study of socio-economic and demographic change over time and across districts. Demographic topics covered include fertility, nuptiality, marital fertility, mortality and migration, as well as composition of the population by age, sex and marital status. In addition, other topics included are socio-economic characteristics of registration districts, such as occupational distributions and indicators of educational levels.

From census materials we obtained, for each registration district within each county (excluding London), demographic and socio-economic data, including

such items as population figures, marital status of the population, and occupational distributions. The reports of the Registrar General provided information such as births, marriages, and deaths for each registration district for the inter-censal periods. The combined information from the census volumes and the reports of the Registrar General allowed for the computation of various demographic rates and indices, such as measures of marital fertility rates, nuptiality, overall fertility, life expectancy, and many more. Many of the measures are available for all census or inter-censal periods; others are available for selected periods only. In some cases, demographic and statistical models were employed to estimate key measures and indicators of demographic change. For example, net migration rates were derived from differences between inter-censal population growth rates and inter-censal rates of natural increase. The great majority of the data and estimates were compiled and computed at the registration district level for the period 1851-1911.

The district-level data is available in a variety of formats, as described below. Among them is a form of district profiles, which we believe will be especially convenient for use by local historians. In addition to the district-level data, decennial inter-county migration flows were estimated for each inter-censal decennial period from 1851 through to 1911. The model used to estimate inter-county migration streams is based on cross tabulations of county-level distributions of populations born in, and enumerated in, each of the censuses.

Contents of the site

England and Wales District Variable File, 1851-1911

The district file (an SPSS data file entitled `england_wales_disvar_1851_1911.sav`) contains data on 578 registration districts in England and Wales from 1841 to 1911. The file thus includes information on populations in all registration districts of England and Wales, with the exception of populations in London districts. All efforts were made to maintain consistency over time in district definitions

For each district, the file contains values for 233 variables; variable labels are included in the file and references to published materials which provide further information on variable definitions are also given where necessary. In addition, a file entitled 'Guide' provides more detailed explanations of some variable definitions and derivations.

Figure 1 is illustrative of an excerpt of the District Variable File, and highlights certain variable values for 11 districts in Surrey. Each line in the file refers to a different district. District names have been shortened for convenience. For full names of districts, as well as changes over time in district definitions and boundaries, the user can see an Excel data file called 'district_names'.

Many of the variables provide a time series of values for the same measure. For example, in Figure 1 we present six different values for implied male and female net migration rates per 1,000 per year at ages 15-44. The six different values refer to rates for the six consecutive decennial periods 1851-1860

Figure 1 Excerpt from district variable file

	district	m@fmig61	m@fmig71	m@fmig81	m@fmig91	m@fmig01	m@fmig11	urbdist	more_vars
1	SURREPSOM	-2.70	10.40	9.20	2.00	8.90	26.80	93	.
2	SURRCHERTSEY	-1.00	5.50	-3.20	1.30	10.20	9.80	91	.
3	SURRGUILDFOR	-2.60	7.60	7.20	6.60	11.30	8.90	85	.
4	SURRFARNHAM	98.00	-19.90	-11.20	17.70	16.90	17.50	82	.
5	SURRHAMBLEDO	-24.80	-13.30	-16.60	-5.40	-0.10	3.80	79	.
6	SURRDORKING	-11.30	-3.90	-13.40	-5.00	-9.50	-4.30	86	.
7	SURRREIGATE	23.40	13.80	-8.60	3.20	3.80	4.10	98	.
8	SURRGODSTON	-17.10	17.10	20.20	-1.90	37.00	14.70	87	.
9	SURRCROYDEN	25.20	49.10	22.40	9.10	14.90	22.60	93	.
10	SURRKINGSTON	20.70	36.50	22.10	19.40	22.50	14.60	97	.
11	SURRICHMOND	9.20	28.90	18.90	13.70	13.80	21.10	98	.
12
13
14

through 1901–1911 (m@fmig61 through m@fmig11). For each district, urban distance (urbdist) is also provided: this is an index which captures the weighted distance from large towns/cities (closest=100).

Other variables which may be of particular interest (not presented in the illustration in Figure 1) include a time series of measures of population density; various measures of fertility levels over time, including Coale’s indices;² characteristics of fertility transition processes, including indirect indicators of the use of fertility control; a time series of estimates of life expectancy and child mortality rates at ages 0-4; a time series of measures of literacy for men and women separately; and a time series of occupational distributions, including detail on proportions of the adult male populations employed in professional, commercial, industrial, textile, mining, domestic service and agriculture.

England and Wales district profiles

Fifty-two Excel spreadsheets present a selection of the district-level variables included in the District Variable File (see above). The districts are organised into 52 counties, with each spreadsheet containing the district-level information for each district in one county. File names refer to the relevant county. For example, the file containing district-level profiles in Surrey is called Surrey.xls. The format of these district profiles is particularly convenient for those interested in local history.

Figure 2 presents an example of one such district profile, for Epsom, Surrey (one of the 578 registration districts included in the District Variable File). Each district profile is divided into an upper and lower section, which are separated from each other by a thick black line. The upper section provides period-specific information, for a series of time frames, indicated by column headings. The lower section also provides period-specific information, but the periods do

Figure 2 Excerpt from district profiles

DISTRICT indices within SURREY indices for EPSOM (2 1) district

Period	1841-50	1851-60	1861-70	1871-80	1881-90	1891-00	1901-10
Mean population	18144	20724	26350	35699	45615	56072	74650
Net migration rate		-2.70	10.40	9.20	2.00	8.90	26.80
No. females per males, ages 15-44		1.09	1.15	1.19	1.28	1.25	1.30
% Women ever married, age 20-24	22.40	24.80	24.40	21.00	16.10	14.50	12.90
Coale nuptiality index IM	0.44	0.44	0.45	0.43	0.40	0.37	0.38
Marital fertility index IG	0.67	0.75	0.75	0.75	0.70	0.62	0.56
Total fertility rate	3.80	4.20	4.50	4.30	3.70	3.10	2.70
% Grooms unable sign name at marriage		29.20	17.80	11.40			
% Brides unable sign name at marriage		13.90	11.90	5.70			
Life expectancy at birth*	45.30	48.10	47.40	47.40	45.10	48.30	46.90
Fertility transition characteristics**		(a)0.77	(b)1880	(c)0.82	(d)1897	(e)0.58	(f)-1.41
Death rate age 0-4***		(a)45.84	(b)39.03	(c)42.44	(c)39.24	(c)36.78	(c)34.9
% Male L.F in selec. occup. 1861****		(a)6.9	(b)6.7	(c)21.7	(d)4.7	(e)5.2	(f)33.3
% Male L.F in selec. occup. 1871****		(a)7.1	(b)9.1	(c)25.6	(d)4	(e)5.1	(f)27.7
% Male L.F in selec. occup. 1881****		(a)6.43	(b)11.17	(c)27.15	(d)3.46	(e)5.65	(f)22.28
% Female L.F in selec. occup. 1871****		(a)4.7	(b)0.4	(c)1.5	(d)8.1	(e)21.1	(f)1.3

***** These percentages do not relate to the time periods indicated above. (a) Is the percentage of female labor force in professional occupations. (b) Is the female percentage in commerce. (c) The percentage of females in industry. (d) The percentage of females in textile. (e) The percentage of females in domestic occupations. (f) The percentage of females in agriculture. (g) The percentage of females engaged mainly in household duties.

* Life expectancies are based on indirect standardization, using the English age-specific mortality rates as standards.

** These indices do not relate to the time periods indicated above. (a) Is the mean value of pre-transition marital fertility (IG levels over the period 1841 to 11, the year of onset of the transition. (b) Is the estimated year of onset of the transition t1. (c) Is the estimated IG value at year t1. (d) Is the estimated year t2 of the end of the transition. (e) Is the estimated IG value at year t2. (f) Is the percentage mean rate of fertility decline between t1 and t2.

*** The rates marked (a) and (b) are specific death rates at age 0-4 in 1861-1870 for males and females respectively. (c) These rates are specific death rates at age 0-4 for both sexes for 1861-70, 1871-80, 1881-90 and 1891-1900 respectively.

**** These percentages do not relate to the time periods indicated above. (a) Is the percentage of labor force in professional occupations. (b) Is the percentage in commerce. (c) The percentage in industry. (d) The percentage in textile. (e) The percentage in mining. (f) The percentage in agriculture. (g) The percentage in medical occupations. (h) Socioeconomic type of district: 1=agriculture, 2=agriculture/textile, 3=agriculture/industrial, 4=industrial, 5=mining, 6=none of the above.

***** These percentages do not relate to the time periods indicated above. (a) Is the percentage of female labor force in professional occupations. (b) Is the female percentage in commerce. (c) The percentage of females in industry. (d) The percentage of females in textiles. (e) The percentage of females in domestic occupations. (f) The percentage of females in agriculture. (g) The percentage of females engaged mainly in household duties.

not conform to the column headings (see asterisked information following the district profile for further details). Users can also refer to the 'Guide' file for further information.

England and Wales: age, sex, and marital status distributions, 1851–1911

We present a set of district-level information on population by age and sex and, where available, by current marital status, as reported in census publications for the years 1851, 1861, 1871, 1881, 1891, 1901 and 1911. This information is contained in the SPSS file `england_wales_agesex_mar_dis_1851_1911.sav`. Current marital status information is available for certain age groups of women only, with the exception of 1901, where estimates are not computerised.

For numbers of women currently married in 1861, 1871, 1881 and 1891, information on age and marital status distributions was taken directly from reported census data. Detailed age-specific information was not available in 1851, 1901 and 1911. Numbers of married women by 5- or 10- year age group were estimated for the years 1851 and 1911, using an iterative process, from county-level information on married women by age together with district-level information on total numbers married at ages 20 and above, and 15–44 respectively, as a first approximation.³ Further file information, including details about age groups, is available in the pdf document entitled 'Guide'.

Figure 3 presents an excerpt from the Age, Sex, Marital Status Distribution file. Each line represents one of the 578 registration districts. Again, for complete district names and boundary changes, see the file 'district_names'. The figure presents numbers of females in selected age groups, as well as numbers of currently married females in selected age groups, for the districts of Surrey in 1851. For example, `fm.00.51` reports the number of females aged 0–4, and `fm.95.51` reports the number of females aged 95 and above. The variable `mr.15.51` reports the number of currently married women aged 15–19, and the variable `mr.55.51` reports the number of currently married women aged 55–64. Numbers of males, by age group, are also available, but marital status distributions of males are not included in the file.

England and Wales inter-county migration files

A set of six SPSS data files contain information on net and gross inter-county migration streams for each of the six decennial periods from 1851–1861 through 1901–1911. For example, the file which refers to the period 1851–1861 is entitled `england_wales_intercount_migration_1851_1861.sav`. Each file contains information on gross and net migration flows between pairs of counties. Because there are 53 total counties (including London), migration flows are recorded for each of the 2,756 (53 x 52) pairs.⁴

The model used to estimate inter-county migration streams is based on cross tabulations of county-level distributions of populations born in and enumerated in each of the various censuses. Regarding information on gross and net migration streams between counties, the user should be aware that streams between

Figure 3 Excerpt from age sex marital status distributions

	district	fm.00.51	fm.05.51	more_ages	fm.95.51	mr.15.51	mr.20.51	other_ages	mr.55.51	more_var
1	SURREPSOM	1175	1013	.	1	0	199	.	560	.
2	SURRCHELTSEY	1075	969	.	2	0	179	.	505	.
3	SURRGUILDFORD	1602	1473	.	1	0	296	.	776	.
4	SURRFARNHAM	1371	1220	.	1	0	403	.	412	.
5	SURRHAMBLEDO	885	872	.	1	0	143	.	500	.
6	SURRDORKING	732	703	.	0	0	107	.	358	.
7	SURRREIGATE	985	825	.	0	0	173	.	383	.
8	SURRGODSTONE	606	536	.	1	0	99	.	277	.
9	SURRCROYDEN	1935	2042	.	1	0	323	.	758	.
10	SURRKINGSTON	1642	1425	.	0	0	310	.	737	.
11	SURRICHMOND	879	787	.	3	0	167	.	387	.
12
13
14

Figure 4 Excerpt from intercounty migration files

	county a	county b	adjacent	distance	n_a_to_b	g_a_to_b	g_b_to_a	reside a	reside b
1	LONDON	SURREY	1	55	-6069	17834	11765	2175369	198072
2	LONDON	KENT	1	97	13628	10224	23852	2175369	470526
3	LONDON	SUSSEX	0	85	3448	7973	11421	2175369	334108
4	LONDON	HAMPSHIRE	0	157	2272	9444	11716	2175369	388257
5	LONDON	BERKSHIRE	0	134	6078	2322	8400	2175369	197275
6	LONDON	MIDDLESEX	1	25	-7428	14563	7135	2175369	146264
7	LONDON	HERTFORDSHIRE	0	60	7069	3348	10417	2175369	172296
8	LONDON	BUCKINGHAMSHIRE	0	102	4704	1839	6543	2175369	142585
9	LONDON	OXFORDSHIRE	0	143	4513	1041	5554	2175369	168981
10	LONDON	NORTHAMPTONSHIRE	0	165	3668	1287	4955	2175369	211004
11	LONDON	HUNTINGDONSHIRE	0	172	1911	340	2251	2175369	59841
12	LONDON	BEDFORDSHIRE	0	110	3020	759	3779	2175369	128897
13									
14									
15									

adjacent counties may include short- as well as longer-distance migrations, whereas streams between non-adjacent counties are usually longer distance.

In addition to the estimates of county-level migration streams, we provide in the relevant computer files county-specific population totals of individuals born in England or Wales and enumerated in England or Wales at census dates. That is, we provide for each county the total numbers of persons in residence in each county in each census, regardless of county of birth. The term 'in residence' is used, although this, in fact, should be understood as 'place where enumerated'. In the article by Friedlander and Roshier (see fn 4), the authors utilised this type of county-level population total in their estimates of various indices of migration, which relate migration streams to population bases, for purposes of standardisation. Note that there are small numbers of persons enumerated in each census for whom exact county of birth is unknown, or who were born abroad. Such individuals are omitted from population counts.

Figure 4 presents an excerpt of inter-county migration streams during the period 1851–1861. For example, the first line refers to migration streams from London to Surrey. The variable 'adjacent' indicates that these two counties are physically contiguous. The variable 'distance' indicates the air distance in miles between the centres of gravity of the two counties. Regarding information on distances between counties, these distances are not very meaningful in the context of adjacent counties.

The variable 'n_a_to_b' represents the estimated net migration flow from a (London) to b (Surrey), 1851–1861. Negative values indicate net outflow from a to b. Positive values indicate net flow into a from b. The variable 'g_a_to_b' represents the estimated gross migration flow from a to b, 1851–1861. The variable 'g_b_to_a' represents estimated gross migration flows from b to a, 1851–1861. The variable 'reside a' represents population in residence in county a, according to the 1851 census, regardless of county of birth. The analogous information for county b is represented in 'reside b'.

Summary

Those interested in analyses of population processes in the districts in England and Wales should direct their attention to the District Variable File. A number of published articles have relied on this data (for a complete list, see the pdf file entitled 'Guide'). Whoever is interested in specific districts, counties or regions may find the district profiles more convenient to use. In addition, researchers whose focus is on historical inter-county migration streams will also find this information available. Overall, this electronic resource makes accessible a large and diverse set of historical census and vital registration data, as well as many quantitative variables, estimated from demographic and statistical models, which shed light on important demographic transformations and which are not otherwise available in any published source.

NOTES

1. Contact e-mail addresses for Dov Friedlander and Barbara Okun are, respectively, dovfri@vms.huji.ac.il and bsokun@vms.huji.ac.il. The postal address for both is Faculty of Social Sciences, Hebrew University of Jerusalem, Mt. Scopus Campus, 91905 Jerusalem, Israel.
2. The clearest explanation of how to calculate these rates is in R. Woods, *Population analysis in geography* (Harlow, 1979), 118–20. For a good demonstration of their use see A. Armstrong, *The population of Victorian and Edwardian Norfolk* (Norwich, 2000), 105–15.
3. Further details can be found in D. Friedlander, 'A technique for estimating a contingency table, given the marginal totals and some supplementary data', *Journal of the Royal Statistical Society*, **3** (1961), 412–20
4. A detailed description of the method used to estimate intercounty migration streams is provided in D. Friedlander and R.J. Roshier, 'A study of internal migration in England and Wales: part I', *Population Studies*, **19** (1966), 239–79.