GLOBAL PROJECTIONS OF HOUSEHOLD NUMBERS AND SIZE DISTRIBUTIONS USING AGE RATIOS AND THE POISSON DISTRIBUTION.

Vic Jennings[†], *University of Melbourne*
Bill Lloyd-Smith, *RMIT*
Duncan Ironmonger, *University of Melbourne*

[†] Address for correspondence: Households Research Unit, Economics Department, University of Melbourne, Victoria 3010, Australia. e-mail: f.soupourmas@ecomfac.unimelb.edu.au

Abstract

Global projections are made for household numbers and size distributions in 2010, 2030 and 2050 . These projections are assembled from 142 individual country projections and a 'rest of world' projection. These projections use the latest UN World population projections with three fertility scenarios (UN 1999) and the "Age ratio-Poisson" model originally developed in Jennings, Lloyd-Smith and Ironmonger (1999), as a "Dependency ratio-Poisson" model. The projections show the total world growth in household numbers is almost independent of the fertility scenario. If these projections are sustained, in the next 50 years, the world number of households will almost treble from 1.3 billion in 1990 to 3.6 billion in 2050. This will occur whether the world population doubles from 5.1 billion to 10.4 billion on the high fertility scenario or only increases by 50 per cent to 7.1 billion on the low fertility scenario. The apparent similarity of total household growth under the various scenarios conceals a phenomenal switch in the size distributions. Urban regions are expected to absorb most of this growth. The paper continues the research on the demography of households by the University of Melbourne's Households Research Unit.

GLOBAL PROJECTIONS OF HOUSEHOLD NUMBERS AND SIZE DISTRIBUTIONS USING AGE RATIOS AND THE POISSON DISTRIBUTION.

Vic Jennings†, *University of Melbourne*
Bill Lloyd-Smith, *RMIT*
Duncan Ironmonger, *University of Melbourne*

This paper is a sequel to Jennings, Lloyd-Smith and Ironmonger (1999), hereinafter referred to as JLI (1999). Projections are given of world household numbers and their size distribution through to 2050. These projections are based upon United Nations (UN) population estimates and projections (UN 1999a,b) and household census data from 1950 to 1995, see for example UN (1997). Further study of the concept of the household is also undertaken.

The procedure we have adopted has its antecedents in the work of John Graunt. In the 1660's he used births, deaths, and various population ratios to estimate families and hence population in London during the periods of the Plague (David 1962). In this paper with the refinements of modern censuses we use age ratios, together with population to estimate households, and with probability to estimate their size distribution. Projections are then made based upon the relationships established.

A recent long term projection of households at the world level was given in the Global Report on Human Settlements 1996 by the United Nations Centre for Human Settlements (Habitat 1996). The projections were based upon data provided by national Statistical Offices. There was no sensitivity analysis provided, nor any assessment of error, nor household size distribution information. Nevertheless the report was a major advance in that it showed, to the authors' knowledge for the first time, a comprehensive long term projection of households for the world assembled from individual country projections.

Many projections of households are made for individual countries or regions within countries. For example, the household projections produced for the State of Victoria in Australia through to the year 2021, see the Department of Infrastructure Victoria (2000), hereinafter referred to as DoI(2000). This State projection is based upon reconciling three methods; the cohort component method, a partly developed household formation method and the housing unit method. Another example is the propensity method (Ironmonger and Lloyd-Smith 1992) which provides detailed household counts by type and has been applied nationally and at State level in Australia. Van Imhoff (1995:273-291) describes LIPRO: a multistate household projection model which is applied to produce household forecasts by type through to 2050 for the Netherlands. In all these cases however cohort and other forms of data are required and this type of information is not readily available across the world. Even in developed countries research is still necessary to improve the methods, see Bartie (2000) and van Imhoff (1995: 348-350).

Population data by age and sex are available across the world for each country and presented in a standard format (UN 1999b). The data extend back in time to 1950 and forward as projections to 2050. In addition, near standardised information on household numbers and their size distributions are available for many countries at various census periods (UN 1955-1997). By bringing the two sets of data together it is possible to

† Address for correspondence: Households Research Unit, Economics Department, University of Melbourne, Victoria 3010, Australia. e-mail: f.soupourmas@ecomfac.unimelb.edu.au

determine the strength of the relationship between households and populations. If a strong relationship could be found between these two types of data then this relationship could be used as a predictive tool for household numbers and size distributions.

Such a relationship was found. In JLI (1999) it was shown that household size distribution was related to the Poisson distribution and that the Poisson parameter used, namely average household size, was in turn related to age ratios. Thus projections of households could be made based on changes in population characteristics.

Population projections to 2050 may be made with some degree of confidence. Changes in the population of individual countries are generally gradual and smooth and many of the existing population are expected to live to the year 2050. In addition considerable efforts are made in most countries to estimate future fertility, mortality and migration upon which population forecasts are based. Assumptions are subject to continuing research. This is in contrast with many technical and economic areas where there is a higher degree of uncertainty surrounding long term projections. Rescher (1998: 193-208) discusses the degree to which various phenomena are predictable.

For this paper household intensities (households per person) were regressed on age ratios to obtain household intensity equations including measures of error. These equations, based on the 1950 to 1995 data, were then used to project household intensities and the numbers of households to 2050 under three alternative UN population scenarios, the low, medium, and high fertility variants (UN 1999b). The external inputs to the model are the UN based population projections.

The adjustment terms for the truncated Poisson distribution were obtained as follows. The UN Demographic Yearbooks, for example UN (1955-1997), together with official government sources provide household size distribution data for many countries from 1950 to the present. For each household size *k* the discrepancies between fitted truncated Poisson proportions and observed proportions were regressed on household intensities. The resulting linear equation provided an adjustment term which when added to the truncated Poisson value for a particular intensity gave the estimate of observed proportion of households of size *k*. An estimate of error was also obtained.

The combined model can be called the “Age ratio – Poisson” model since it uses age ratios to determine household intensities, which in turn generate adjusted truncated Poisson distributions for household size.

At least six significant figures are used in calculations to enable a reconciliation of totals.

Statistics of Household Numbers and Long-Term Projections

The most numerous and ultimately influential decision making units in the economy are not governments, firms, or multinational corporations, but households. A recent review of the demand for demographic data by the Australian Bureau of Statistics (ABS) found that for many purposes, the demand for up-to-date statistics on the numbers of households was stronger than the demand for statistics on people. The ABS has been a leader in providing current estimates of the numbers of households and in making household projections (ABS, 1996 and 1999).

Although much of economic modelling and forecasting is concerned with outcomes over a relatively short-term horizon of the next one to five years, for many environmental issues and for the planning of infrastructure such as roads, ports (sea and air), railways, water supply and the urban built environment, projections of likely developments over the next 25 to 50 years are needed. Indeed for some purposes, such as the actuarial assessment of the viability of social security arrangements, projections are made for a 75-year horizon (Lee, 2000). And although in this paper we use the UN World population projections for the next 50 years, supplementary UN projections provide a 150-year horizon to 2150 (UN 1998a).

Data on Population and Age structure used for Projections

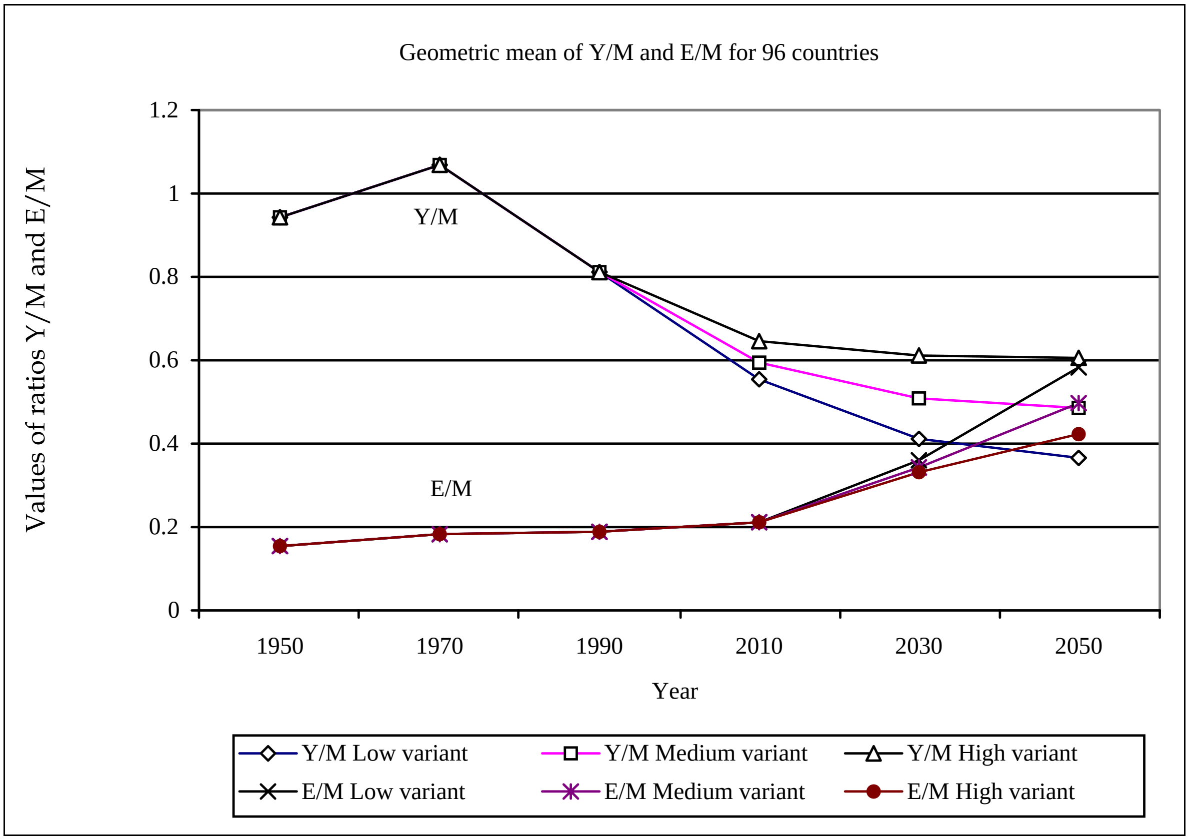
The projection years are 2010, 2030 and 2050. The reference date upon which projections were based was 1990 since it was the latest time at which a wide range of both household and population data could be obtained. The population data was obtained from UN World Population Prospects, the 1998 Revision (UN 1999b). The projected population through to the year 2050 (low, medium and high fertility variants) was used as the basis for projecting the number of households. The variants were regarded as providing feasible scenarios and enabled a measure to be obtained of sensitivities of household projections to changes in fertility (UNFPA 1999:31).

The population in households is generally less than the total resident population. For each separate country at a particular time the total population, the population in households, and household numbers were collated. It was found that in a number of countries the population in households was recorded as being the same as the total population. By comparing population tables in UN Population projections (UN 1999b) and those in the UN Demographic Year Books, for example UN (1997) it was ascertained whether the information given in the latter was likely to be the household population or total population. If the information given was total population then a reduction factor was applied. This was calculated as follows.

There were 68 countries (2.5 billion persons) for which the population in households and the total population were reported to be different. The average (unweighted) proportion of persons in households to total population was calculated to be 0.972824404 for these countries. This is the reduction factor used in this paper.

Household projections are made in the context of the radical changes in the relative proportions of young and elder persons projected over the next 50 years. In Figure 1 the variation over time of the geometrical mean of two age ratios is shown. These ratios are used later in the estimate of household intensities. For 96 countries representing 83% of the world's population at 1990 the geometrical mean of the youth/middle person ratio (Y/M) has been declining since about 1970 and the elders/middle person ratio (E/M) is projected to climb from about 2010.

Figure 1 Estimates from 1950 to 1990 and projections through to 2050 of the geometric mean of the ratio Y/M, youth (0-19) to persons (20-59); and E/M, elders (60+) to persons (20-59); for low, medium and high fertility variants for 96 countries.



Youth (*Y*) is defined to be a person aged 0-19 years, a middle person (*M*) is aged 20-59 years, and an elder person (*E*) is aged 60+ years. The elder age cut off point of 60+ differs from 50+ used in JLI (1999). *M* is of the order of 2.6 *W* where *W* is the number of women aged 20-49 as used in JLI(1999) and *E*(60+) is approx 0.6 *E*(50+). It is shown in JLI (1999) that *E*(50+) /*W* and *Y*/*W* are related to the median age and the total fertility rate respectively. Thus the ratios used in this paper correspond approximately to other measures of fertility and ageing.

In 1990 youth were four times as numerous as elders but in 2050 this ratio is projected to be between 1.5 for the high variant case to 0.7 for the low variant case. Similar ratios to those in Figure 1 are obtained for the estimated total world population. In the past the opposing changes in these ratios has been accompanied by a general movement to increased household intensity. See Figure 3 where observed household intensity is plotted against (E/M)/(Y/M) = E/Y.

The population projections in UN (1999b) show the youth ratio (*Y/M*) declining for most countries from a range of values from 0.4 to 1.6 in 1990 to a narrow range of 0.4 to 0.6 in 2050 for the medium fertility variant, and for the low fertility variant to between 0.3 to 0.5. For the elderly age ratio (E/M) the medium fertility projections rise from an extreme value of just over 0.4 in 1990 to around 1.0 in 2050. For some countries, by 2050 the population aged 60+ is equal to the population aged 20-59 years.

The Household Intensity Model

Average household size is the parameter used in the truncated Poisson distribution for determining the allocation of persons to households. However the authors have found that its inverse, described as household intensity or households per person or households per 1000 persons has a number of advantages in developing projections at country-wide level. The equations relating households to age ratios are more simply expressed using household intensity rather than household size. In addition the plot of the truncated Poisson distribution against intensity gives a better spread when household intensity exceeds 250 households per 1000 persons (average household size 4) as seen in Figure 6. This is particularly desirable as the general trend is to higher household intensities, with currently a number of countries exceeding 400 households per 1000 persons. Thirdly, household intensity relates to other per capita measures commonly used such as energy per capita or Gross Domestic Product per capita.

Data used to determine constants of the household intensity model

The estimated population data for the years 1950 to 1990 provided the basis for determining the constants of the model, and for measuring error.

There are 140 countries (4.9 billion persons) where the authors have at least one set of data which have both household counts and population data by age and sex. Of these 140 countries there are 96 countries where there is at least two sets of data, 82 countries where household counts and population data can be estimated for both 1980 and 1990 and 14 countries where household counts and population data can be estimated for both 1970 and 1980.

Table 1 Selection of countries according to the household and population data available

Number of countries used in projections	At least one household size distribution in the period 1970 to 1995	Hshld. counts at 1970	Hshld. counts at 1980	Hshld. counts at 1990	Hshld. counts One point from 1970 to 1995 per country	Population from 1950
140					x	x
2						x
82			x	x		x
14		x	x			x
112	x					x

Since household counts were generally not available exactly at 1970, 1980 or 1990, in order to obtain estimates at these times linear interpolations were made from the closest available data. There were two countries, Nigeria, and Ethiopia, with large populations for which household counts were not available. Because of their size it was considered desirable to estimate households for these countries using equations (1) and (2) and add them to the list of 140 countries. There are 112 countries where household size distribution data are also available. This method of selection is detailed in Table 1. The list of countries used is shown in Table A1 in the Appendix.

The estimated model

Following JLI (1999) it is now possible to obtain two regression models with household intensity as dependent variable and various age ratios as independent variables. The first is based upon the estimate for population and households for the 140 countries described above and the second is based upon the estimate for the 96 countries described with two points. The two formulae provide a means of cross validation

Each of the two approaches has advantages. The set of 140 countries with one time point for each country provides a wider diversity of countries and therefore any estimate based upon this will more closely represent the world population. On the other hand the data for the 96 countries (at two time points) provide more information per country. By choosing two time points the variation within a given country over time can be taken into account.

To estimate the constants of the model we used household intensity, *h*, as the number of households per person and defined the population age groups *Y*, *M* and *E* of the total population, *P*, as:

Y = population aged 0-19 years
M = population aged 20-59 years, and
E = population aged 60+ years.

The regression equations obtained were for the 1 × 140 country case equation (1) and for the 2 × 96 country case equation (2)

$$h = 0.082 + 0.064 * (M / Y) + 0.468 * (E / M) \qquad R^2 = 0.87 \qquad (1)$$

$$h = 0.169 + 0.109 * (M / Y) + 0.363 * (E / M) - 0.251 * (M / P) \qquad R^2 = 0.89 \qquad (2)$$

Both equations (1) and (2) are similar in form to equation (26) in JLI (1999). However in equation (2) the M / P term is added. This additional factor is significant at the one per cent level using relevant *t* tests and F tests and is the proportion of adults aged 20-59 years (*M*) to the total population (*P*). The estimates of intensities obtained for equations (1) and (2) are averaged for each country for which estimates are prepared. This is equivalent to averaging the two estimates of the number of households for each country.

The observed intensities adjusted to the decade years for the period 1950 to 1990 were compared with the average intensities obtained from equations (1) and (2) and plotted in Figure 2. The observed intensities ranged in value from 144 households per thousand for Algeria to 468 in the case of Sweden.

For each of the five data sets the difference between the observed and estimated household intensities were found to be approximately normally distributed. The average difference over the data for a decade year was not significantly different from zero using a *t* test. The variances of the differences for each data set were not significantly different from each other on an F test using the variance of the 2 × 96 country data set as denominator. Thus the variability in the differences does not change significantly over time. The standard error is shown on the graphs in Figure 2. The combined data set was later used in evaluating equation (3),

For the purposes of predicting statistical error in the household intensity projections using the average of equations (1) and (2), the same 2×96 data set was used as for determining equation (2). Since there was duplication of countries the degrees of freedom for calculating standard error were reduced to 190. The standard error of the average difference was 0.00177. If one assumes that similar variation occurs in the future as in the past then there is a 95% probability that the observed values of mean intensity will lie within the range $\pm 1.96 \times 0.00177 = \pm 0.00346$ of the projections of intensity. This corresponds to ± 1.34 per cent of 0.2587, the average observed intensity over the 192 data sets. Other sources of error are discussed later.

Another way of measuring error in the estimates is as follows. Of the list of 142 countries there are 82 for which there is available at 1990 both the observed and estimated number of households. The mean of the 82 observations for the number of households is 10531796 and for the estimated number of households is 10584517. This is a difference of 52721 while the standard error of the average difference is 150733. On a *t* test this difference is not significant.

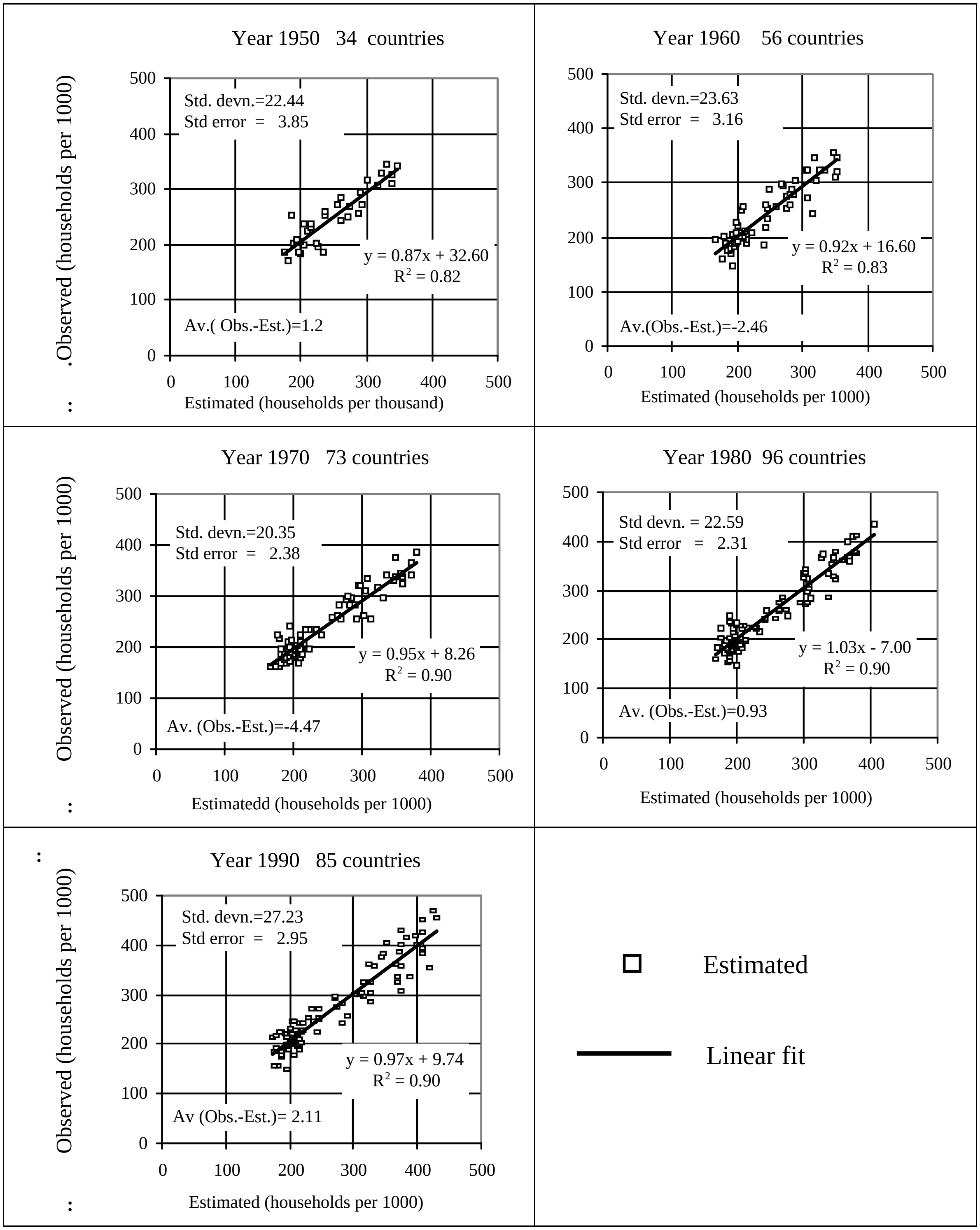
Equations (1) and (2) have coefficients of determination over 0.85 and utilise ratios which are central to representing population characteristics, ie fertility and ageing. It is anticipated that any other equations that might be sought to project household intensities would need to include *Y/M* and *E/M* as explanatory variables or the ratio *E*/*Y*.

A linear equation using the ratio *E*/*Y* as independent variable compares well with equations (1) and (2) as an estimate of household intensity over the range 0.25 to 1, see equation (3). If family households are associated with *Y*, ie larger households, and *E* is associated with smaller households, then this is not surprising, although for small *Y* some distortion may occur. For all the intensity data (344 cases) shown in Figure 2 the overall formula using *E*/*Y* is

$$h = 0.305(E/Y) + 0.170 \quad \text{with } R^2 = 0.89 \qquad (3)$$

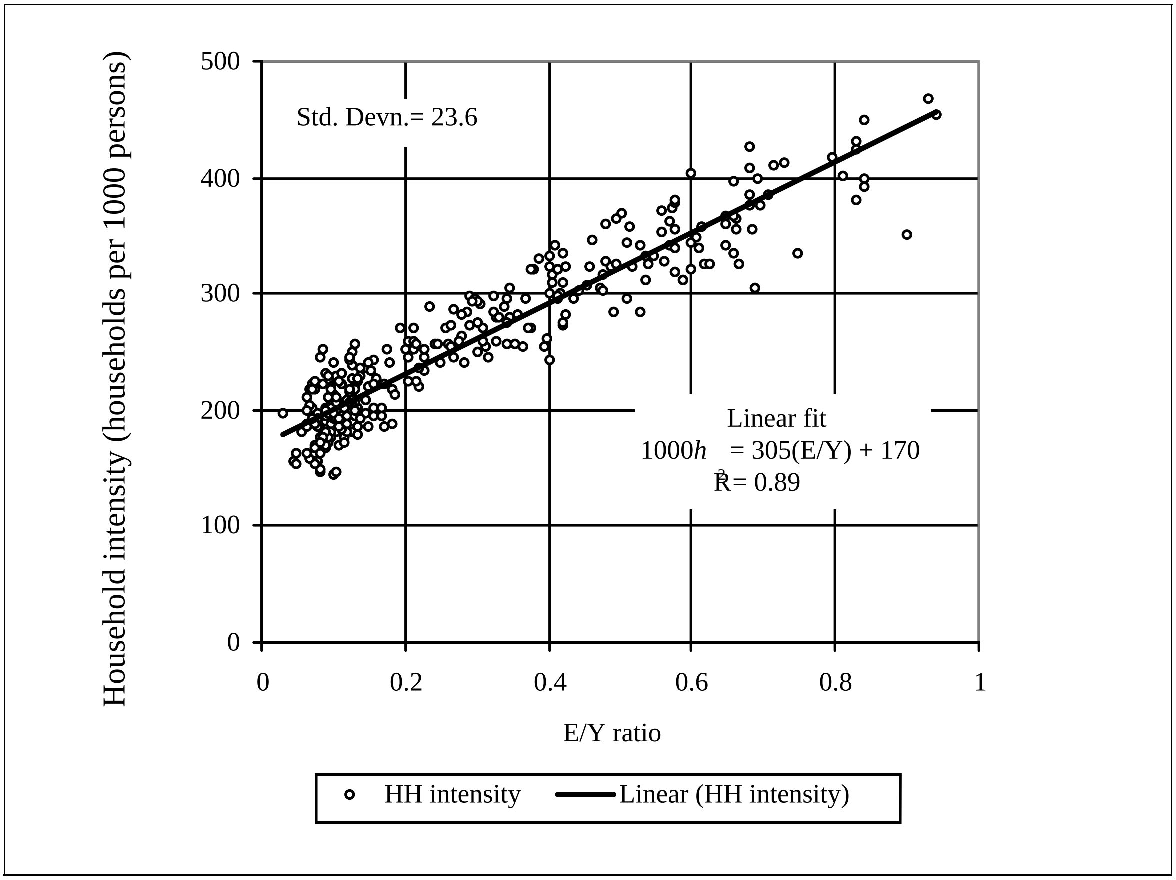
This relationship may be seen graphically by plotting the observed households per thousand against *E/Y* see Figure 3. This provides a convenient approximate estimator of household intensity. Equation (3) implies that a small percentage increase in (h-0.170) is equal to a small percentage increase in *E* less a small percentage increase in *Y*.

Figure 2 Observed household intensity (households per thousand) against estimated intensity for 34 countries. at 1950, 56 countries at 1960, 73 countries at 1970, 96 countries at 1980, and 85 countries at 1990.



Equation (3) is a more parsimonious model than the model involving equations (1) and (2), and has an equal, even slightly better coefficient of determination, and the ratio can be directly plotted as an independent variable. However the projections presented in this paper are based upon averaging the estimates obtained from equations (1) and (2) since these equations shows the separate contributions of the age components and will be less amplified if Y is small and E is larger than M.

Figure 3 Household intensity (households per 1000 persons) against E/.Y, 344 cases



For individual country projections a more precise estimate of the household numbers may be obtained as follows. Adjust the intercepts in equations (1) and (2) for each country so that the estimated number of households at 1990 equals the observed number. Use these adjusted equations to calculate the total number of households for all countries as before. This approach however implies a judgement concerning the appropriateness of the resulting individual country adjustments to the world model in the long term. It is beyond the scope of this paper to evaluate such judgements. The approach is also not parsimonious. This approach was therefore not used.

Upper bound to household intensity

In the projections through to 2050 there are 22 cases (listed below) where household intensities marginally exceed 700 per thousand. There were 19 low fertility cases at 2050 which had an average Y/M=0.30 and average E/M=0.95. A household intensity of 700 per thousand is equivalent to an average household size of 1.429 persons per household. This implies a very large proportion of one-person households and few 3+ person households. For an adjusted truncated Poisson distribution with this mean, 68.9 per cent would be size 1, 24.5 per cent size 2 and 6.6 per cent size 3+ .

We decided that 700 households per thousand would be used as an upper bound to all projections. This is equivalent to putting an upper bound to one-person households of two-thirds of all households with no more than 48.2 per cent of the population living alone (0.689/1.429 = 0.482)

Of the total of 1278 projections there were 22 cases in which the ceiling was applied. These were Italy for the low fertility variant at 2030 and Italy and Spain for the medium fertility variant at 2050. For the low fertility variants at 2050 the ceiling was applied to Austria, Belarus, Bulgaria, Czech Republic, Germany, Greece, Hong Kong (China), Hungary, Italy, Japan, Latvia, Portugal, Romania, Russian Federation, Slovakia, Slovenia, Spain, Switzerland., Ukraine, No high fertility variant cases reached the upper bound.

Projections of Household Intensities and Households

Equations (1) and (2) were then used to project household intensities for the 140 countries for which household count data are available together with Nigeria and Ethiopia. The rest of the world was treated as a quasi country. Thus there were a total of 143 'countries' which when added together gave the total for the world. The total population of the 142 countries was 5.04 billion at 1990 and the rest of the world 0.226 billion giving a total world population of 5.266 billion persons. The data for China included that for Taiwan.

For each of the 143 countries.10 household counts were obtained. These comprise an estimate of the number of households at 1990, and three projections for each of the year's 2010, 2030 and 2050 based upon the low, medium and high fertility scenarios of UN, (UN 1999 b). The total number of households was then calculated for each case.

Household intensity for the 'rest of the world' was obtained as follows. For each age category the population estimates were obtained by subtracting the relevant population counts for the total of 142 countries from the world estimates. The same projection procedure was then applied as for individual countries. The steps used to calculate the numbers of households of various sizes for the world are summarised in Table 2.

The results are shown in Table 3. and graphed in Figure 4. Habitat (1996) projections are included in the table for comparison. For 2010 the Habitat projections of household numbers are close to those in this paper but for 2030 and 2050 they are 8 per cent and 12 per cent lower respectively. Adjusting for the higher population projections used, which are based upon the medium variant of the UN 1994 population projections (UN 1994) they are for 2030 and 2050 10 per cent and 13 per cent lower. No detailed explanation is given in that reference as to the method for making these projections, except that it is noted that they are derived from individual country projections. The UN 1994 world population projections (medium variant) are higher than the UN 1998 projections by about 3.5 per cent for 2010, 7 per cent higher for 2030 and 10 per cent higher for 2050.

The projection of intensities by the Department of Infrastructure Victoria, DoI(2000) can be compared with projections using the method adopted here. At 1996 the DoI projections of household intensity were 6 per cent higher, at 2011 they were 2 per cent lower and at 2021 9 per cent lower. At 2021 household intensity projected by the DoI is 427 households per thousand.

Table 2 Outline of procedure for obtaining projections of household intensity

		Year 2010 2030 2050	Year 2010 2030 2050	Year 2010 2030 2050	Year 2010 2030 2050	Year 2010 2030 2050	Year 2010 2030 2050
1	Equation selection	Eqn 1	Eqn 2	Eqn 1	Eqn 2	Eqn 1	Eqn 2
2	UN fertility variant	Low	Low	Medium	Medium	High	High
3	Projection of household numbers by country	a	b	c	d	e	f
4	Mean of pair of projections of household numbers by country	=(a+b)/2		=(c+d)/2		=(e+f)/2	
5	Projection of household numbers for 142 countries grouped together	=(A+B)/2 =AB A is the sum over all 'a' , similarly B ,C etc		=(C+D)/2 =CD		=(E+F)/2 =EF	
6	Projection of household numbers for remainder of world treated as one quasi country	R(a)	R(b)	R(c)	R(d)	R(e)	R(f)
7	Mean of projection of household numbers for remainder of world	=(R(a) +R(b))/2 =RaRb		=(R(c) +R(d))/2 =RcRd		=(R(e) +R(f))/2 ReRf	
8	Forecast of world household numbers (8)= (5) + (7)	=AB+RaRb		=CD+RcRd		=EF+ReRf	
9	Household intensity	(AB+RaRb)/ Population		(CD+RcRd)/ Population		(EF+ReRf)/ Population	

Figure 4 shows a comparison of the projected movements in the overall world household intensity for the different fertility scenarios. For the low fertility variant case average world household intensity is expected to double by year 2050 to 507 households per thousand persons while in the high fertility case the intensity will increase by about 40 per cent to 358 households per 1000 persons.

For the three fertility variants the ratio of the highest to lowest projected household numbers at 2050 is 1.03, see Table 3. This contrasts with the much higher sensitivity of population projections to the fertility variants. The projected population for the low fertility projection is about 0.8 times the medium fertility projection and the high fertility projection about 1.2 times the medium fertility projection. Both these effects have been found by the authors to be carried through into most individual countries.

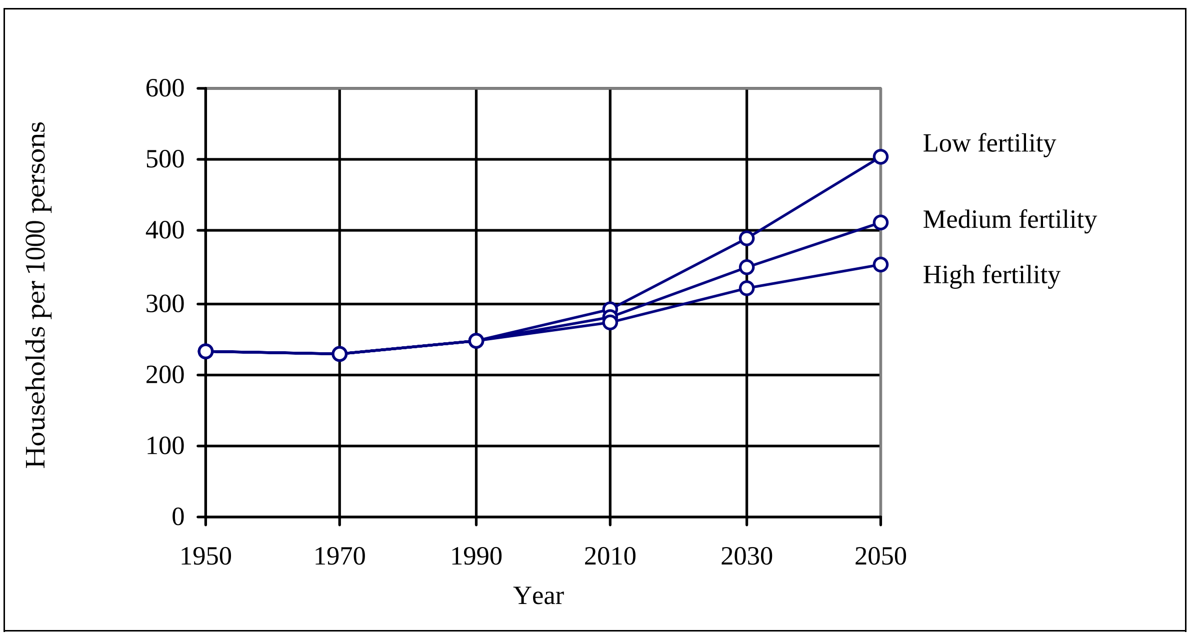
This result can be shown in summary by considering the annualised growth rates for households and population through to 2050 from 1990. In Figure 5 the growth rates were compared for each of the 142 countries described earlier. The dashed line indicates equal household and population growth rates. In every case, the household growth rate exceeds the population growth rate; no observation falls on or below the dashed line

However, the observations shift to the right, closer to the dashed line of equal growth rates as the scenario changes from low to medium to high fertility. The mean population growth rate is 0.6 per cent a year in the low fertility case, 0.9 per cent in the medium case and 1.2 per cent in the high case. In all these scenarios the household growth rate is around 1.9 per cent per annum. The graphs indicate that for the medium and low fertility variants, population growth rate has to be negative for household growth rate to be zero.

Table 3 Projected household population and total number of households for the world based upon the 1998 low, medium and high fertility population projections of the United Nations

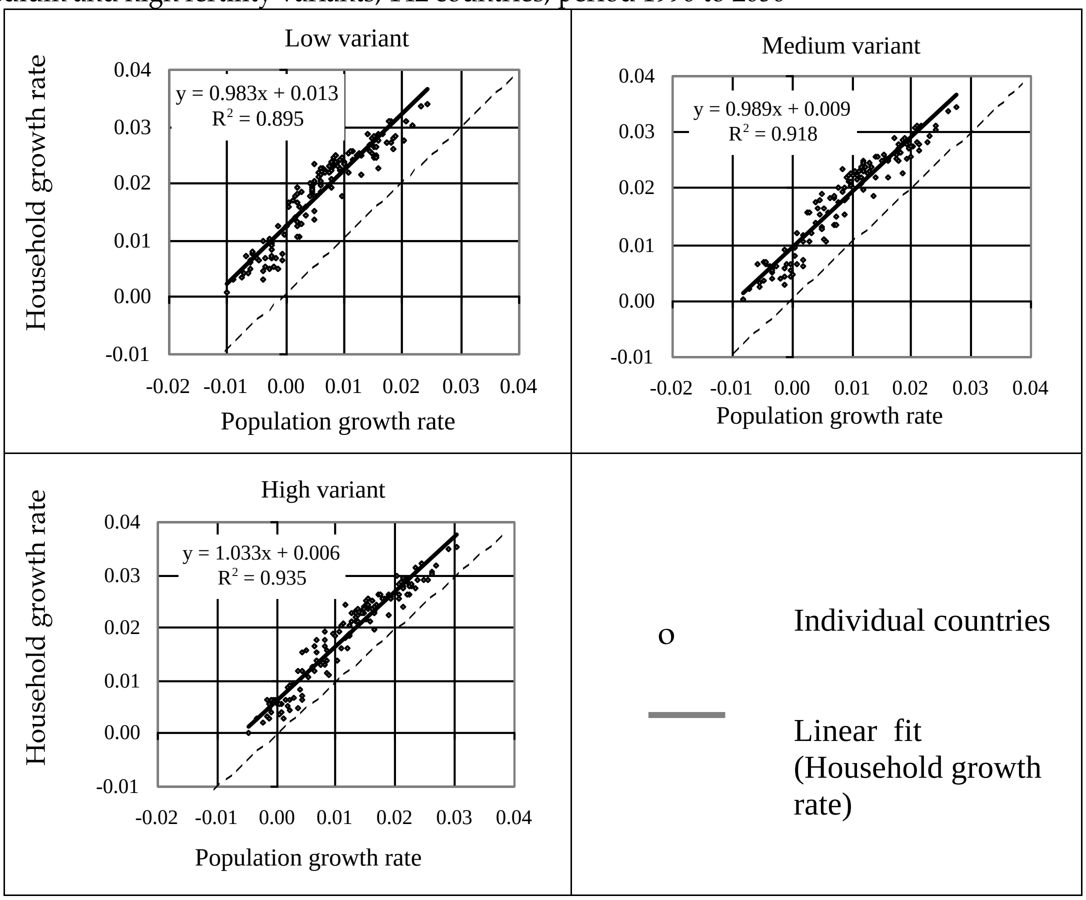
Year & fertility variant	Households	Population in Households	Household intensity	Households Mean size	Habitat 1996 Households
	Number	Number	Households per thousand	Persons per household	Number
1990 all	1 279 916 416	5 123 322 327	250	4.00	1 264 620 000
2010 LOW.	1 926 328 708	6 440 654 010	299	3.34	
2010 MED.	1 908 854 742	6 610 120 994	289	3.46	1 892 782 000
2010 HIGH.	1 896 719 082	6 776 962 325	280	3.57	
2030 LOW.	2 868 596 924	7 196 303 148	399	2.51	
2030 MED.	2 794 601 318	7 891 533 082	354	2.82	2 576 020 200
2030 HIGH	2 792 577 908	8 609 082 525	324	3.08	
2050 LOW	3 623 850 768	7 143 688 913	507	1.97	
2050 MED.	3 609 426 748	8 666 986 006	416	2.40	3 220 037 000
2050 HIGH	3 716 249 527	10 383 656 270	358	2.79	

Figure 4 Projections of world household intensity (households per thousand persons) by low, medium and high fertility variants



By 2050 for the low fertility variant the projections show that three quarters of the world's countries and about three quarters of the world's population will live in countries where household intensity exceeds that of most countries now ie 400 households per thousand. In the medium fertility variant the corresponding figure is one half and for the high fertility variant about a quarter.

Figure 5 Household growth rate per annum against population growth rate per annum for low, medium and high fertility variants, 142 countries, period 1990 to 2050



The Household Size Distribution Model

Household size distribution is modelled in this paper using the adjusted truncated Poisson distribution as seen in Figure 6. Following JLI (1999) for a particular number of households in a population there is an allocation of persons to households in a pseudo-random manner. This produces a distribution of the proportions of various sizes similar to the truncated Poisson distribution. See equation (7) in JLI (1999) shown here as equation (4).

$$F(k,\mu) = e^{-\lambda}\frac{\lambda^k}{k!}\frac{\mu}{\lambda} + L(k,\mu) \qquad (4)$$

where k is the number of persons per dwelling, μ is the average household size and λ is related to μ by the formula $\mu = \dfrac{\lambda}{1-e^{-\lambda}}$, The $L(k,\mu)$ are functions of μ satisfying $\sum_k L(k,\mu) = 0$.

As discussed earlier in this paper it was decided to use the inverse of the average household size, household intensity, *h*, as the parameter. Thus equation (4) becomes

$$F(k,h) = e^{-\lambda} \frac{\lambda^k}{k!} \frac{1}{\lambda h} + L(k,h) \qquad (5)$$

More precisely we take the difference between observed proportion of size *k* and the fitted truncated Poisson value for size *k* and regress that difference on the household intensity as a linear function of *h*. In symbols
$L(k,h) = \alpha_t + \beta_t h + \varepsilon_k$ for suitable constants α_k and β_k.

For size *k* we rewrite the observed proportion as O_k and the fitted truncated Poisson as t_k. Then the regression equation becomes

$$O_{kj} - t_{kj} = \alpha_k + \beta_k h_j + \varepsilon_{kj} \qquad (6)$$

where $h_j = \mu_j^{-1}$ is the household intensity for country *j*, and ε_{kj} is the error for country *j*. Given that *n* countries are used in the analysis we have the constraint

$$\sum_{j=1}^{n} \varepsilon_{kj} = 0 \text{ and in addition} \sum_{k=1}^{5+} \varepsilon_{kj} = 0 \qquad (7)$$

ie the total sum of discrepancies for all sizes for one country is zero.

The discrepancies of the estimates from the observed percentages shown in Figure 6 have been found by the authors to be approximately normally distributed and thus standard statistical tests can be applied. The values of α_k and β_k are found by ordinary least squares regression. As seen in Figure 6 there were 112 countries with at least one set of household distributional data. These data were used by the authors to determine these constants. In this way we can model the discrepancies between observed and theoretical proportions of sizes 1, 2, 3, 4, and 5+. Since the total number of persons in households and the total number of households are known, the value of '5+' can be calculated as being the ratio of persons in households not in households sizes 1 to 4 divided by the number of households of size 5 or more. The mean size of households of size 5+ has a different value for observed and truncated Poisson for each country. The resultant adjustments to the truncated Poisson are shown in equations (8) to (12).

The linear adjustments were

Size 1 $$L(1,h) = \frac{6.76 - 13.93h}{100}$$ R^2=0.10 s=0.0333 (8)

Size 2 $$L(2,h) = \frac{5.61 - 16.33h}{100}$$ R^2=0.25 s=0.0231 (9)

Size 3 $$L(3,h) = \frac{4.42 - 24.63h}{100}$$ R^2=0.56 s=0.0180 (10)

Size 4 $$L(4,h) = \frac{-6.72 + 27.69h}{100}$$ R^2=0.37 s=0.0299 (11)

Size 5+ $$L(5+,h) = \frac{-10.08 + 27.21h}{100}$$ R^2=0.36 s=0.0297 (12)

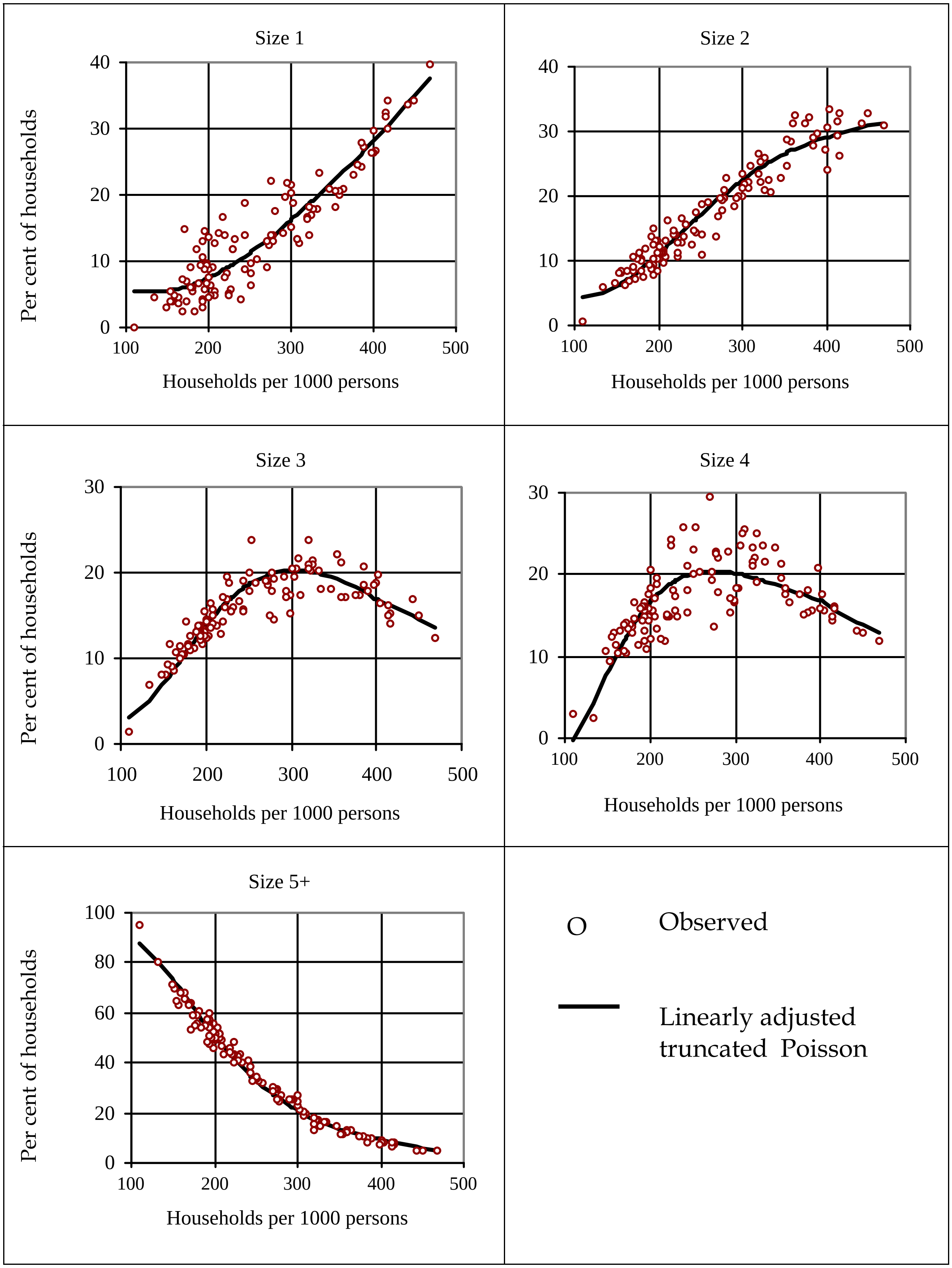
R^2 is the coefficient of determination and s is the standard deviation for the residual variation $\varepsilon(k)$.

Equations (8) to (12) are incorporated into truncated Poisson values for household intensity less than or equal to 0.4. For intensities greater than 0.4 (400 households per 1000 persons) the values of equations (8) to (12) at intensity of 0.4 are used, firstly since our observations show discrepancies stabilise at intensities above 0.4, and secondly this procedure maintains continuity of values.

The observed distribution of household sizes 1 to 5+ are plotted against household intensity as a percentage of total households for each of 112 countries for which data is available over the period 1960 to 1995, see Figure 6. The linearly adjusted truncated Poisson estimates are also shown.

Generally the estimated distributions follow the observed. It is noticeable that the fit improves for household intensities above 300 household per 1000 persons. As household intensity increases the percentages for size one and two increase while size 5+ shows a decrease. Sizes three and four have turning points at intensities between 250 and 300 households per 1000 persons.

Figure 6 Households of size 1 to 5+ as a percentage of total households by household intensity for each of 112 countries. Observed and linearly adjusted truncated Poisson estimates shown



Projections of Household Size Distributions

The projection of the distribution of household sizes for the world is obtained as follows. For each country the household intensity already obtained is substituted in equation (5) and equations (8) to (12) to obtain the estimated proportions of households of each size. Given the household population and household intensity the number of households of each size for each 'country' can be calculated. The number of households of each size are totalled across the 143 'countries' to give the world totals. The detailed procedure is outlined in Table 4. The results are summarised in Table 5 and plotted in Figure 7. The detailed projections for individual countries have been included as a series of appendix tables in a paper presented to the World Project LINK meeting in Oslo, Norway in October 2000 (Ironmonger, Jennings and Lloyd-Smith 2000).

Table 4 Outline of procedure for obtaining household projections by size

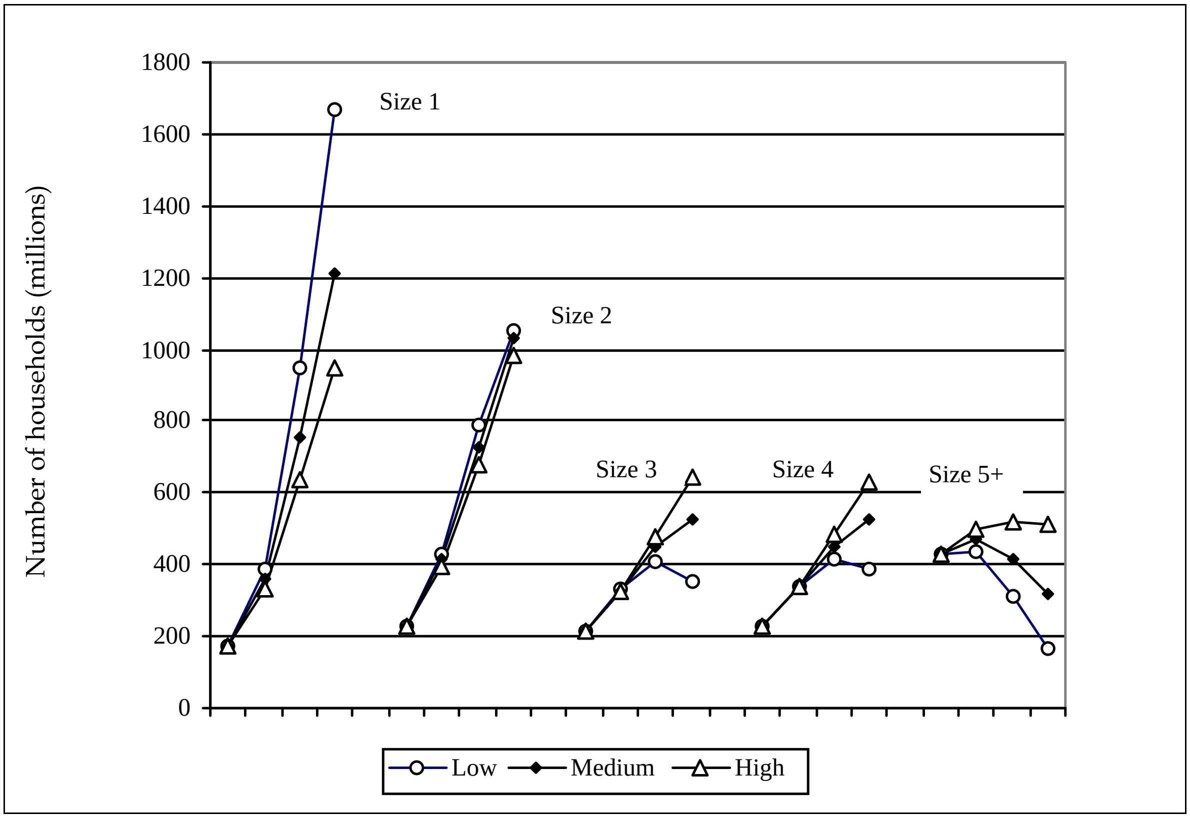
<table>
<tr><td></td><td></td><td>Year
2010
2030
2050</td><td>Year
2010
2030
2050</td><td>Year
2010
2030
2050</td><td>Year
2010
2030
2050</td><td>Year
2010
2030
2050</td><td>Year
2010
2030
2050</td></tr>
<tr><td>1</td><td>Equation selection</td><td>Eqn. 1</td><td>Eqn. 2</td><td>Eqn. 1</td><td>Eqn. 2</td><td>Eqn. 1</td><td>Eqn. 2</td></tr>
<tr><td>2</td><td>UN fertility variant</td><td>Low</td><td>Low</td><td>Medium</td><td>Medium</td><td>High</td><td>High</td></tr>
<tr><td>3</td><td>Forecast of household numbers by country</td><td>a</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td></tr>
<tr><td>4</td><td>Mean of pair of projections of household number s by country</td><td colspan="2">=(a+b)/2</td><td colspan="2">=(c+d)/2</td><td colspan="2">=(e+f)/2</td></tr>
<tr><td>5</td><td>Determine household intensity h by country including quasi country</td><td colspan="2">h(ab)</td><td colspan="2">h(cd)</td><td colspan="2">h(ef)</td></tr>
<tr><td>6</td><td>Use equation (5) to calculate per cent of sizes 1 to 5+ by country and hence number of households of each size</td><td colspan="2">Size 1 | Size 2 | Size 3 | Size 4 | Size 5+</td><td colspan="2">Size 1 | Size 2 | Size 3 | Size 4 | Size 5+</td><td colspan="2">Size 1 | Size 2 | Size 3 | Size 4 | Size 5+</td></tr>
<tr><td>7</td><td>Add all 143 countries together to obtain number of households of each size for world and hence world projection</td><td colspan="2">Size 1 | Size 2 | Size 3 | Size 4 | Size 5+</td><td colspan="2">Size 1 | Size 2 | Size 3 | Size 4 | Size 5+</td><td colspan="2">Size 1 | Size 2 | Size 3 | Size 4 | Size 5+</td></tr>
<tr><td>8</td><td>Check against total previously calculated in Table 4</td><td colspan="2">Sum over sizes
=(AB+RaRb)</td><td colspan="2">Sum over sizes
=(CD+RcRd)</td><td colspan="2">Sum over sizes
=(EF+ReRf)</td></tr>
</table>

Table 5 Projected number of households by size based upon the population projections of the UN World Population Prospects, The 1998 Revision (low, medium and high fertility variants)

Households	Year	Low Number	Medium Number	High Number	Low %	Medium %	High %
Size 1	1990	176 416 101	176 416 101	176 416 101	13.8	13.8	13.8
Size 1	2010	388 378 325	361 230 063	335 306 322	20.2	18.9	17.7
Size 1	2030	947 562 338	751 773 240	638 689 514	33.0	26.9	22.9
Size 1	2050	1 665 172 288	1 209 088 142	950 541 159	46.0	33.5	25.6
Size 1 ratio*		9.4	6.9	5.4			
Size 2	1990	228 902 744	228 902 744	228 902 744	17.9	17.9	17.9
Size 2	2010	431 607 679	411 950 459	395 370 372	22.4	21.6	20.8
Size 2	2030	790 765 568	724 655 812	677 636 968	27.6	25.9	24.3
Size 2	2050	1 052 916 058	1 031 212 828	984 253 517	29.1	28.6	26.5
Size 2 ratio*		4.6	4.5	4.3			
Size 3	1990	213 476 306	213 476 306	213 476 306	16.7	16.7	16.7
Size 3	2010	329 464 415	326 531 440	325 354 911	17.1	17.1	17.2
Size 3	2030	408 076 321	452 586 020	474 766 027	14.2	16.2	17.0
Size 3	2050	354 967 522	525 817 347	641 682 580	9.8	14.6	17.3
Size 3 ratio*		1.7	2.5	3.0			
Size 4	1990	230 890 548	230 890 548	230 890 548	18.0	18.0	18.0
Size 4	2010	338 584 452	338 447 494	339 526 750	17.6	17.7	17.9
Size 4	2030	413 050 061	452 124 589	482 665 435	14.4	16.2	17.3
Size 4	2050	386 096 184	523 715 674	629 689 276	10.7	14.5	16.9
Size 4 ratio		1.6	2.3	2.7			
Size 5+	1990	430 230 709	430 230 709	430 230 709	33.6	33.6	33.6
Size 5+	2010	438 293 838	470 695 285	501 160 727	22.8	24.7	26.4
Size 5+	2030	309 142 636	413 461 657	518 819 964	10.8	14.8	18.6
Size 5+	2050	164 698 716	319 592 757	510 082 995	4.5	8.9	13.7
Size 5+ ratio*		0.4	0.7	1.2			
All households	1990	1 279 916 409	1 279 916 409	1 279 916 409	100.0	100.0	100.0
All households	2010	1 926 328 708	1 908 854 742	1 896 719 082	100.0	100.0	100.0
All households	2030	2 868 596 924	2 794 601 318	2 792 577 908	100.0	100.0	100.0
All households	2050	3 623 850 768	3 609 426 748	3 716 249 527	100.0	100.0	100.0
All HH ratio*		2.8	2.8	2.9			
Popn. in hsehlds	1990	5 123 322 327	5 123 322 327	5 123 322 327			
Popn. in hsehlds	2010	6 440 654 010	6 610 120 994	6 776 962 325			
Popn. in hsehlds	2030	7 196 303 148	7 891 533 082	8 609 082 525			
Popn. in hsehlds	2050	7 143 688 913	8 666 986 006	10 383 656 270			
All popn. in HH ratio*		1.4	1.7	2.0			

Note:
ratio* is the ratio of the 2050 figure to the 1990 figure for the given variable.

Figure 7 Projection of the number of households (millions) in the world from 1990 to 2010, 2030, 2050: low, medium and high fertility variants.



Note:
All curves are a four point sequence beginning at the lowest point at 1990, then to 2010,, then to 2030, and then finally to 2050.

*Overall result*s

Figure 7 shows the dramatic increase in the size one and size two households over the 60 years from 1990 to 2050. The ratio of households of size 1 at 2050 is over nine times that at 1990 for the single person households. On the other hand the number of large households hardly increases at all. Overall the combined growth in households of size one and two is over five times while the growth in households of size 3+ is approximately two times.

The UN 60 year population growth projection to 2050 (UN 1999) varies widely from 39 per cent (low fertility) to 103 per cent (high fertility). However this paper's projection shows that for household growth there is only a variation from 182 per cent (medium fertility) to 190 per cent (high fertility). There is a very large rise in single person households of 1489 million in the low fertility scenario and about half that rise of 774 million in the high fertility scenario. For two person households the rises are simila*r* under different fertility scenarios (824 million, low and 755 million high). For size 3+ households there is virtually no rise for low fertility (31 million) but for high fertility the rise is 907 million. Thus the apparent similarity of total household growth under the various scenarios conceals a phenomenal switch in the size. For size one only, the authors have found that sensitivity to fertility scenarios shown for the world data carries through to the 112 countries for which separate data are available.

In terms of the population in households the change is equally dramatic. In 1990 there were 57 per cent of the world's population living in households of size 5+. By 2050 this is projected to reduce to 10 per cent for the low fertility variant and 29 per cent for the high fertility variant. This contrasts with an increase in the combined population in size one and two households from 12 per cent of the household population in 1990 to 53 per cent (low fertility) and 28 per cent (high fertility) in 2050.

Error

The confidence interval for the proportions of household size *k* is derived from two independent sources of variation. The first source is already calculated and is shown in equations (8) to (12) and is due to the error in the estimate of $L(k,h)$ for households of size *k*. However household intensity is subject to random variation as discussed above. For a given household size *k*, the adjusted truncated Poisson $F(k,\bar{h})$ depends on the household intensity through equations (5) and (6) subject to the constraints (7). Given an expected value and variance for the household intensity $\bar{h}$ and the estimate of $F(k,h)$, we can determine the expected value and variance for $F(k,\bar{h})$ by means of formulae described in Stuart and Ord (1987: 323-324). Thus the variance for the proportions of household size k given the variability in h, is the sum of these two sources of variance. The 95 per cent confidence limits for household percentages at an intensity of approximately 420 households per 1000 persons are: size 1, ± 3.0 per cent, for size 2, ± 1.5 per cent, for size 3, ± 2.2 per cent, for size 4 , ± 5.4 per cent, and for size 5+, ± 9.1 per cent. Confidence limits for size 5+ based upon population rather than households has been calculated by the authors to be approximately ± 1.5 per cent.

Discussion

The household concept and household size distribution

From the data analysed in this paper about 97 per cent of the world's population live in households. It appears to be a widespread characteristic of human beings that they distribute themselves as members of households, and in a manner well described by variants of the Poisson distribution depending only on average household size. This distribution has persisted over a wide variety of social and economic conditions around the world for at least 50 years with some earlier documented examples. In this paper for the period 1970 to 1990 the observed versus fitted household size distributions were compared, see Figure 6. In addition comparisons were made to determine if the same pattern of variability occurred over time for 54 countries with two time points a decade apart for which we have information on household size distribution. Similarly 25 countries with two time points twenty years apart were examined. We found that the patterns of variability were quite stable from 1950 to 1990. Possible explanations for this are given in JLI (1998). It is therefore likely that a similar pattern will persist for the next 50 years even though there may be substantial changes in economic, social, climatic and physical conditions.

The truncated Poisson distribution is a good candidate to be considered as the defining distributional characteristic of households. With the definition of households enlarged to include the zero size household, an empty dwelling as described in JLI (1999), household size is represented by the non-negative integers. The full Poisson distribution can then be used as a comprehensive framework for the size distribution of households.

On the basis of our Poisson model, for a given average household size it can be shown that the proportion of all households which are of size k (where $k \geq 0$) is precisely the proportion of persons (in households) who are members of households of size $k+1$, what we term a 'shadow effect'. In particular the proportion of size zero households is equal to the proportions of persons in size one households. This can be used to check vacancy rates and for other purposes.

Size zero households are important. In the random allocation of balls to cells model on which the truncated Poisson model is based, see JLI (1999), the proportion of empty cells (vacant dwellings or zero size households) is $e^{-\lambda}$. This can be taken as the implied vacancy rate on the assumption that the full household size distribution is a stable Poisson distribution. These dwellings must be accessible to households. For a more extensive discussion of vacancy rates under various pricing conditions see Igarashi (1991). The implied total stock of dwellings equals $1/(1-e^{-\lambda})$ or μ/λ times the number of households. For average household size 4.0 which is the estimated world average at 1990, see Table 3, the factor is 1.02, ie vacant dwellings are an extra two per cent on occupied dwellings. Whereas, for the medium fertility variant for the world for 2050 with an average household size near 2.4 the factor is 1.14, vacant dwellings are an extra 14 per cent of occupied dwellings. Thus following the model the trend to small households implies a substantial increase in the vacancy rate.

In this discussion it is assumed that a person lives in only one abode at a time although a person may over time or even over a month live sequentially in several residences or none. In some countries,eg India, houseless households represent a significant proportion of all households (Bhandare and Mukhopadhyay 1996:196). To a lesser extent this can occur in cities such as London (Editor, Focus on London 2000: 49-50). Bartie (2000) notes that a dwelling is defined in terms of building type whereas a household is defined in terms of characteristics of residents. A survey of household definitions over 35 countries in Europe is given in Keilman (1995: 113-116) and for the world in UN (1989: 3-10, 17-44) and in UN (1997: 98-104).

In Keilman (1995: 113-116) it is pointed out that one and two person household counts may be subject to reporting misallocations. Government regulation and social acceptability may encourage persons to report either as a member of a one person household or a two person household. Blurring of the distinction between membership of a one or two person household can also occur because of informal partnering which may occur with older persons, and dual households where there are separate bedrooms and common eating.

It is therefore possible to imagine an over-count of single person households in our projections, although the estimates shown in Figure 6 do not indicate that this is significant. It is interesting to compare the well documented projection to 2050 of single person households for the Netherlands (van Imhoff 1995:288) of 51 per cent (realistic scenario) with the projection of 51.6 per cent (medium fertility variant) for the Netherlands made for this paper.

Bias in data collection

Data are provided for countries that have data in a comparable form. See Table D in UN (1989: 17-44) for a comparison of definitions over many countries. Those countries that do not provide data are generally likely to be poor, and this may be associated with high birth rates. Thus there is likely to be some bias arising from the data used towards lower birth rates. However since over 80 per cent of the world's population is included in the individual country counts this bias is unlikely to substantially influence the results.

Population and household growth

The results show that population growth is much more sensitive to fertility assumptions than household growth, see Table 3. The effect of the variation in population projections is to alter average household size and hence the distribution of household sizes, rather than the total number of households. This effect has been found to carry through to the 142 countries used in the projections.

A likely explanation for this effect is as follows. Generally a country with a high population growth rate usually has high average household size whereas a country with a low population growth rate usually has low average household size, see Figure 4a in (JLI 1999). When projecting households these two factors tend to cancel. De Beer (1995:259-260) argues that in order to obtain the extremes for household numbers, rather than the stability observed, the population variants may need to be chosen differently to those required for choice of maximum high and low population numbers. In this paper the aim is to project household numbers based on a range of population projections rather than to find a maximum feasible range.

Fertility, mortality and immigration

There are considerable uncertainties in current long term projections of population. Lee (2000) reviewed 'The 1999 Technical Panel' report on assumptions and methods concerning projections of fertility, mortality and immigration particularly in relationships to the US social security system. Although there was a great deal of uncertainty in forecasting fertility the assumption on fertility was not changed. However in the light of bio-medical advances and statistical evidence a decision was made to raise the projected improvement in mortality.

It appears to the authors that the conditions leading to lower fertility rates such as uncertainties in income and activities through the life course, migration to cities and incompatible social and economic objectives are likely to become more prevalent. For a review of the many factors involved see UN (2000) and the discussion by McDonald (2000: 1). The assumptions in the world population projections used in this paper are discussed in UNFPA (1999:31-32). Demographic change through to 2050 in the European Union is discussed in Eurostat (1997). There is a wide ranging literature on fertility and ageing. A few examples particularly relevant to the projections in this paper are: low fertility rates in Europe in Lesthaeghe and Willems (1999); the long term effect of the timing of fertility decline on population size in O'Neill, Scherbov and Lutz (1999); UN World Population Projections to 2150 in UN (1998a).

A number of countries with previously high fertility have now commenced a downward path. For example Rashad (2000) in Table A1 shows that Arab countries (17 number) have shown an average drop in total fertility rate of about 30 per cent over the ten years from 1983 to 1993; Sather and Casterline (1998) identifies the onset of fertility transition in Pakistan. UNFPA (1999:29-30) indicates that from 1980 to 1990 total fertility rate has been dropping over many countries. The rate of change in fertility is indicated in 55 countries in UN Population Division (2000:43-44).

There appears to be a catch up effect to the developed world in fertility behaviour in some of the developing world. This is consistent with the trends shown in Figure 1 which provides an overview of the projected changes occurring in ratios of young and old with a general trend downwards in fertility. From a different point of view Goldstein and Schlag (1999) have discussed the evidence that mean age at childbearing is increasing at roughly the same rate as life expectancy. This implies that in a general way Y/M and E/M shown in Figure 1 are related.

Eurostat have completed projections of population through to 2050 for 18 European countries based upon low, baseline and high fertility scenarios (Eurostat 1997). The lowest limit chosen in the projections at 1995 was 1.2 for Spain yet in 1997 Spain had a total fertility rate of 1.15 (Eurostat 1999a:102-103). Comparing the baseline population projections against the UN medium variant case shows the Eurostat projections to be about 11 % higher than those of the UN. This is probably due to the higher fertility rates chosen.

Changes in social and political conditions which may herald a return to sustained higher fertility rates appear to be decades away. As in other slowly changing factors such as the environment, detrimental effects have to be demonstrated to a substantial proportion of the population before any major change occurs. For the purpose of this paper it is therefore considered that the medium to low fertility variants are more likely to be realised in the next fifty years than the high fertility variant.

The uncertainty in fertility projections contrasts with assumptions about life expectancy. While life expectancy is projected to increase the main uncertainty will be the degree to which the utilisation of medical knowledge and life style knowledge can prolong human life. Since most people would prefer a longer life the only possible reversal in the trend would be a major catastrophe. The UN assumptions do take into account the prevalence of diseases such as AIDS. See UN (1999a:4).

Examination of UN (1999b) shows that for the population projections for the world as a whole at 2010 for age 15+ population, for 2030 at age 35+ and at 2050 at age 55+ are very similar (and in many countries identical) for the three fertility scenarios. This implies mortality is independent of birth rates. However it is possible that some of the resources freed up because there are less children may be reallocated to helping prolong life for the elderly. If this were the case the UN projections of the elderly may need adjustment depending on the fertility variant.

International Migration

The sum total of international migration is of course zero. However the household projections for regions and the distributions of household sizes may be affected by regional differences in migration assumptions, and their age and income profile. The overall level of international migration is not large in proportion to the world population. Slotnik (1998) has pointed out that over the period 1965-1996 international migration has been nearly stable at about 2.1 to 2.3 per cent per annum of the world population.

Europe is projected to show decreases in population from about 722 million (1990) to 628 million (2050) for the medium fertility scenario and 550 million for the low fertility variant. But with a well-developed infrastructure, the possibility of its under-utilisation, and an ageing population, there may be some incentive for an increased migrant intake. This is despite the long held difficulties of reconciling different cultures (UN 2000:178) . 'Retirement' migration may increase with the ageing of the world's population. This is already evident in some countries and may become much more widespread. For the case of Australia see Bell (1992:300).

The concentration of household growth in urban regions.

Population growth in the future is projected to occur mainly in urban regions. For example in World Urbanisation Prospects (UN 1998b: 96-105) it is projected for the medium fertility variant (UN 1997a) that for the period 1990 to 2030 urban population will grow from 2.28 million to 5.10 billion while rural population will grow from 3.00 billion to 3.25 billion.

This implies a concentration of household growth in urban regions. However the data available on growth in the number of households in urban areas as distinct from population are not readily available world wide. This is perhaps partly because the definition of 'urban' varies (Habitat 1996:14-17). However there are a number of reasons why most of future household growth is expected to be urban areas, see for example Habitat (1996: 418,419). Grouping of households for efficiency and cost and social interaction create urban regions.

Currently about half of the world's population are in urban areas (UN 1998:2) and since urban households tend to be smaller than rural households at least half of all households are urban. Thus we can expect that the household growth rate in urban regions will be greater than that overall and as seen in Figure 7 much of this growth will be in small households. See also Table 27 in UN (1997b) which shows that average household size for those countries that provide data is generally smaller in urban areas than in rural areas. UNFPA (1999:39).also refers to the smaller family size in urban regions. However there is still work needed to clearly identify and define urban areas, see Brockerhoff (1999) and Richardson, Brunton and Roddis (1998).

Monitoring and future adjustment

New population and household information will come available from time to time. This should be incorporated into the data base and the equations recalibrated and/or augmented and projections updated. More sophisticated methods of analysis may then prove of benefit. As we have seen the sizes most sensitive to changes in population assumptions will probably be single and two person households, and size 5+ households. However the distinction between being a member of a one or two person household could become rather fuzzy. The current proportions of population by age group and household position as given in Eurostat (1999b: 42-49) may give an indications of the future for many countries undergoing a fertility transition.

The relatively rapid change in population characteristics from an *E/Y* of 0.25 to 1 will not necessarily produce uniform changes in household size across the world. Sachs (2000) has pointed out the large differences in income and technological adaptability of various countries. This may necessitate occasional revision in some of the projections.

Conclusion

Since the projection method used is rather different to those hitherto published the emphasis in this paper has been on methodology as the results indicate major changes in household distributions from the past. This paper sets down a procedure for transforming UN population projections of 142 individual countries and the ‘rest of the world’ into household projections. These projections are then assembled into world household projections.

It has been found that there is a strong relationship between household intensities and population ratios for data back to 1950. Examination of household size distributions from 1950 to 1995 also indicate that households or more precisely household groups (inclusive of singles and zero size households) appear to have a universal and characteristic pattern. This is best described by variants of a Poisson distribution depending only on mean household size. Projected numbers of households are not sensitive to differences between three UN population scenarios, while size distributions are sensitive to these differences. Given these scenarios the rate of growth of household numbers through to 2050 is shown to be almost independent of the population growth.

Most of this growth in households is expected to take place in urban areas. The Poisson model implies that vacant dwellings will be about an extra 14 per cent of occupied dwellings (medium variant) by 2050 in order to allow household movement to take place unimpeded. This additional implied requirement for housing can be easily overlooked in urban policy formulation.

The household is evidently a part of two worlds, the physical where via dwellings it is a node in a variety of infrastructure networks (roads, telecommunications, housing) and the social where it is an elemental group in communities. Thus change in household numbers and their size distribution can be seen as an important indicator of changes in society.

It is possible in the future that variables other than those related to the life course may become more important in influencing household size. Investigation of countries such as a number in Europe with current high household intensity may provide a pointer for other countries as to how socio-economic factors may interact with fertility, mortality, household size and type. The upper limit of 700 households per thousand persons chosen for the calculations would warrant investigation. It would be of interest to investigate why household size distributions are generally closer to observed for household intensities greater than 300 households per 1000 persons. Finally household intensities in individual countries differ from the projections provided by the global equations. It would be worthwhile to investigate whether these differences are likely to persist for individual countries.

References

Australian Bureau of Statistics (ABS) 1996 *Household Estimates, Australia 1986, 1991- 1994 Catalogue No. 3229.0*. Canberra: Australian Bureau of Statistics

Australian Bureau of Statistics (ABS) 1999 *Household and Family Projections, Australia 1996 to 2021 Catalogue No. 3236.0*. Canberra: Australian Bureau of Statistics

Bartie, D. 2000. *Production of Post-Censal Small Area Estimates of Households; a Preliminary Investigation. Demography Working Paper 2000/1*. Canberra: Australian Bureau of Statistics.

Bell, M. 1992. *International Migration in Australia 1981-198*6 Produced for the Joint Commonwealth/State /Territory Population and Immigration Research Program, Bureau of Immigration Research. Canberra: Australian Government Publishing Service.

Bhandare, S.S. and J.K.Mukhopadhyay (eds.) 1996 *Statistical Outline of India 1996-97*. Mumbai: Tata Services Limited, Department of Economics and Statistics.

Brockerhoff, M. 1999. Urban Growth in Developing Countries: A Review of Projections and Predictions. *Population and Development Review* 25(4):757-778.

Burch, T. K. 1995. Theories of Household Formation : Progress and Challenges. In E. van Imhoff, et al. (eds). *Household demography and household modelling*. New York: Plenum Press.

David, F.N. 1962. *Games, Gods and Gambling.* London: Charles Griffin & Co. Ltd.

De Beer, J. 1995. National Household Forecasts for the Netherlands. In van Imhoff, et al. (eds). *Household demography and household modelling*. New York: Plenum Press.

Department of Infrastructure Victoria 2000. *Victoria in future: Methodology and Assumptions*. Melbourne: Department of Infrastructure.

Editor, Focus on London 2000. *Focus on London 2000.* London: Social and Regional Division, Office for National Statistics.

Eurostat 1997. *Statistics in Focus: Population and social conditions 1997-7.* Luxembourg: Statistical Office of the European Communities.

Eurostat 1999a. *Demographic Statistics: Data 1960-99 1999 Edition*. Luxembourg: Statistical Office of the European Communities.

Eurostat 1999b. *Demographic Statistics: Data 1995-98 1998 Edition*. Luxembourg: Statistical Office of the European Communities.

Goldstein, J. R. and S. Schlag 1999. Longer Life and Population Growth. *Population and Development Review* Vol. 25, No. 4 December 1999: 741-747.

Habitat (United Nations Centre for Human Settlements) 1996. *An Urbanizing World: Global Report on Human Settlements, 1996*. Oxford: Oxford University Press.

Ironmonger, D.S., Jennings, V.E. and Lloyd-Smith, C. W. 2000 Long Term Global Projections of Household Numbers and Size Distributions for LINK Countries and Regions. *Paper presented at the Project LINK meeting*. Oslo, Norway on 3-6 October 2000.

Ironmonger, D.S. and C.W. Lloyd-Smith. 1992. Projections of households and household populations by household size propensities. *Journal of the Australian Population Association* 9(2):153-171.

Jennings, V.E., Lloyd-Smith, C. W. and Ironmonger, D.S. 1998. *A New Method in Household Demography*. Research Paper Number 624, Melbourne: Department of Economics, University of Melbourne.

Jennings, V.E. , Lloyd-Smith, C. W. and Ironmonger, D.S. 1999. Household size and the Poisson Distribution. *Journal of the Australian Population Association*, 16(1/2):65-82.

Keilman, N., 1995. Household Concepts and Household Definitions. In van Imhoff, *et al.* (eds). *Household demography and household modelling*. New York: Plenum Press.

Lee, R. 2000. Long –Term Population Projections and the US Social Security System. *Population and Development Review* 26(1):137-143.

Lesthaeghe, R. and P. Willems 1999. Is Low Fertility a Temporary Phenomenon in the European Union? *Population and Development Review* 25(2):211-228.

McDonald, P. 2000. Gender Equity, Social Institutions and the Future of Fertility. *Journal of Population Research* 17(1):1-16.

O'Neill, B.C., S. Scherbov and W. Lutz 1999. The Long Term Effect of the Timing of Fertility Decline on Population Size *Population and Development Review* 25(4): 749-756.

Rashad, Hoda 2000. Demographic Transition in Arab Countries. *Journal of Population Research* 17(1):83-101.
Rescher, N. 1998. *Predicting the Future.* Albany, N.Y. State University of New York Press,
Richardson, A.J. , P.J. Brunton and S.M.Roddis. 1998. The calculation of perceived residential density. *Road and Transport Research.* 7(2):3-15.
Sachs, J 2000. Sachs on Globalisation: A new map of the world *The Economist* June 24th, 2000 :95-97.
Sathar, Z. A. and J.B. Casterline. 1998. The Onset of Fertility Transition in Pakistan. *Population and Development Review*. 24(4): 773-796.
Slotnik, H. 1998. International Migration 1965-96: An Overview. *Population and Development Review.* 24(3): 429-468.
Stuart, A. and J.K.Ord. 1987. *Kendall's Advanced Theory of Statistics.* Volume 1, Fifth Edition. London: Griffin
United Nations (UN) 1955-1997. *1955, 62, 63, 71, 73, 76, 82, 87, 90, 95 Demographic Year Book.* New York.
UNFPA 1999.*The State of World Population 1999* United Nations Population Fund. New York.
United Nations (UN) 1989. *1987 Demographic Year Book*. New York.
United Nations (UN) 1994. *The Sex and Age Distribution of the World Populations: The 1994 Revision.* New York.
United Nations (UN) 1995. *World Population Prospects: The 1994 Revision*. New York.
United Nations (UN) 1997a. *The Sex and Age distribution of the World Population: The 1996 Revision.* New York.
United Nations (UN) 1997b. *1995 Demographic Yearbook*. New York.
United Nations (UN) 1998a. *World Population Projections to 2150*. New York.
United Nations (UN) 1998b. *World Urbanization Prospects: The 1996 Revision*. New York.
United Nations (UN) 1999a. *World Population Prospects: The 1998 Revision*, Volume 1: Comprehensive Tables. New York.
United Nations (UN) 1999b. *World Population Prospects: The 1998 Revision*, Volume II: Sex and age. New York.
United Nations (UN) Population Division 2000.*Below Replacement Fertility.* Population Bulletin of the United Nations Special Issue Nos. 40/41 1999. New York.
Van Imhoff E. 1995. LIPRO: A multistate household projection model. In van Imhoff, *et al.* (eds). *Household demography and household modelling*. New York: Plenum Press.

Table A1 Countries used in the 143 country data set. Reference ‘1’ is to those countries which are also used in the 2 × 96 country data set , and reference ‘d’ is to those countries which are also used in the 112 country household distribution data set

Afghanistan	d	Dominican Rep	1	Madagascar	1	Russian Fed.	d
Algeria	1	Ecuador	1	Malawi		Rwanda	1,d
Argentina	1,d	Egypt	1,d	Malaysia	1,d	Samoa	
Australia	1,d	El Salvador	1	Maldives		Saudi Arabia	d
Austria	1,d	Fiji	1	Mali	1,d	Singapore	1,d
Bahamas, The	1,d	Finland	1,d	Malta	d	Slovakia	d
Bahrain	1,d	France	1,d	Martinique	1,d	Slovenia	d
Bangladesh	1,d	Ethiopia		Mauritania		Solomon Isles	1
Barbados	1,d	Fr. Polynesia	d	Mauritius	1,d	South Africa	
Belarus		Germany	1,d	Mexico	1,d	Spain	1,d
Belgium	1,d	Greece	1,d	Moldova Rep.		Sri Lanka	d
Belize	1	Guadeloupe	1,d	Morocco	1,d	Sudan	d
Benin	1	Guam	1,d	Mozambique	d	Sweden	1,d
Bolivia	1,d	Guatemala	1,d	Myanmar	d	Switzerland	1,d
Botswana	1,d	Guinea		Namibia	d	Syrian Arab Rep.	1,d
Brazil	1,d	Guyana	1,d	Nepal	1,d	Tajikistan	d
Brunei	1	Hungary	1,d	Neth. Antilles	1,d	Tanzania U.Rep.	1
Bulgaria	1,d	India	1,d	Netherlands	1,d	Thailand	1,d
Burundi	1,d	Indonesia	1,d	New Caledonia		Trinidad &	1,d
Cameroon	d	Iran	1,d	New Zealand	1,d	- Tobago	
Canada	1,d	Iraq	d	Niger	d	Tunisia	1
Cape Verde	d	Ireland	1,d	Nigeria		Turkey	1,d
Cen.African Rep.	d	Israel	1,d	Norway	1,d	Uganda	d
Chile	1,d	Italy	1,d	Oman		Ukraine	1,d
China	1,d	Jamaica	1	Pakistan	1,d	United Arab	d
China, Hong Kong	1,d	Japan	1,d	Panama	1,d	-Emirates	
– SAR		Jordan	d	Papua - N.G.		United Kingdom	1,d
Columbia	1	Kazakhstan	d	Paraguay	1,d	United States	1,d
Comoros		Kenya		Peru	1,d	Uruguay	1,d
Congo	d	Korea Rep.	1,d	Philippines	1,d	Vanuatu	1,d
Costa Rica	1,d	Kuwait	1,d	Poland	1,d	Venezuela	1,d
Cote d'Ivoire	d	Kyrgyzstan		Portugal	1,d	Viet Nam	1,d
Croatia	d	Latvia		Puerto Rico	1,d	Yemen	
Cuba	1,d	Luxembourg	1,d	Qatar		Yugoslavia	1,d
Cyprus	1,d	Macau	1,d	Reunion	1,d	Zambia	1,d
Czech R.	1,d	Macedonia -		Romania	1,d	Zimbabwe	
Denmark	1,d	TFY				Rest of world	

Note:

There are five countries not listed separately in United Nations Population Prospects , the 1998 Revision,which are included in the 112 country household distribution set and not listed above. These are French Guinea, Kiribati, Saint Lucia, Seychelles and Tonga.