International Institute for
Applied Systems Analysis
Schlossplatz 1
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342
Fax: +43 2236 71313
E-mail: publications@iiasa.ac.at
Web: www.iiasa.ac.at

Interim Report IR-08-002

Probabilistic Household Projections based on an Extension of the Headship Rates Method with an Application to the Case of Russia

Sergei Scherbov (scherbov@iiasa.ac.at)
Dalkhat Ediev (dalkhat.ediev@oeaw.ac.at)

Approved by

Wolfgang Lutz (lutz@iiasa.ac.at)
Leader, World Population Program

January 28, 2008

International Institute for Applied Systems Analysis
Registration number: ZVR 524808900

Contents

Abstract

The paper presents a probabilistic method for projecting the number of households and their distribution by size. The method combines probabilistic population projection with a probabilistic headship rates model. In order to distribute the households by size, we use recently developed models for conditional proportions of households of different sizes among households of the same size or bigger. Models are approbated on the case of Russia with the fertility scenario assuming the considerable success of demographic policies recently introduced in the country. The parameters for household models are estimated from the 2002 census using bootstrap procedures and from assumptions about the possibility of headship rates at young ages to reach levels reported for Sweden. Our results show significant changes in the future distribution of private households in Russia. Despite the overall decline in the number of households, our results imply a persistent shortage of housing infrastructure for four-person households. Typically these would be the households of two parents with two children, i.e., those families that are the focus of the recently introduced demographic policies.

Acknowledgments

The paper presents joint work between IIASA and the Vienna Institute of Demography of the Austrian Academy of Sciences which was supported in part by the EU TACIS project, grant contract 2006/121-795.

The authors would like to thank Wolfgang Lutz for his valuable suggestions and comments.

About the Authors

Sergei Scherbov is a Senior Research Scholar at IIASA's World Population Program and Research Group Leader at the Vienna Institute of Demography of the Austrian Academy of Sciences.

Dalkhat Ediev is a Senior Scientist at the Vienna Institute of Demography of the Austrian Academy of Sciences.

Probabilistic Household Projections based on an Extension of the Headship Rates Method with an Application to the Case of Russia

Sergei Scherbov and Dalkhat Ediev

1 Introduction

Household projections are important for planning purposes and for analyzing the implications of population dynamics for consumption, labor, ecology, etc. (MacKellar et al. 1995; O'Neill and Chen 2002; Perz 2001; Prskawetz et al. 2004). In some areas, like housing and urban planning, projections of the distribution of households by size are of key importance, as they are relevant to decisions involving substantial long-term public and private expenditures. It is important for such applications to have a better knowledge of the trends expected in the future as well as the *uncertainty* accompanying such trends. Understanding the uncertainty of a household's prospects is important, as it is not always possible to easily and quickly adjust investment decisions to deviations from the projected trend. This issue is also important for developing demographic and taxation policies based on family and household composition. This work utilizes recent advances in modeling household distribution by size and in probabilistic projections to develop probabilistic household projections for Russia in 2005-2050.

In household projections, internally consistent reflection of the expected future is of core importance. In a scenario with fertility rising due to successful policy based on second-order births, the number of households with four and more members must rise. Similarly, in an aging population, two- and single-person households should increase in number due to an increased proportion of elderly. Deriving consistent distributions of households by type and of persons by household position is not a simple task even when these distributions are based on census data, i.e., when we know that the data source is consistent by itself. It is even less easy in the case of projections, as the projected, i.e., artificial, population proportions may lack consistency. A common way to tackle such inconsistencies would be to apply some adjustments or reconciliation procedures, which restore consistency with respect to the most important distributions at least. The problems are worse in the case of *probabilistic* projections: First, it is easy to get unrealistic population structures while generating future populations at random and, second, the application of reconciliation procedures is not a desired option as it may uncontrollably distort the probabilistic assumptions upon which the projection is built. In view of these problems and keeping in mind the importance of practical applications, we build our work on the simplest method for household projection and confine the complexity of the output to the number of households and their distribution by size only. We use the most traditional and simple headship rates method based on *age-*

specific headship rates, i.e., on proportions of household heads, without considering the status of the head, except for the age. In particular, we do not differentiate heads by sex and by type or size of the household, as the headship rates broken down by these characteristics would be erratic and highly correlated to fertility and mortality assumptions as well as to the cultural context. Age-specific headship rates, in contrast, are more stable and less correlated to projection assumptions. We use these rates to generate the overall number and, consequently, the average household size. From the average size, the distribution of households by size is generated, using recently developed models. Such a strategy pays off, as the population size and, therefore, the average size and distribution of households by size adequately reflect the implications of the population projection assumptions.

The case of Russia deserves special attention for several reasons. During the last century the country passed through many dramatic social and economic disturbances, which made deep imprints in the age structure of the population and have serious implications for demographic prospects (Ediev 2001). An almost inevitable de-population of the country and changes in the age structure may have significant and, sometimes, contradictive effects on the prospects of households of different sizes. At the same time, the country is facing an urgent need for better planning and improvement of the infrastructure and living arrangements in particular, hence, the importance of understanding the prospects and uncertainty of household dynamics in the future.

2 Methodology and Data

This work is based on an extension of the conventional headship rates method (United States National Resources Planning Committee 1938; United Nations 1973). Several rationales support this choice. The headship rates method and its extensions are widely used by government agencies, despite progress in more sophisticated modeling of households. Age-specific headship rates happen to be remarkably stable indices, which vary only moderately despite significant demographic developments observed in many populations. Changes in fertility and mortality have only a limited effect on headship rates. Population age structure is a primary source of variations of household numbers and distribution (see the Appendix for some analytical results in support of this view and implications for a stable population). At older ages mortality and morbidity may play a more significant role as a factor of headship rates dynamics. Yet this effect may be neglected in a study of the overall number of households, to which this work is devoted. More importantly, headship rates at young ages may increase considerably due to earlier separation of youth from parental households. We take the latter effect into account, assuming a possible increase in headship rates at young ages.

The need for probabilistic household forecasting has been acknowledged elsewhere (see, for example, De Beer and Alders 1999; Leiwen and O’Neill 2004). De Beer and Alders forecast uncertainty in the future number of households and introduce a number of assumptions regarding institutional population, probability of changes in the age at leaving the parental home, assumptions about the conditional probability of changes in the percentage of people living alone, etc. In order to derive these assumptions, a very good information base should be available. From the data available for Russia, deriving such a distribution would require too much subjective judgment. Leiwen and O’Neill propose an extension of the headship rates method by introducing age- and household size-specific headship rates. The latter rates were proposed to be derived as functions of demographic indicators, such as the propensity of leaving home, marriage, divorce, fertility rates, and mortality. Such an approach seems to be promising, as it demands less data and fewer model assumptions compared to the micro-simulation approach. It also allows us to address the role of demographic events in a household's formation. In some applications, however, there might not be enough data for the model. Also, the extension to the method may require special reconciliation procedures in order to guarantee internal consistency of the projection, which may limit its application, especially in probabilistic projections. For example, the total population in private households obtained from their model distribution by size might be inconsistent with the size of the actual population in private households. Another potential drawback is the use of parameters, which are quite volatile and involve non-trivial correlations between them. For example, size/age-specific headship rates may vary considerably across time and regions, depending on the prevailing fertility levels, while age-specific headship rates derived regardless of the household characteristics are usually much more stable, i.e., the former rates are negatively correlated. Usage of such model parameters may worsen performance and robustness of the probabilistic model and increase demand for data availability and quality.

This paper presents an approach that is based on the extension of the headship rates method. The extension we use is based on deriving the distribution of households by size from the overall average household size, which, in turn, is derived from the conventional age-specific headship rates. The approach was proposed by Gisser (1986a, 1986b) and has been used in Austrian household projections ever since. One advantageous feature of the approach is that the average household size indirectly reflects the demographic developments, even though headship rates might be less sensitive to those developments. In particular, changes in fertility assumed in population projection will, in fact, affect population size and, thereby, the average household size. Unfortunately, like many other extensions of the headship rates method, the approach may eventually result in inconsistent projections, and special reconciliation procedures are to be used, which somewhat limits its usage in probabilistic projections. For example, the sum of the proportions of households of different sizes may deviate from one, or the population totals obtained directly from the age structure or from the distribution of households by size may differ considerably. Merits of the approach may be used to a wider extent, based on recent developments of models for conditional shares of households among households of the same or larger size and for average sizes of such households (Ediev 2007) (see details below).

We use conventional age-specific headship rates (eventually generated at random, as described below) to derive the number of households from the projected population by age and to obtain the average household size. Then we apply a conditional shares approach to derive the number of households by size.

The general scheme of household projection adopted in this paper is presented in Chart 1. Basically, two tasks are identified: making population projections and projecting the number of households by size, based on population projections. In the case of probabilistic projections, this sequence is repeated a given number of times.

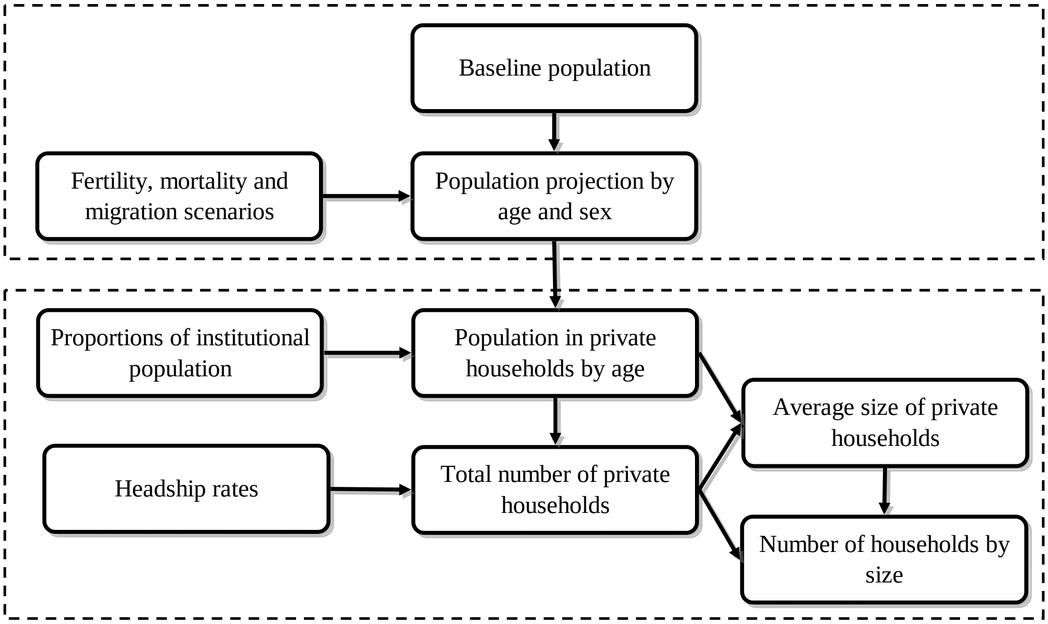


Chart 1. Projection of the number of households.

2.1 Population projections

Population projections were prepared using a probabilistic approach. This approach has been applied successfully in many instances to project the population at national, macro-region and global levels (Lutz et al. 1997, 2001, 2004; Lutz and Scherbov 1998; Keilman et al. 2002). Three main approaches to probabilistic projections are proposed in the scientific literature. The first approach is based on the time-series analysis of past vital rates; the second is based on the analysis of past projection errors; and the third is based on expert opinion. A good overview of these approaches is given by Bongaarts and Bulatao (2000) and Booth (2006). The three approaches are not mutually exclusive but often complementary. In particular, expert judgment is implicitly or explicitly considered in all of them. The third approach, the one adopted here, explicitly uses expert opinion. Expert-based population projections were first proposed in the scientific literature by Lutz et al. (1996). Further use and development of the method can be found in Lutz et al. (1997, 1999, 2001, 2004) and Lutz and Scherbov (1998).

There are many sources of uncertainty in the future development of fertility, mortality and migration. The recent introduction of a new demographic policy in Russia makes the situation even more uncertain. The main aim of the policy was to increase the number of second births. It is not clear how the population will react to the new measures aimed at fertility stimulation. It is not clear whether the number of second births will increase in cohorts of women or whether a shift in the birth calendar will occur without essential changes in the completed fertility of cohorts.

In our projections we assumed that population policy will bring certain positive results. We assumed that this will lead, first, to a shortening of the interval between first and second births and, second, to an increase in the number of second births by 50 percent. Those assumptions result in the increase of projected mean values of period

TFR (total fertility rate) to 1.5 in 2008, peaking at 1.76 in 2014 and declining afterwards and remaining constant at the level of 1.7 starting from 2027. The range of uncertainty in 2050 covers the interval of TFR from 1.25 to 2.15 children per woman.

In the case of life expectancy, we assumed that the lower end of the 90 percent range corresponds to no future increase in life expectancy for both males and females. The upper end corresponds to growth in life expectancy of about two years per decade for females and 2.8 years per decade for males, thus decreasing the gap that exists between life expectancy of males and females. This result in mean predicted values of life expectancy in 2050 equals to 71.3 and 81.7 years of life for males and females, respectively.

The mean predicted value of the number of net migrants was considered constant and equal to 126,000 people coming annually to Russia. The range between 0 and 256,000 is assumed to cover 90 percent of all the future outcomes of net migration.

In order to generate the required distributions of the future path of fertility, mortality and migration, we adopted the method used by Lutz et al. (2001). The starting year of projection was 2005. The data on population, fertility, mortality and migration for this year were utilized. Age-specific fertility rates were preliminarily smoothed using the mixed Gamma distribution function. Age-specific mortality rates were smoothed using the Heligman-Pollard mortality schedule. Projections were made for single-year age groups and are thus carried out on a yearly basis.

2.2 Deriving the number of households by size

After making the population projections, the next step is to obtain the total number of private households. To do that we apply age-specific institutional population proportions to the projected population in order to obtain the population living in private households. Then we apply age-specific headship rates to the population living in private households in order to get the total number of private households (Chart 1). Proportions of the institutional population were fixed at the level observed in the 2002 census.[1] Headship rates for the starting year are derived from 2002 census data using the bootstrap method. In order to avoid biases caused by artificial geographic compositions generated in bootstrap, we pull stratified samples, pooling together regions with similar average sizes of households. Two groups of regions were defined: those with an average household size below 3 persons and those above 3. After 2005, headship rates at young ages are allowed to increase. This is probabilistically done in two steps. First, for 2050 we set a target value for headship rates distributed uniformly between the lower and upper levels, which are derived consequently from the level set for 2005 and from the data on headship rates for Sweden (United Nations 1997). For the years between 2005 and 2050, the headship rates are set to follow linear increasing trends.

Based on the generated number of households, the average size is calculated as the ratio of the number of households to the population in private households. The α method from (Ediev 2007) is applied to size after size:

---

[1] The number of people in institutional households comprises 1.6 percent of all the population. This percentage varies across age and sex.

$$\nu_{k/k+} = e^{-\alpha_k \cdot \eta_k}, \tag{1}$$

where $\nu_{k/k+}$ is the conditional share of households with *k* members among households of the same or larger size, η_k is the average size of such households minus *k*, and α_k are model parameters. The parameters α_k are obtained from regressions against the average household size, which are also derived from the bootstrap procedure based on stratified data of the 1994 micro-census.

The procedure begins with the smallest households, i.e., one-person households. The average size for households of the next size is obtained recurrently by subtracting the number and the population residing in the households of the preceding size.

$$\eta_{k+1} = \frac{N_{k+} - k \cdot H_k}{H_{k+} - H_k} - (k+1) = \frac{\eta_k}{1 - \nu_{k/k+}} - 1, \tag{2}$$

where H_k and H_{k+} are the numbers of households of size *k* and of the same or larger size; N_{k+} is the population residing in households with *k* or more members.

3 Results

To study the sensitivity of the results to assumptions concerning headship rates, we used three different approaches in developing our projections. The first approach was based on directly applying headship rates obtained from the Russian census. Our probabilistic projection set contained 1,000 simulations. For each simulation we stored age-specific population distributions for every projected year. Then we applied fixed age-specific headship rates obtained for Russia as a whole to each population composition, thus deriving the total number of households.[2] In the next step we calculated the average size of the household for each simulation and distributed the total number of households by the number of households of each size. Since we used probabilistic age-specific population distributions, we also obtained the probabilistic distribution of the number of households of different sizes.

In a second approach, the algorithm of the distribution of the number of households by size was similar, except we used random headship rates. They were obtained using the bootstrap procedure described above and without using the Swedish rates to set target values for headship rates in 2050. In the third approach, we used headship rates from the bootstrap procedure for the base year and allowed for a possible increase by 2050 as described above.

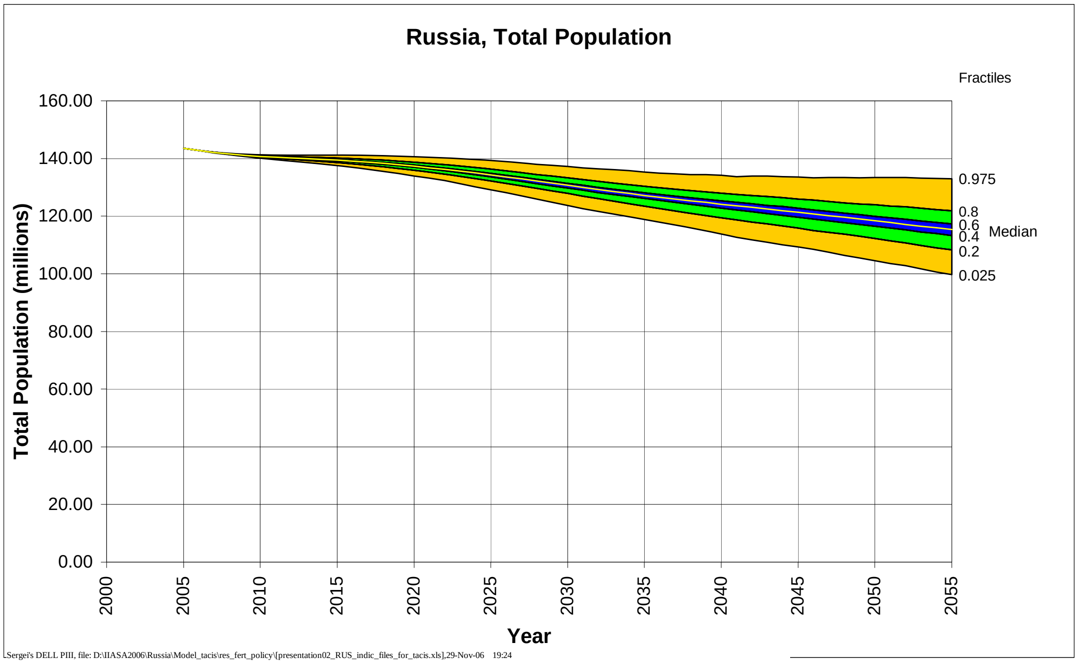
The resulting distributions of households by size were close in these approaches with the random headship rates approach having a slightly higher variance. Thus we will present the results of only those cases where we applied random headship rates with the Swedish rates used to set a target level for headship rates in 2050, up to which the rates increase linearly in 2005-2050.

---

[2] Since we are interested in the population living in private households, we adjusted the projected population with the proportion of people living in private households obtained from the 2002 census.

4 Population and Households: A General Overview of Prospects

First of all, let us look at the probabilistic population projection for Russia. Figure 1 presents the fractals of this distribution. As we observe from this figure, there is virtually no chance for population growth in the future. This is predefined by very low fertility and high mortality levels. Low fertility will also have implications for the total number of households. The total number of households is projected to decline in the long term (Figure 2). In the short term, the next 10-15 years, the median number of households is going to increase slightly, even though the fertility level is low. After that period, a steady decline is expected and by 2050 the number of households may fall to 47 million from 52.5 million in 2005. The 95 percent uncertainty range will spread from 41.5 to 52.4 million households in 2050.



Sergei's DELL PIII, file: D:\IIASA2006\Russia\Model_tacis\res_fert_policy\[presentation02_RUS_indic_files_for_tacis.xls],29-Nov-06 19:24

Figure 1. Probabilistic population projection for Russia.

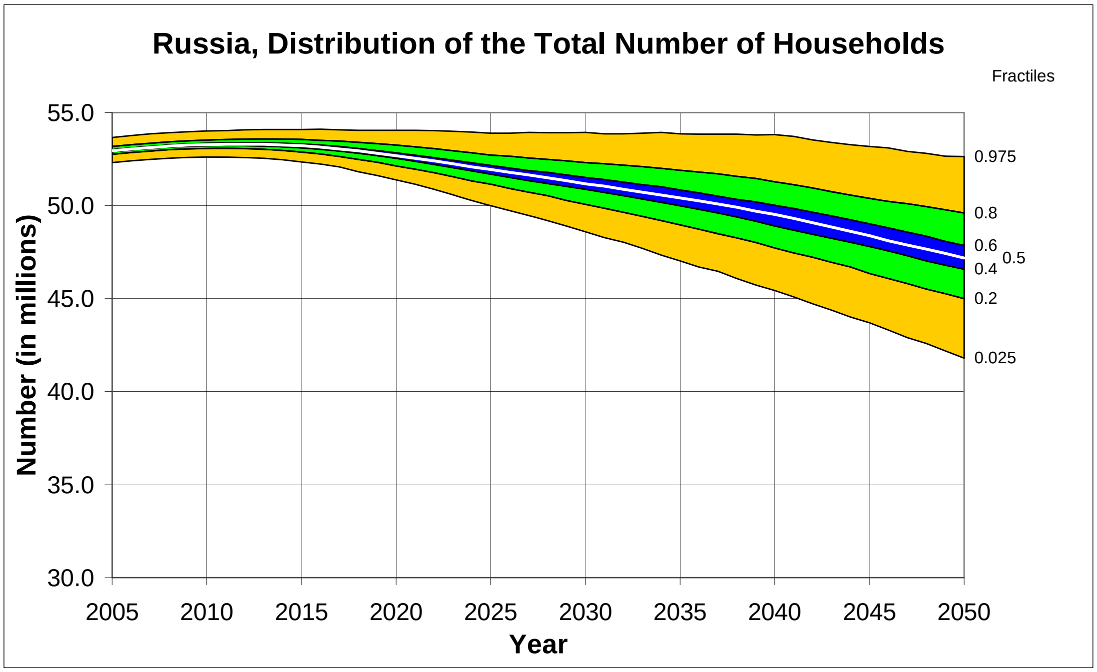


Figure 2. Distribution of the total number of households in Russia, projection for 2005-2050.

Not only is the total number of households projected to decline, but households will become smaller (Figure 3). Median household size falls from 2.69 in 2005 to 2.5 in 2035 and after that stays almost constant. In 2035, the 95 percent prediction interval includes households with sizes between 2.4 and 2.6 members. In general the decline in fertility leads to the decline of an average household size. This process is also partially due to the aging of the population structure as elderly tend to live in smaller households.

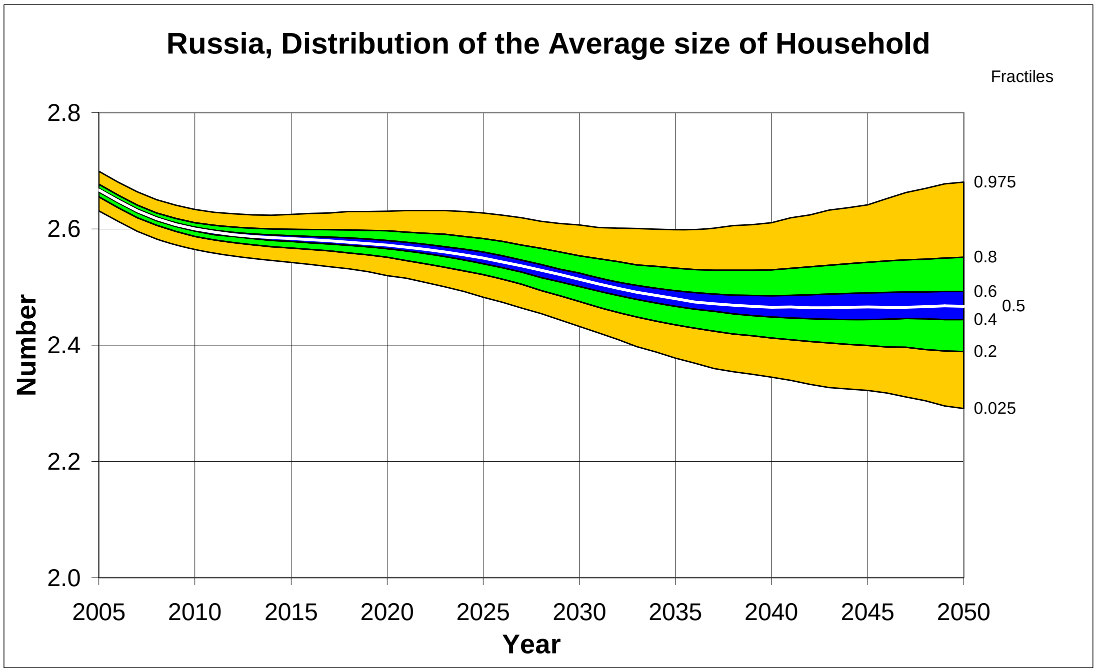


Figure 3. Distribution of the average household size in Russia, 2005-2050.

Even though the total number of households is expected to decline, we may expect diverse trends if we study the dynamics of households of different sizes (Figures 4-8). In the near future we may observe the rise in the number of households of size one from 11.5 million in 2005 to almost 13 million in 2035 (Figure 4). Households of size two and three show either no change or a very slight decrease in the near future with a moderate decrease by 2050 (Figures 5 and 6). The strongest decline will be observed in households with four and more members (Figures 7 and 8). The small variation of the number of households with four members at the base year is mere coincidence: The number of these households is at the maximum level, given the observed population size and average household size; therefore, this number is less sensitive to changes in model parameters. We may expect that households of size four will decline by 20 percent and of size 5 by 60 percent by 2050. Typically that would be households consisting of two parents and two or three children.

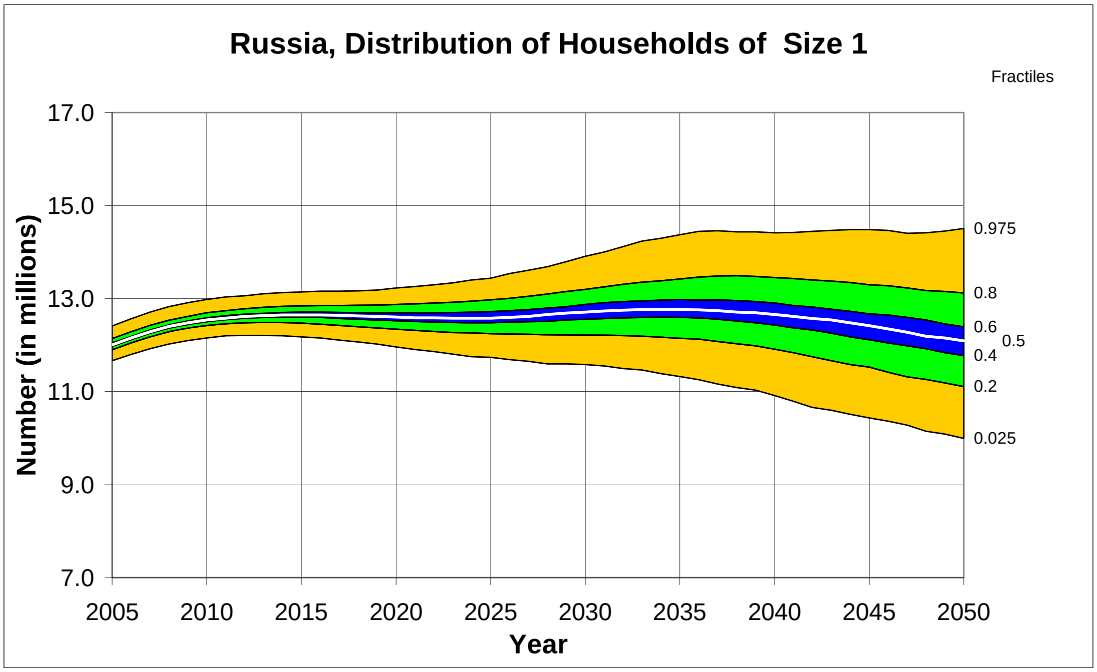


Figure 4. Distribution of size one households in Russia, 2005-2050.

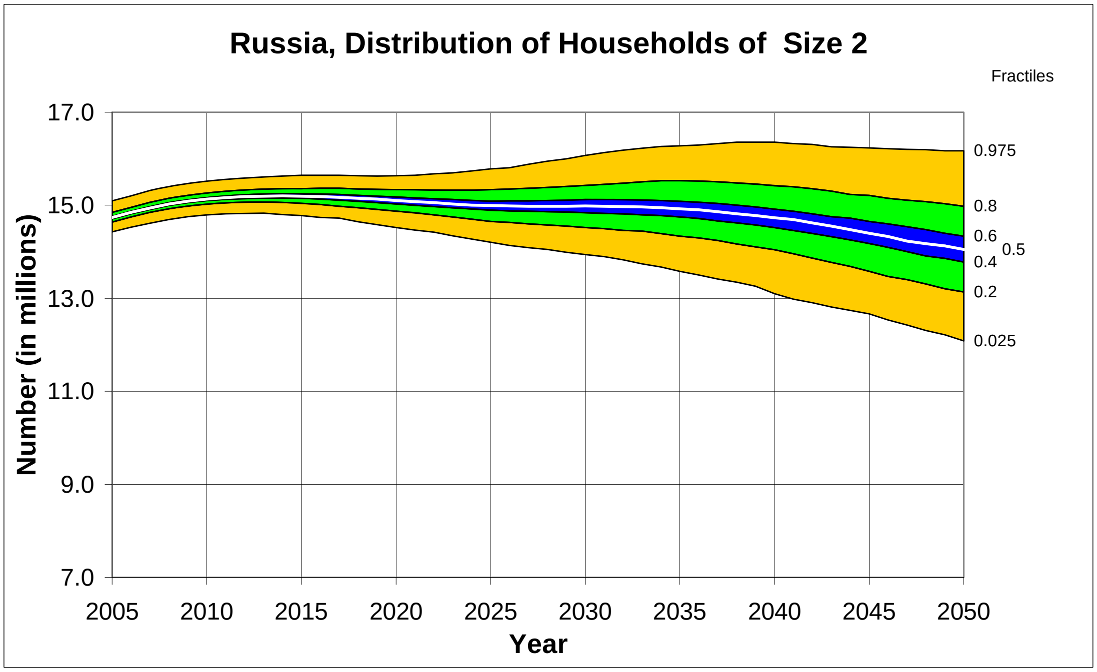


Figure 5. Distribution of size two households in Russia, 2005-2050.

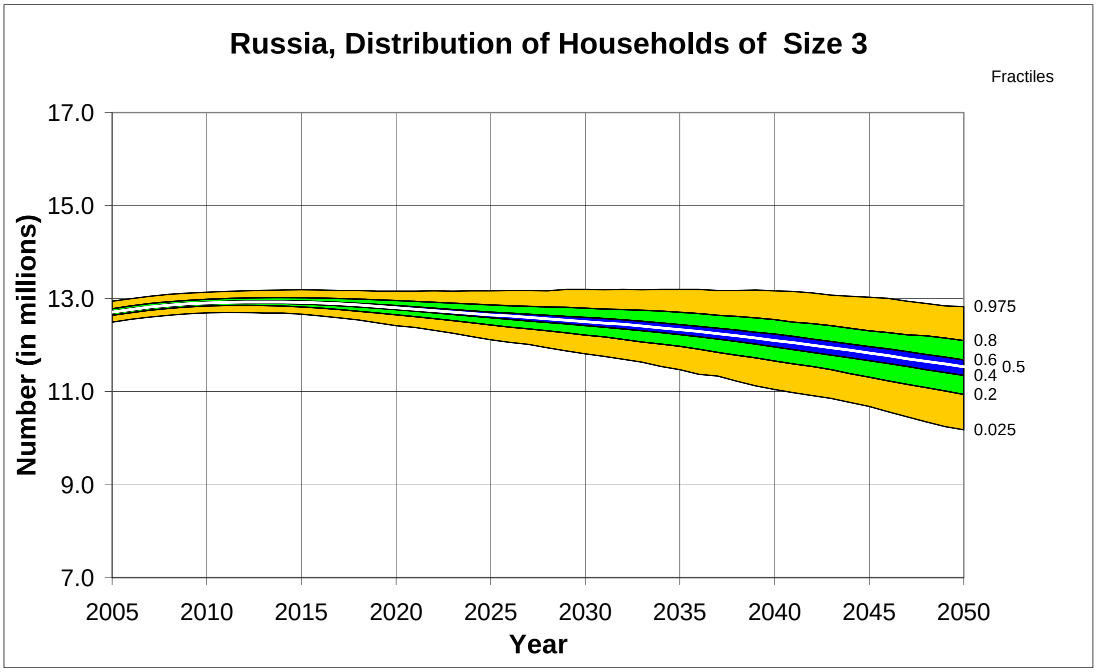


Figure 6. Distribution of size three households in Russia, 2005-2050.

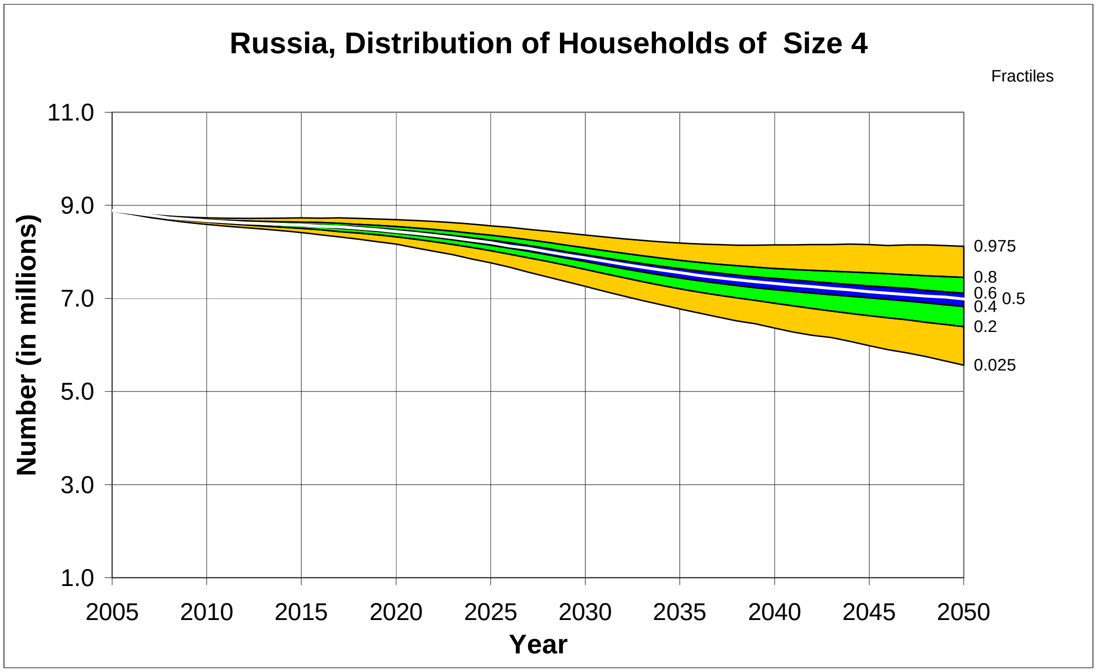


Figure 7. Distribution of size four households in Russia, 2005-2050.

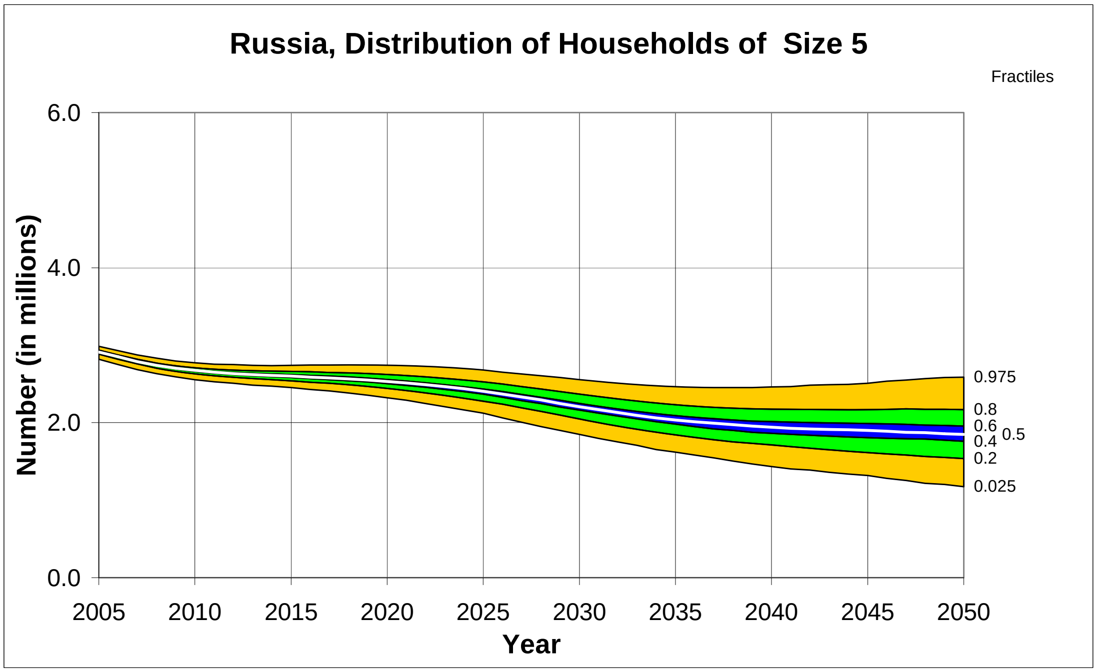


Figure 8. Distribution of size five households in Russia, 2005-2050.

Another way to look at the future distribution of households by size and to track the uncertainty associated with those distributions, is to present the distribution of households by size for a particular time point (Figures 9-10). From Figure 9 we may observe that there is virtually no chance that the number of households with four and more members will be higher in 2025 than it was in 2005. In 2050 a similar statement could be made regarding households with three and more members (Figure 10).

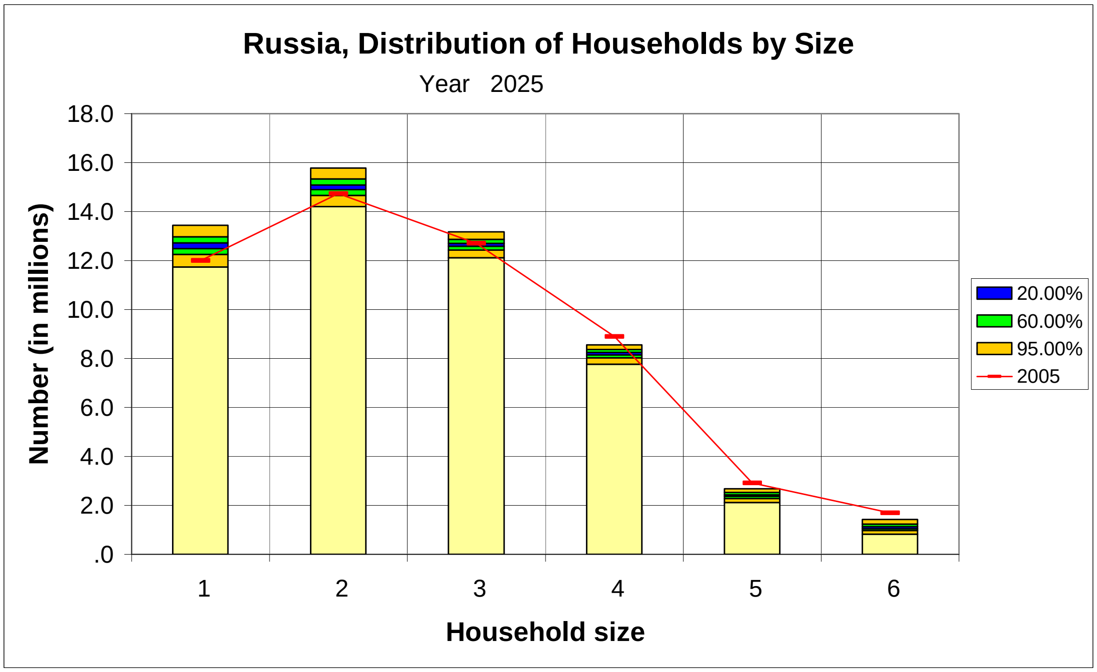


Figure 9. Distribution of household size in 2025 in Russia.

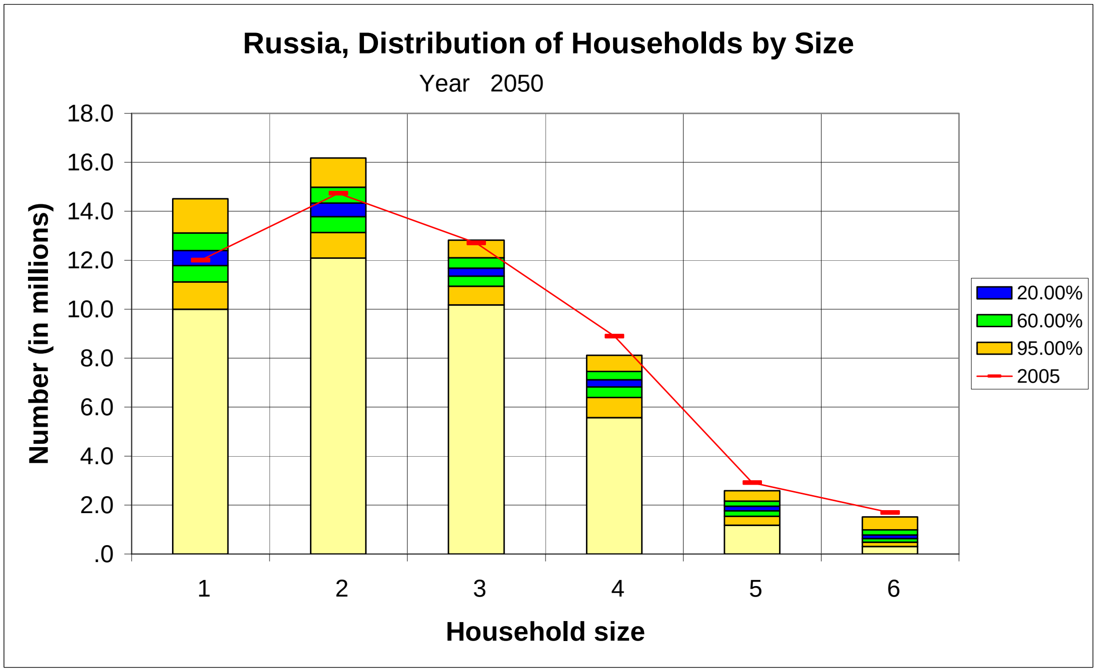


Figure 10. Distribution of household size in 2050 in Russia.

5 Conclusions

In this paper we presented the first probabilistic projections of the number of households in Russia. How can these projections be used? What type of questions are we able to answer with these results?

One of the extremely important issues in Russia today is the availability of housing. Many families and households live in apartments where several people share one room. However, social norms of housing per person exist, depending on the size of the household.[3] Using these norms and assuming that they stay constant in the future, it is relatively easy to calculate the probabilistic demand for housing in Russia. Figure 11 presents the results of these calculations. The dotted line markers designate the existing availability (in 2002) of housing in millions of square meters that is occupied by households of different sizes. The vertical bars represent the demand for housing by household size calculated using social norms standards.

As we see from Figure 11, households with one or two members occupy even more housing space than would correspond to social norms. There might be several reasons for that. First of all, the distribution of housing is extremely uneven. Two households of the same size may live in very different housing conditions. However one of the explanations of excessive available housing is that many of the households of this type consist of an elderly person living alone. Usually this person will have a bigger apartment, since at one time he/she was living together with a spouse and probably children. The children have since left home, the spouse has died, and the apartment or house (usually in a rural area) is occupied by only one person.

The most alarming situation is the availability of housing for households with four members. Typically that would be two parents living with two children. Since the demographic policy adopted in Russia today is aimed at a second child, housing facilities for households consisting of four members should be available. Our projection suggests that the number of households of size four is expected to decline. Despite the decline in numbers, however, these households will face a shortage of housing supply, if the situation with housing availability does not improve (Figure 12). Even if there might be enough housing existing in 2025 for households with five and more members, due to a very strong decline in the number of such households, the lack of housing for households with four members will almost certainly be a problem, if there is no considerable increase in the housing stock for households with four members.

---

[3] Decree of the President of the Russian Federation Nr. 425 of April 28, 1997 “On Housing and Utility Sector Reform in the Russian Federation”.

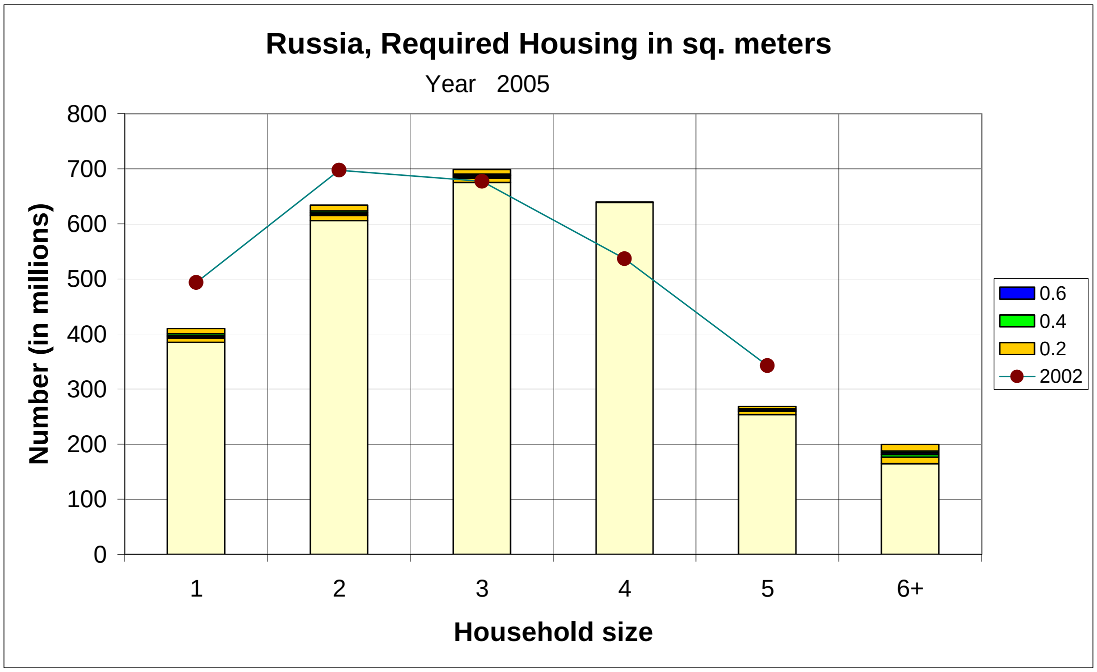


Figure 11. Supply and demand for housing in Russia in 2005.

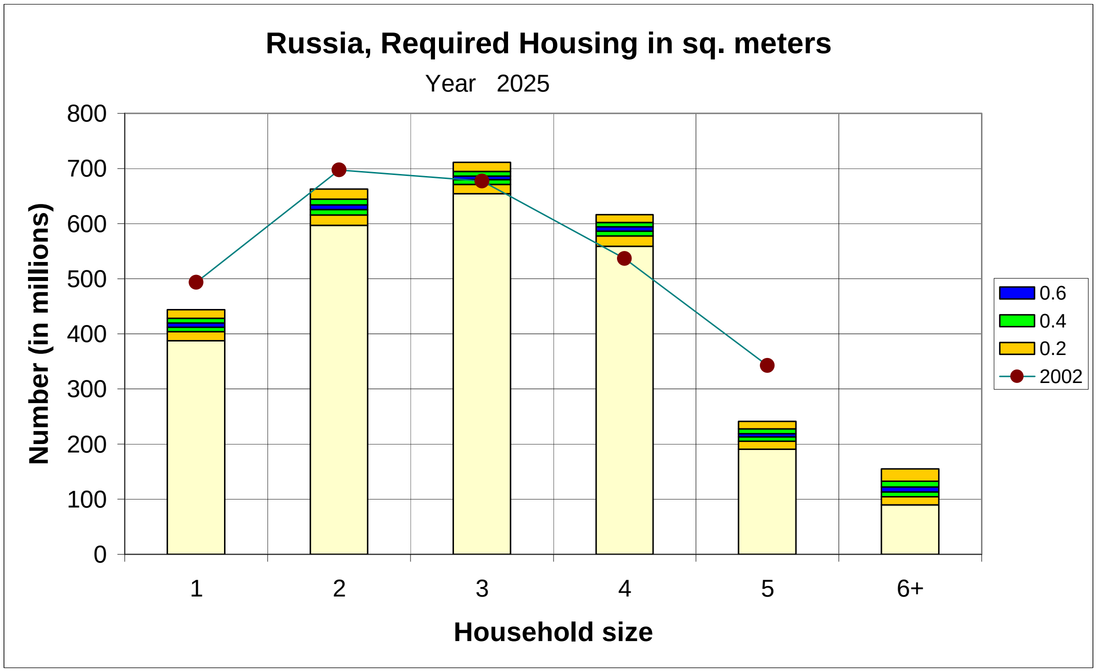


Figure 12. Supply and demand for housing in Russia in 2025.

6 References

Bongaarts J. and R.A. Bulatao, Eds. 2000. *Beyond Six Billion. Forecasting the World's Population*. Washington, D.C.: National Academy Press.

Booth, H. 2006. Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting* 22(3): 547-581.

De Beer, J. and M. Alders. 1999. Probabilistic Population and Household Forecasts for the Netherlands. Paper presented at the Joint ECE-Eurostat Work Session on Demographic Projections, Perugia, May 3-7, 1999.

Ediev, D.M. 2001. Application of the demographic potential concept to understanding the Russian population history and prospects: 1897-2100. *Demographic Research* 4(9): 289-336. http://www.demographic-research.org/volumes/vol4/9/

Ediev, D.M. 2007. On Projecting the Distribution of Private Households by Size. Working Paper 04/2007. Vienna Institute of Demography of the Austrian Academy of Sciences. http://www.oeaw.ac.at/vid/download/WP2007_04.pdf

Gisser, R. 1986a. Vorausschätzung der Haushalte nach Bundesländern bis 2015. *Statistische Nachrichten* 41(4): 279-288.

Gisser, R. 1986b. Vorausschätzung der Haushalte nach Grösse und Bundesländern bis 2015. *Statistische Nachrichten* 41(5): 378-383.

Keilman, N., D.Q. Pham, and A. Hetland. 2002. Why population forecasts should be probabilistic – illustrated by the case of Norway. *Demographic Research* 6: 409-453.

Leiwen, J. and B.C. O'Neill. 2004. Toward a new model for probabilistic household forecasts. *International Statistical Review* 72(1): 51-64.

Lutz, W., W. Sanderson, and S. Scherbov. 1996. Probabilistic population projections based on export opinion. Pages 397-428 in W. Lutz (ed.), *The Future Population of the World: What Can We Assume Today?* London: Earthscan.

Lutz, W., W. Sanderson, and S. Scherbov. 1997. Doubling of world population unlikely. *Nature* 387: 803-805.

Lutz, W. and S. Scherbov. 1998. An expert-based framework for probabilistic national projections: The example of Austria. *European Journal of Population* 14: 1-17.

Lutz, W., W.C. Sanderson, and S. Scherbov. 1999. Expert-based probabilistic population projections. Pages 139-155 in W. Lutz, J.W. Vaupel, and D.A. Ahlburg (eds.), *Frontiers of Population Forecasting.* A supplement to *Population and Development Review,* Vol. 24 (1998).

Lutz, W., W.C. Sanderson, and S. Scherbov. 2001. The end of world population growth. *Nature* 412: 543-545.

Lutz, W., W.C. Sanderson, and S. Scherbov. 2004. The end of world population growth. Pages 17-83 in W. Lutz, W.C. Sanderson, and S. Scherbov (eds.), *The End of World Population Growth in the 21st Century*. London: Earthscan.

MacKellar, L., W. Lutz, C. Prinz, and A. Goujon. 1995. Population, households and CO_2 emissions. *Population and Development Review* 21(4): 894-865.

O'Neill, B.C. and B.S. Chen. 2002. Demographic determinants of household energy use in the United States. Pages 53-88 in W. Lutz, A. Prskawetz, and W.C. Sanderson (eds.), *Population and Environment: Methods and Analysis*. A supplement to *Population and Development Review,* Vol. 28, 2002.

Perz, S.G. 2001. Household demographic factors as life cycle determinants of land use in the Amazon. *Population Research and Policy Review* 20: 159-186.

Prskawetz, A., L. Jiang, and B.C. O'Neill. 2004. Demographic composition and projections of car use in Austria. *Vienna Yearbook of Population Research* 2004: 274-326.

United States National Resources Panning Committee. 1938. *The Problems of a Changing Population.* Washington, D.C.: Government Printing Office.

United Nations. 1973. *Methods of Projecting Households and Families*. New York: United Nations.

United Nations. 1997. *1995 Demographic Yearbook*. New York: United Nations.

Appendix: Private Households by Size in a Stable Population

1. Headship rates

Let $P(x;t)$, $H(x;t)$, $M(x;t)=P(x;t)-H(x;t)$, and $h(x;t)=\frac{H(x;t)}{P(x;t)}$ be the population in private households, household heads, non-head members of households, and headship rates at age *x*.[4] We suppose the following simplified model determining the evolvement of these functions. First, the dynamics of the number of heads is determined by mortality and also by the formation of new households. Death of the head implies that all other household members move to other existing households, rather than forming a new household.[5] Second, the formation of new households occurs by separation from existing households and happens at some fixed age-specific rates $g(x)$. Third, we apply the same age-specific death rates to both heads and non-head members. These simplifying assumptions allow separating the two processes and lead to the following differential equation for the population of non-head members of households:

$$\frac{\partial}{\partial t}M(x;t)+\frac{\partial}{\partial x}M(x;t)=-\mu(x;t)M(x;t)-g(x)M(x;t). \tag{A1}$$

From Eq. (A1), which is written in terms of the non-head population only, it is possible to derive the following relation for that population:

$$M(x;t)=M(0;t-x)e^{-\int_0^x(\mu(y;t-x+y)+g(y))dy}=P(x;t)e^{-\int_0^x g(y)dy}, \tag{A2}$$

where we suppose that there are no heads of age zero (i.e., $P(0;t)=M(0;t)$) and use the following traditional relation for the dynamics of the size of birth cohort:

$$P(x;t)=P(0;t-x)e^{-\int_0^x \mu(y;t-x+y)dy}. \tag{A3}$$

The population of heads may be obtained as the difference between the population total and the non-head population:

$$H(x;t)=P(x;t)-M(x;t)=P(x;t)\left(1-e^{-\int_0^x g(y)dy}\right). \tag{A4}$$

This allows obtaining the headship rates:

---

[4] For the sake of simplicity and also to avoid uncertainty related to the sex of the household head, we do not address sex, although it may be added to the study.

[5] In fact, the emergence of new households due to the death of the head of an existing household may be reflected indirectly in the proposed model through the age-specific rates of changing status from "non-head" to "head".

$$h(x;t)=\frac{H(x;t)}{P(x;t)}=1-e^{-\int_0^x g(y)dy}. \qquad \text{(A5)}$$

Hence, headship rates are constant and do not directly depend on the reproduction regimen of the population as long as the age-specific rates of separating to new households are fixed. This result may be extended to the case of a varying rate of new household formation $g(x;t)$:

$$h(x;t)=1-e^{-\int_0^x g(y;t-x+y)dy}. \qquad \text{(A6)}$$

In this more general case, again, mortality and fertility are not directly involved in headship rates. In the model proposed, the reverse transitions from ‘head’ status to ‘non-head’ status were neglected. Hence, the solutions in Eqs. (A5) and (A6) – under non-negative rates of transition from non-head status to head status – are ever increasing by age. In real populations, there is a slight decrease in the headship rates for the oldest old ages, as the elderly may join the households of their kin rather than continuing to keep their own households. However, this decline in headship rates may also offer more options for stating the ‘household head’ in the census in households with several generations cohabiting together and also reflect the cohort effects on headship rates.

In any case, the headship rates seem to be much less sensitive to variations in reproduction regimes compared to, say, population size and age structure. This explains the remarkable stability of headship rates in human populations and also provides a rationale in support of the headship rates method. This point is also supported by empirical data (e.g., Leiwen and O’Neill 2004; Ediev 2007): Age-specific headship rates are remarkably stable, when no details concerning the household size or type are concerned.

2. Average household size

Since age-specific headship rates have become less sensitive to changes in the reproduction regimen, one may study the consequences of a stable population’s age structure for number, average size, and distribution of households by size assuming some fixed age profile of the headship rates. Let $h(x)$ be the headship rate at age x, which we assume to be fixed for all populations to be considered. The average household size, which – under the model proposed – determines their distribution by size, may be written as follows for a stable population:

$$n^s=\frac{\int_0^\omega Bl(x)e^{-\rho x}dx}{\int_0^\omega Bl(x)e^{-\rho x}h(x)dx}=\frac{\int_0^\omega l(x)e^{-\rho x}dx}{\int_0^\omega l(x)e^{-\rho x}h(x)dx}, \qquad \text{(A7)}$$

where B are births in the stable population, $l(x)$ is the survivorship function, ρ is the Malthusian parameter (or Lotka’s coefficient), and ω is the maximum lifespan.

Headship rates are nil for children and grow rapidly to the level of about 0.6 by the age of 25-30. Therefore, one may use the following approximate for headship rates in Eq. (A7) in order to simplify the relation:

$$h(x) \approx \begin{cases} 0, \ x \le x_{\min}, \\ h^*, \ x > x_{\min}. \end{cases} \tag{A8}$$

Substituting this into Eq. (A7), we have:

$$n^s \approx \frac{\int_0^{\omega} l(x) e^{-\rho x} dx}{h^* \int_{x\min}^{\omega} l(x) e^{-\rho x} dx} = \frac{\int_0^{x\min} l(x) e^{-\rho x} dx + \int_{x\min}^{\omega} l(x) e^{-\rho x} dx}{h^* \int_{x\min}^{\omega} l(x) e^{-\rho x} dx} =$$

$$= \frac{1}{h^*}\left(1 + \frac{\int_0^{x\min} l(x) e^{-\rho x} dx}{\int_{x\min}^{\omega} l(x) e^{-\rho x} dx}\right). \tag{A9}$$

This expression indicates that there is a lower limit for average household size in a stable population:

$$n^s \ge \frac{1}{h^*}. \tag{A10}$$

For the usual case of headship rates of about 0.6 at most adult ages, this implies that the *average household size in a stable population may not be lower than about* 1.67, which – given the models proposed for household distribution by size – has apparent implications for limiting the proportions of households of different sizes.

The expression on the right-hand side of Eq. (A9) depends on mortality and on the reproduction regimen of the population. To make these relations more explicit, let us use the following simplifying approximation. Let us consider, that the survivorship function may be approximated by a piece-wise constant function:

$$l(x) \approx \begin{cases} 1, \ x \le e_0, \\ 0, \ x > e_0, \end{cases} \tag{A11}$$

where e_0 is the life expectancy at birth. Using this approximation, one may get from Eq. (A9):

$$n^s \approx \frac{1}{h^*}\left(1 + \frac{\int_0^{x\min} e^{-\rho x} dx}{\int_{x\min}^{e0} e^{-\rho x} dx}\right) = \frac{1}{h^*}\left(1 + \frac{1 - e^{-\rho x_{\min}}}{e^{-\rho x_{\min}} - e^{-\rho e_0}}\right) = \frac{1}{h^*}\frac{e^{\rho e_0} - 1}{e^{\rho(e_0 - x_{\min})} - 1}. \tag{A12}$$

For stable populations with a reproduction level close to simple replacement, i.e., with Lotka's coefficient close to zero, life expectancy at birth is the main factor of variations in average household size:

$$n^s \approx \frac{1}{h^*}\frac{e_0}{e_0 - x_{\min}}, \tag{A13}$$

when $\rho e_0 << 1$.

For a wider range of stable populations, one may use Eq. (A12) to study the variations of average household size.[6] Figure A1 presents the results of the calculations using $x_{\min} = 25$, $h^* = 0.6$. The figure shows explicitly that the declines in average household size were due mainly to improvements in life expectancy and fertility decline. Both processes are tightly linked to the process of demographic transition. Hence, the demographic transition – apart from cultural changes and reassessments of family values – has caused the decline in average household size. Note, however, that during the first stages of the transition, when mortality decline results in improvements of Lotka's *r*, the average household size might be relatively stable or even growing. Later on, however, a decline in average household size must follow.

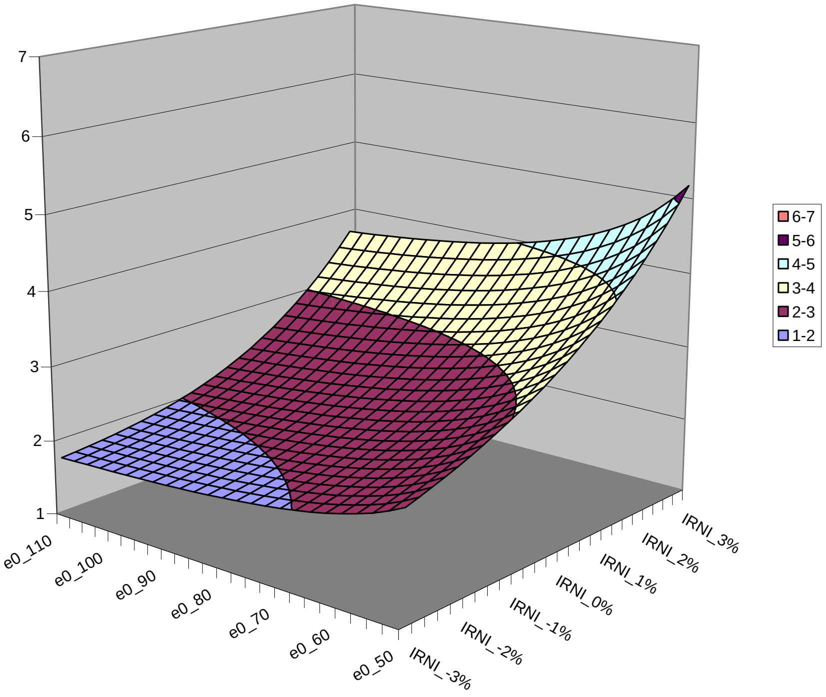


Figure A1. Approximates of average household size in a stable population as a function of life expectancy at birth and of Lotka's coefficient (IRNI – intrinsic rate of natural increase).

---

[6] Numerical simulations show that the approximation Eq. (A12) works well and provides results very close to those obtained directly from Eq. (A7).

3. Distribution by size

The distribution of households by size may be derived from their average size (Ediev 2007) and has been described in this paper. Figure A2 presents the results of estimating the proportion of households of different sizes in a stable population with varying fertility and with life expectancy fixed at a level of 80 years. Figure A3 presents the results for a stable population with varying mortality and with replacement fertility. Changes in the population age structure associated with a fertility decline have a negative effect on the proportion of households with four and more members, and a positive effect on proportion of one- and two-person households. The proportion of households with three persons, however, varies only moderately even within the remarkably wide range of fertility levels. A rise in life expectancy has nearly the same effect on household distribution by size. Hence, a simultaneous fall in fertility and rise in life expectancy, as was observed for many populations, enhance the effects of both processes on household dynamics. In particular, the mere change in population age structure during the demographic transition process seems to be the main factor of emergence of the modern distribution of households with a declining share of large households and a dramatic growth of the share of one-person households. This is illustrated in Figure A4, which presents the proportion of one-person households for a stable population with different combinations of fertility and life expectancy.

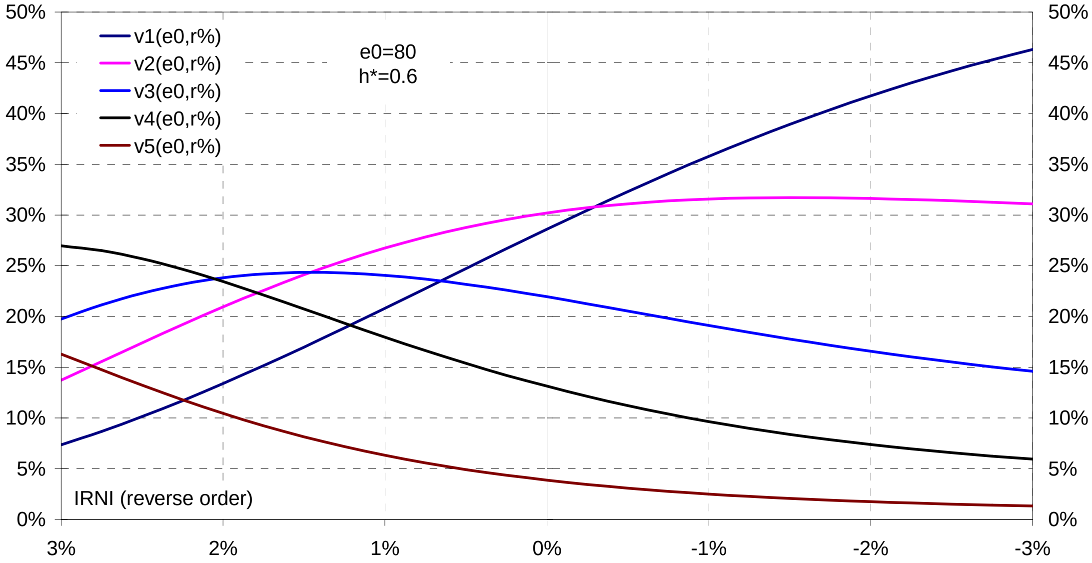


Figure A2. Proportion of private households of sizes one to five in a stable population as a function of Lotka's coefficient (IRNI) with a life expectancy at birth fixed at 80 years.

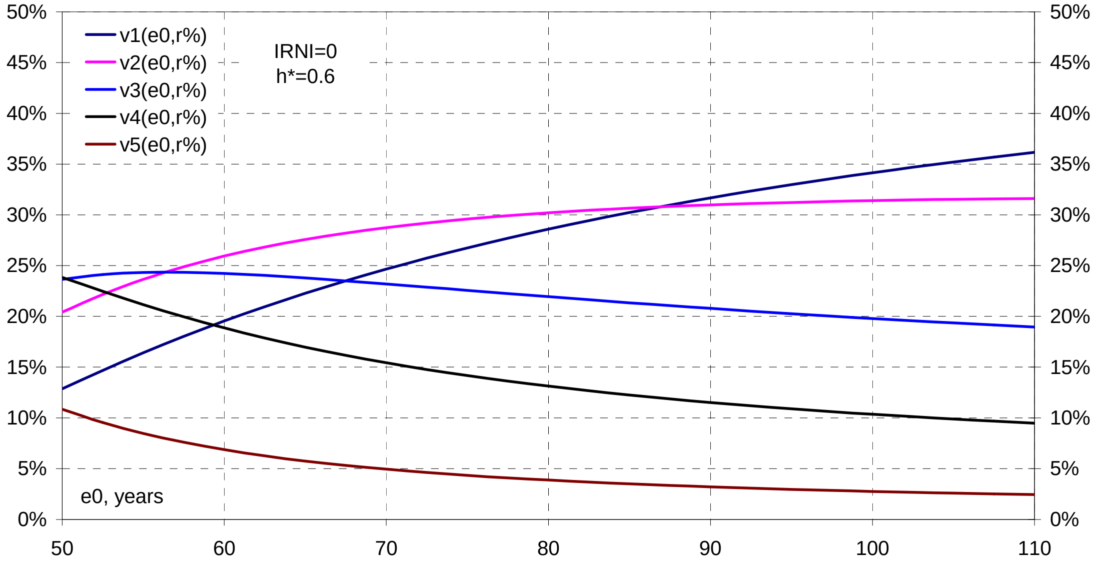


Figure A3. Proportion of private households of sizes one to five in a stable population as a function of life expectancy at birth under replacement fertility.

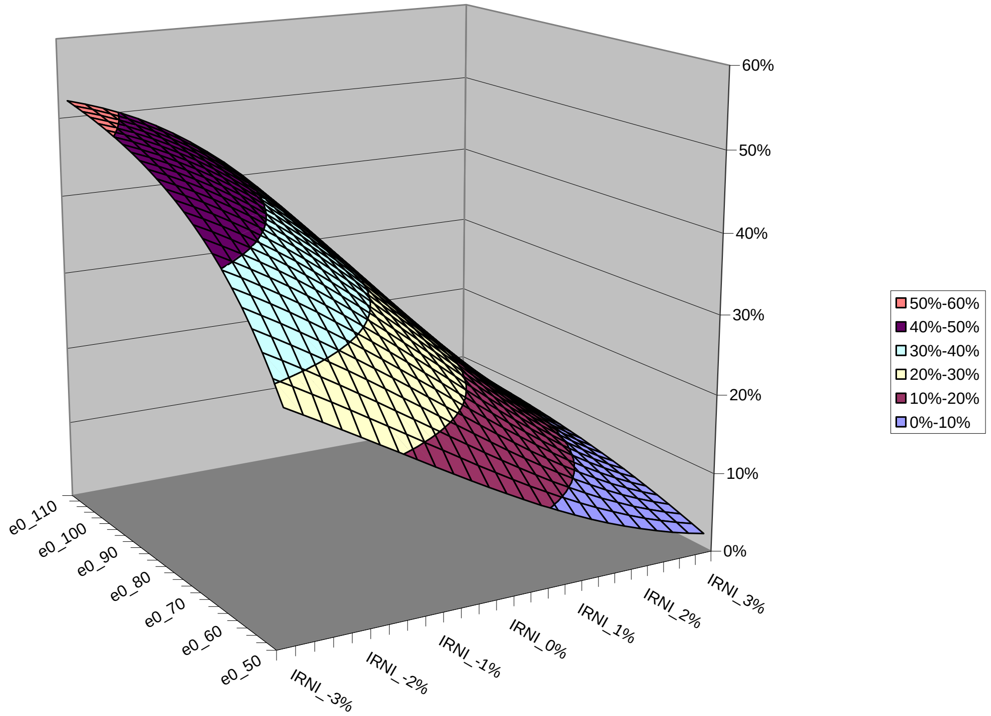


Figure A4. Proportion of single-person households in a stable population as a function of life expectancy at birth and of Lotka's coefficient (IRNI).