Dmitriev Mikhail

STRUCTURAL ANALYSIS OF FERTILITY IN RUSSIA

Academic advisor – Belianin A.V., PhD

1. Demographic situation in Russia

1.1. Introduction: Historical Trends

It is a commonly known fact that Russian population is steadily decreasing since 1992. Although it is generally accepted that population decline was caused by political and economic reasons of the 1990s, data show that Russia completed its second demographic transition earlier – in the 1950s and 1960s. Figure 1 shows the dynamics of *total cohort fertility* – the mean number of birth per women of a particular year of birth (cohort). As can be seen, even for women born in 1935 the average number of children was lower than 2.3, which is the level of simple reproduction of population.

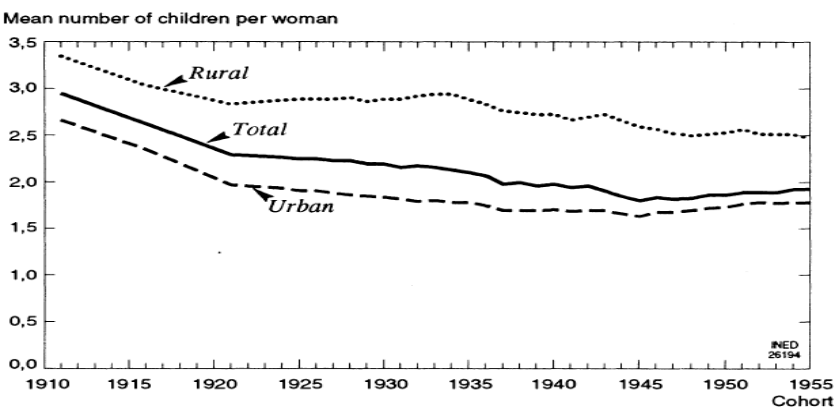

Figure 1. Mean Number of Children per Woman for each cohort from 1910-1950. Source: Avdeev and Monnier (1995).

This situation is threatening, and is barely getting better in the last few years. Is this process irreversible? What can be done to increase the rate of fertility in Russia? What kind of measures should be implemented? The present paper addresses these questions, and seeks an answer to them using econometric static and dynamic econometric techniques.

1.2. Russia in the context of international tendencies

In comparison with other developed countries, the low fertility rate of Russia is absolutely typical. Demographically, Russia is just an ordinary country whose transition to the new model of family took place a little later but had been significantly faster than in most developed countries. For instance, in France that period took about 120 years, for typical European countries – only 40-60 years (Vishnevsky, 2006). At the end of the 1990s Russia was supposed to complete its second demographic transition from high birth and death rates to low birth and death rates.

There was a rise in fertility rates in the 1980-s, thanks to simulative demographic policy in the late USSR (higher state subsidies, extended maternity leaves etc). However, those measures had had only short run effect: families just moved planned births to earlier ages, but did not increase their family sizes over the entire fertility period. Later, this forward move of planned births did contribute to sharp decrease of the birth rates in the 1990-s.

1.3. Demographic Situation in the 1990s and empiric Research on Russian fertility

Micro-data in the analysis of Russian fertility are used not more than 15 years. Basic database for such research is Russian Longitudinal Monitoring Survey (Kohler and Kohler, 2001; Grogan, 2003; Roshina and Boykov, 2005; Roshina, 2006). Other researches used data of population micro-cencus (Kharkova and Andreev, 2000) or data of international project “Parents and children, men and women in family and society” (Zaharov e.a., 2007).

The major puzzle that researchers were interested in was a sharp decrease in the birth rate from the beginning of the 1990s. There are several explanations to that puzzle.

Demographic explanation: Decrease of fertility rate is a long-run tendency and economic crisis just accelerated tendencies that anyway would have taken place, albeit they would have been a little slower in its absence (Kharkova and Andreev, 2000).

Adaptation explanation: Philipov and Kohler (2001) show that the Russian model of reaction to economic and social factors is converging to that of the Europeans. With the transition to the market economy typical families are organized later than in the Soviet Union. These authors forecasts that fertility rate would decrease in the short run and than in the long-run return to current values.

Deferring explanation: Heleniak (1995) and Rimashevskaya (1997) blame economic crisis and destruction of social and economic support of the big families as a main reason of the decrease in the birth rate. Because of these difficulties, births are delayed till the end of crisis.

Reproductive behavior in contemporary Russia has several peculiarities, including:

- Importance of social norms (e.g. Roshina, 2006): it is generally accepted in Russia to have at least one child and at most two.
- Similarity of rural and urban fertile behavior, although fertility rate is still slightly higher in the countryside (Avdeev and Monnier, 1995)
- Imporance of housing conditions (Maleva and Sinyavskaya, 2006).
- Absence of correlation between income of woman and number of her children. A notable exception is Buhler (2004), who found negative correlation between households income and probability of birth of the first child. By contrast, additional income (temporary, or from second or third job) positively influences fertility.

This paper presents results of both structural and reduced-form analysis of fertility factors using RLMS data. Stochastic dynamic programming approach can explain the nature of birth-planning and role of uncertainty in that process, whereas reduced-form model can control for a much larger set of potential determinants. Hence the two approaches complement, rather than contradict each other. Moreover, we also use structural dynamic discrete-choice modeling with permanent heterogeneity, which is probably the first implementation of this method when dealing with the Russian data.

2. Data

2.1. Data Description

For estimation of fertile behavior we used the RLMS (Russian Longitudinal Monitoring Survey) database. It includes data for 10 years from 1994 to 2005 (from 5th to 14th rounds). The database consists of 20 files (2 files for each year), which contain individual and household data, respectively. Individual data contain very detailed information about each person: age, nationality, education, job, health, smoking and alcohol preferences… Questions relevant for our purposes contain information about women's reproductive behavior – e.g. whether the woman would like to have a new child. Household file contains information about housing conditions, short information about each member of households and relationships between different members. For further information about data the reader is referred to the full text of the thesis.

Unfortunately, the list of questions changed from year to year. Moreover, the subset of questions concerning reproductive behavior has been shortened sharply from 1999 in the individual questionnaires. From 2004 the individual file even seased to contain information about the number of children born by a particular woman. However, this information can still be obtained indirectly from household files.

The sample was restricted to women from 16 to 39 years for the reduced form models and from 16 to 35 years in the structural form models. A specific contribution of the paper is the method of construction of our dependent variable, (*birth*), which was set to one in year *T* if the baby had been conceived between surveys at *T* and *T*+1, otherwise it is equal to zero. This definition is more elaborated than the standard one, as its derivation requires combination of information from household file and different individuals. However, we believe our method is better than the standard definition of dependent variable *birth* in period *T*, which is set equal to one if number of children recorded in the survey at year *T*+1 exceeds corresponding value at *T*. The reason is that this method may be misleading: suppose that three surveys were conducted in November of 2004, 2005 and 2006. If a particular woman decided to have a baby on March 2005 and conceived it at the same month, then she would give a birth at January 2006 and it would be recorded on November 2006. In this way, variable *birth* will be set equal to one for 2005, although other variables this year could have already been affected by the fact of woman's pregnancy (she is likely to smoke and drink less, could take a leave at her job etc.), whereas other variables which might have affected that decision in 2004 (e.g. increase of income or living space) are disregarded.

The list of independent variables included

- Woman characteristic: age, education, number of children, employment status, income, being on maternity leave, alcohol and smoking preferences, life satisfaction, self-estimated health status.
- Characteristics of her husband (only for married women): age, employment status, income, alcohol and smoking preferences, life satisfaction, self-estimated health status.
- Characteristics of the household: size, status of the settlement, average living space per one member of the household, number of facilities in the apartment, other revenues of the household.

All incomes are deflated to the level of November 1994. In the reduced form models logarithms of real incomes were used; in structural form models we used real (deflated) data.

3.Reduced Form Models

3.1. Methodology

It is supposed that the woman can get different utilities from different decisions. She compares two situations: when she has additional child in the family and when her family size stays the same. The fact that at a given period the woman doesn't give birth implies that utility from additional child is lower than in the current state of the family. Utility depends on observable and unobservable factors. Using econometric method it is possible to estimate the parameters of utility functions that best fit the observable decision patterns.

Observable factors include age (five categories); quantity of already born children and square of that number; employment status (three categories); education (four categories); drinking (five categories); health (three categories); real income; additional income of the household; dummy variables for equality of woman's real income and additional income to zero; average living space per person in the household (five categories); number of other members of the household (excluding woman herself, her husband and her children); number of additional desirable children and its square; settlement status (four categories), federal region (7 categories); year dummy variables. Models are estimated on three samples: total sample, subsamples of unmarried and married.

Separate care was taken to fix measurement errors. If total income of the household was lower than the sum of woman and her husband's income, its value was replaced by this sum. Observations with the highest 0.5% income were excluded from the sample.

Explanatory variables can be divided into three broad groups: (1) variables affecting costs of upbringing a child, (2) variables that characterize financial and timing restrictions of the household, (3) factors that detail preferences towards quantity and quality of children. Some of variables can be referred to several groups. Dummy-variable of being married influence the costs of upbringing, as some husbands help their wives in dealing with pregnancies and childcare. It also affects financial opportunities of the household (e.g. non-labor disposable income of the married women), preference for the number of children (through the preferences of the husband) etc. Education also can be referred to all groups through higher opportunity costs of upbringing, higher financial opportunities of educated people, and non-traditional preferences and demand for higher quality of children instead of their quantity (last group of factors). In the first group number of other members of the household, personal income of an employed woman can be listed, alongside with the desired number of chilcren. Husband income and additional income of the household, living space measures belong to the second group.

We have estimated the panel probit random effect model, which assumes that unobservable preferences equal to the sum of the individual effect and random shock $\varepsilon_i + e_{it}$. Individual effect is permanent across periods for a given individual. It is supposed to be normally distributed across individuals. Random shock e_{it} is also normally distributed random variables independent across periods and individuals and also uncorrelated with individual effect. Then probability of a given individual to have a child equal

$$\Pr(birth = 1 \mid X_i) = \Pr(U_{birth=1} = \beta_0 + \sum \beta_{ik} X_{ik} + \varepsilon_i + e_{ik} > 0) = \\ = F\left(\frac{\beta_0 + \sum \beta_{ik} X_{ik} + \varepsilon_i}{\sigma_e}\right), \quad (1)$$

where *X* is the matrix of covariates and *a* is the choice index. All coefficients are determined using MLE estimation given by

$$\sum_{i=1}^{N} \ln \int_{-\infty}^{\infty} \prod_{t=1}^{n_i} \left(F\left(\frac{\beta_0 + \sum \beta_{ik} X_{ik} + \varepsilon_i}{\sigma_e}\right)\right)^{a_{it}} \times \\ \times \left(1 - F\left(\frac{\beta_0 + \sum \beta_{ik} X_{ik} + \varepsilon_i}{\sigma_e}\right)\right)^{1-a_{it}} \frac{1}{\sigma_\varepsilon} f\left(\frac{\varepsilon_i}{\sigma_\varepsilon}\right) d\varepsilon_i \to \max_{\beta, \sigma_e, \sigma_\varepsilon} \quad (2)$$

Coefficients of the ordinary pooled probit model were also estimated.

3.2. Estimation Results

Estimation results of the pooled probit model and probit random effect model are presented in tables 4 and 5 of the Appendix. In the lowest row of the table 5 there are tests that check whether the contribution of individual effect in total variation is significant from zero. The null hypothesis about equality of individual effect to zero cannot be rejected at the 5% level only for the sub-sample of married women. Signs and coefficients' significance is the same for pooled probit model and probit random effect model even for married women.

Age, marriage, employment, number of children and number of additional desirable children, number of additional members of family, settlement status were significant determinants of fertility, whereas use of contraception, living space, health, education, income, and year dummies were not.

Negative influence of age can be seen only for 30-34 and 35-39 age groups. By contrast, for the baseline group (from 16 to 19, omitted) it is essentially the same as for women up to 29 years. Country and small towns have higher rate of births than big cities.

Employed women (omitted category) have higher probability of giving a birth than not employed. Positive significance of employment can explain why education is insignificant. Employed women usually are more educated and have higher income, therefore they all may correlate, what decreases probability of their significance. Another explanation lies in social rules: employed women can get additional financial opportunities on their maternity leaves, whereas working experience may increase one's ability to place a child in a good kindergarten.

Probability of having additional child non-linearly depends on the number of those already born. Second child is less likely to be born than the first, the third one is less likely to be born than the second (although difference in probabilities is smaller), but the forth and next one have higher probabilities to be born than the previous ones. The number of desirable children is positively significant what is natural. At the same time, this effect should not be taken too literally, as this variable could be correlated with other regressors.

General patterns of fertilities are similar between sub-samples. However, in the sub-sample of unmarried women a decrease of birth rates begins later than for married women, whose fertility decreased from 20-24 years. Thus, early marriage is likely to lead to the birth of the first child, while among unmarried women the relationship among age, bith and number of children is unclear, but 70 % of unmarried women don't have children and only 8% have two or more children. Significance of marriage may be explained by the fact that women with higher demand for children prefer being married; but the characteristic of their husbands are not significant.

Contraception is negatively significant only for married women. Consumption of alcohol positively affects the probability of having a baby only for unmarried women, which indirectly explains non-planned character of births for that group. Economic factors are significant for married women at the 10% level. The same is true of average living space. For unmarried women these variables are not significant.

4. Dynamic discrete choice structural models

4.1. Basic algorithms

In dynamic discrete choice structural models, agents are forward-looking and maximize expected intertemporal payoffs. The parameters to be estimated are structural in the sense that they describe preferences of agents and beliefs about technological and institutional constraints.

Dynamic discrete choice models are, in a sense, ubiquitous: they can be applied to almost any problem of applied microeconomics. Decision to retire, to accept or reject new job, to give a birth etc., – they all depend not only of current but also of future benefits expected by the decision-maker. In these models the individual is maximizing expected sum of his or her discounted utilities:

$$\max_{\{a_t\}_{t=0..T}} E_0\left[\sum_{t=0}^{T}\beta^t(a_t u_{1,t}+(1-a_t)u_{0,t})\right] \quad (3)$$

where a_t is the a binary choice indicator made at *t* and $\{a\}_{t=0..T}$ is the sequence of such choices over the life time.

The basic papers of dynamic discrete choice structural models include Wolpin (1984) on mortality and fertility, Keane and Wolpin (1994) on occupational choice, Rust (1987) on optimal bus engine replacements, Pakes (1986) on patent renewal, Miller (1984) on job matching, and many others.

4.1.1. Rust (1987) Engine Replacement Model

In a widely cited paper, Rust (1987) has estimated the time of bus engine replacement in Madison, Wisconsin. The model is based on the following assumptions (Aguirregabiria and Mira, 2007, p.6):

ASSUMPTION AS (Additive separability): The one-period utility function is additively separable in the observable and unobservable components:

$$U(a, x_{it}, \varepsilon_{it}) = u(a, x_{it}) + \varepsilon_{it}(a) \qquad (4)$$

where $u(a, x_{it})$ is the utility function dependent on the choice indicator *a* and a vector of time- and agent-specific correlates x_{it}, and $\varepsilon_{it}(a)$ is a zero mean random variable with the real line support.

ASSUMPTION IID (iid unobservables): The unobserved state variables in ε_{it} are independently and identically distributed over agents and over time with CDF $G_\varepsilon(\varepsilon_{it})$ which has finite first moments and is continuous and twice differentiable in ε_{it}.

ASSUMPTION CI-X (Conditional independence of future *x*): Conditional on the current values of the decision variable and the observable correlates, the next period's correlates do not depend on the values of the current ε's: $CDF(x_{i,t+1} \mid a_{it}, x_{it}, \varepsilon_{it}) = F_x(x_{i,t+1} \mid a_{it}, x_{it})$. We use θ_f to represent the vector of parameters that describe the transition probability function F_x.

ASSUMPTION CI-Y (Conditional independence of *y*): Conditional on the values of the decision and the observable state variables, the value of the payoff variable y is independent of ε: i.e., $Y(a_{it}, x_{it}, \varepsilon_{it}) = Y(a_{it}, x_{it})$. The vector of parameters that describe Y is θ_Y.

ASSUMPTION CLOGIT: The unobserved state variables { ε_{it} (*a*): *a* = 0, 1, ..., *J*} are independent across choice alternatives and have an extreme value type 1 distribution.

ASSUMPTION DIS (Discrete support of x): The support of x_{it} is discrete and finite: $x_{it} \in X = \{x(1), x(2), ..., x(|X|)\}$ with $|X|<\infty$.

Subject to these and a few more technical assumptions, the future value function can be expressed as a function of observable variables only:

$$\overline{V}(x_{it}) = \int \max_{a\in A}\{u(a,x_{it})+\varepsilon_{it}(a)+\beta \sum_{x_{i,t+1}} V(x_{i,t+1}|a,x_{i,t})f_x(x_{i,t+1}|a',x_{it})\}dG_\varepsilon(\varepsilon_{it}) \quad (5)$$

Then the probability of decision a is

$$P(a \mid x,\theta) = \int I\{u(a,x_{it})+\varepsilon_{it}(a)+\beta \sum_{x_{i,t+1}} \overline{V}(x_{i,t+1}|a,x_{i,t})f_x(x_{i,t+1}|a,x_{it})>$$
$$> u(a',x_{it}) + \varepsilon_{it}(a') + \beta \sum_{x_{i,t+1}} \overline{V}(x_{i,t+1}|a',x_{i,t})f_x(x_{i,t+1} \mid a',x_{it})\}dG_\varepsilon(\varepsilon_{it}) \quad (6)$$

Moreover, assuming that ε have extreme value type 1 distribution,

$$\overline{V}(x_{it}) = \log\left\{\sum_{a=0}^{J} \exp\left[u(a,x_{it}) + \beta\sum_{x_{i,t+1}} V(x_{i,t+1}| a, x_{i,t})f_x(x_{i,t+1}| a, x_{it})\right]\right\} \quad (7)$$

and correspondingly, the probability of each outcome equals

$$P(a \mid x_{it},\theta) = \frac{\exp\{v(a,x_{it})\}}{\sum_{j=0}^{J} \exp\{v(j,x_{it})\}}, \quad (8)$$

where

$$v(a,x_{it}) = u(a,x_{it}) + \beta \sum_{x_{i,t+1}} \overline{V}(x_{i,t+1}|a,x_{i,t})f_x(x_{i,t+1} \mid a,x_{it}) \quad (9)$$

Although Rust's model was the most popular among economists because of its simplicity, it is prone to some criticism – in particular, it cannot deal with unobserved heterogeneity. Further, although the iid assumption fits well for bus engines is likely to oversimplify the world of humans.

4.1.2. Backward induction algorithm

The following version of the backward induction algorithm has been suggested by Keane and Wolpin (1994) for finite number of periods T. The value function is

$$V(S(t),t) = \max_{\left\{\{a_k\}_{\tau=t..T}\right\}_{k\in K}} E\left[\sum_{t=\tau}^{T} \beta^t (\sum_{k\in K} d_{kt}u_{kt})|S(t)\right] \quad (10)$$

where a_{kt} =1 if alternative *k* is chosen at time *t*, otherwise 0; u_{kt} is a reward from the chosen strategy; $S(t)$ is the state space that consists of all factors (known to the individual but not the researcher) that affect current rewards or the probability distribution of any of the future rewards.

The value function can be written as

$$V(S(t),t) = \max_{k\in K}(V_k(S(t),t)) \quad (11)$$

and, respectively, $V_k(S(t),t)$ is the alternative-specific expected lifetime reward or value function, expressed according to the Bellman equation (Bellman, 1957)

$$V_k(S(t),t) = U_k(S(t),t) + \beta E(V(S(t+1),t+1) \mid t, d_k(t) = 1) \quad (12)$$

Here $E(V(S(t+1),t+1) \mid t, d_k(t) = 1)$ is the so-called Emax function which is of our major interest. Starting from the Emax function for the last period:

$$E(V(S(T),T) \mid T-1, a_k(T-1) = 1) = \\ = \int_{\varepsilon_{kT}} ... \int_{\varepsilon_{1T}} \max(U_1(T)..U_k(T)) f(\varepsilon_{1T}..\varepsilon_{kT}) d\varepsilon_{1T}..d\varepsilon_{kT} \quad (13)$$

and using backward induction with formulas (11), (12) and (13), we can find the value of Emax function for any point in the state space.

In practice, calculation of Emax may be problematic due to the limitation of computational capacities. Keane and Wolpin suggested solving these problems through simulation and interpolation. First, they compute integrals using simulation (later, Bingley and Lanot (2004) found that for a given distribution of ε these integrals can be expressed analytically) and interpolation. They calculate Emax for a subset of sample space and then regress that function on the set of given factors, and predicted values from this regression to the remaining points of the state space.

The most important advantage of Keane and Wolpin method is its ability to deal with a permanent unobserved heterogeneity. Keane and Wolpin (1994) show that thei method is applicable for a finite number of agents' types and for any type of dynamic programming problem using weighted maximum likelihood.

4.2. Structural fertility models

Applying the standard dynamic model of fertility to our data, we assume that individuals maximize their expected utility over life time at each period from τ to T.

$$\max E_{\tau} \sum_{t=\tau}^{T} \beta^{t} U(N_t, M_t, X_t, H_t, u_t, \theta) \quad (14)$$

As before, β is a discounting factor, U is current utility, M_t is the number of children, X_t is consumption, H_t is leisure time, θ is an individual-specific random term, N_t is a dummy variable equal to one if the birth was given. Utility is maximized under constraint

$$I_t = X_t + P_t M_t + cN_t \quad (15)$$

Where, in addition to the parameters defined above, I_t is income of the household, P_t is consumption per one child, and c are costs of birth. The function (14) is maximized in each period by choosing X_t and H_t.

Heckman and Willis (1976):
$U = W(\psi N_t, X_t) - f(u_t)$

Wolpin (1984):
$U = W(M_t, X_t, \theta)$
$= (\alpha_1 + \theta) M_t - \alpha_2 M_t^2 + \beta_1 X_t - \beta_2 X_t^2 + \gamma M_t X_t$; γ any sign

Hotz and Miller (1984):
$U = W(M_t, Z_t)$; $Z_t = Z(H_t, X_t, \zeta_t)$
where Z_t is household production
ζ_t is a random error

Rosenzweig and Schultz (1985):
$U = W(N_t, M_t, X_t, H_t, \theta)$
$= \phi_1 N_t - \phi_2 N_t^2 + \alpha_1 M_t - \alpha_2 M_t^2 + \beta_1(\theta) X_t - \beta_2 X_t^2 + \delta_1 H_t - \delta_2 H_t^2 + \gamma H_t M_t$; γ any sign

Newman (1988):
$U = W(M_t, X_t, u_t)$
$= \alpha_1 M_t - \alpha_2 M_t^2 + \beta_1 X_t - \beta_2 X_t^2 + \gamma M_t X_t + \rho_1 u_t - \rho_2 u_t^2$; γ any sign

Leung (1991):
$U = W(M_t, X_t) - f(\pi(u_t))$
where $\pi(.)$ is the probability of a birth

Table 1. Different specifications of te utility functions. Source: Arroyo and Zhang (1997).

Several variants of this model are listed in Table 1; the specification we use is presented in the next subsection.

4.3. Model

Current utility function is supposed to depend on woman consumption and number of children:

$$U_t = f(M_t, X_t) + eN_t \tag{16}$$

To be concrete

$$U_t = (\alpha_1 + w)M_t + \alpha_2 M_t^2 + \alpha_3 M_t^3 + \beta_1 X_t + \beta_2 X_t^2 + eN_t \tag{17}$$

Where M_t is the number of existing children, X_t is consumption, N_t is a dummy equal to 1 if woman gave a birth, e is a random shock independent across individuals and periods, and w is an index of being fond of children, which is a random variable across individuals and constant across periods. Consumption is defined by

$$X_t = I_t - P_t M_t - cN_t \tag{18}$$

and the family size evolves according to

$$M_{t+1} = M_t + N_t \tag{19}$$

The individual solves

$$\max_{\{N_t\},t=1..T} E\{\sum \delta^t [(\alpha_1 + w)M_t + \alpha_2 M_t^2 + \alpha_3 M_t^3 + \\ + \beta_1(I_t - P_t M_t - cN_t) + \beta_2(I_t - P_t M_t - cN_t)^2 + eN_t]\} \tag{20}$$

Or using the Bellman principle,

$$\max_{N_t}(-\beta_1 c + \beta_2 c^2 - 2\beta_2(I_t - P_t M_t))N_t + \delta(E_t V(M_t + N_t, S_{it})) \tag{21}$$

Taking into account finiteness of time horizon backward induction algorithm is used. For the last period

$$V(T,M) = \int_{-\infty}^{\infty} f(e)[\max_{N_T}\{(\alpha_1 + w)M_T + \alpha_2 M_T^2 + \alpha_3 M_T^3 + \beta_1(I_T - PM_T - cN_T) + \\ + \beta_2(I_T - PM_T - cN_T)^2 + \gamma M_T(I_T - PM_T - cN_T) + eN_T]de \tag{22}$$

For earlier periods:

$$V(t,M) = \int_{-\infty}^{\infty} f(e)[\max_{N_t}\{(\alpha_1 + w)M_t + \alpha_2 M_t^2 + \alpha_3 M_T^3 + \beta_1(I_t - PM_t - cN_t) + \\ + \beta_2(I_t - PM_t - cN_t)^2 + \gamma M_t(I_t - PM_t - cN_t) + eN_t + \delta V(M_{t+1}, t+1)]de \tag{23}$$

That integrals are calculated using Monte-Carlo simulation. In such a way the matrix of value function for all space is created. This is part is traditionally called Nested Fixed Point (NFPX) algorithm, which has been implemented in Gauss, using procedure IMPVALUE.PRC, which can also be used for the calculation of likelihood:

$$\sum_{i=1}^{N} \ln \int_{-\infty}^{\infty} \prod_{t=1}^{n_i} \left[F\left(\frac{\beta_1 c+\beta_2 c^2 - 2\beta_2 c(I_{it}-PM_{it})-\gamma M_{it}c+\delta\Delta V(\beta x_{it},\delta,w_i,\sigma_e)}{\sigma_e} \right) \right]^{a_{it}}$$

$$\times \left[1-F\left(\frac{\beta_1 c+\beta_2 c^2 - 2\beta_2 c(I_{it}-PM_{it})-\gamma M_{it}c+\delta\Delta V(\beta x_{it},\delta,w_i,\sigma_e)}{\sigma_e} \right) \right]^{1-a_{it}}$$

$$\times \frac{1}{\sigma_\varepsilon} f\left(\frac{w_i}{\sigma_\varepsilon}\right) dw_i \rightarrow \max_{\beta,\delta,\sigma_e,\sigma_\varepsilon} \quad (24)$$

where $F(.)$ is a culmulative standard normal distribution, $f(.)$ is a density of the standard normal distribution. This procedure is realized in the file DP_FUNCTION.PRC.

Realization of the whole algorithm is shown in the Figure 2; while the Idea of the algorithm and structure of files are based on Duan (2000).

Implementing this algorithm lead us to the following results:

σ_w	σ_e	δ	α_1	α_2	α_3	β_1	β_2	γ	P	c
0.97	0.99	0.68	0.52	-0.26	0.03	0.48	0.48	0.55	0.45	0.56

Table 2. The MLE Parameters. Log-Likelihood=-3067.944

All coefficients appear to be significant (jointly and separately). Looking at the results of the model it can be seen that the probability of a second child is minimal, while the true probability values (frequency data) show that third child is most unlikely. It is important that high value of variances of unobserved effects contribute to the major non-planned character of childbearing decisions.

Number of Children	0	1	2	3	4	5	6	7
Actual Probability of Birth	0.079	0.048	0.019	0.019	0.105	0.000	0.250	1.000
Predicted Probability of Birth	0.078	0.048	0.053	0.081	0.127	0.220	0.709	0.794

Table 3. Actual and Predicted Probabilities of Birth Given Number of Children.

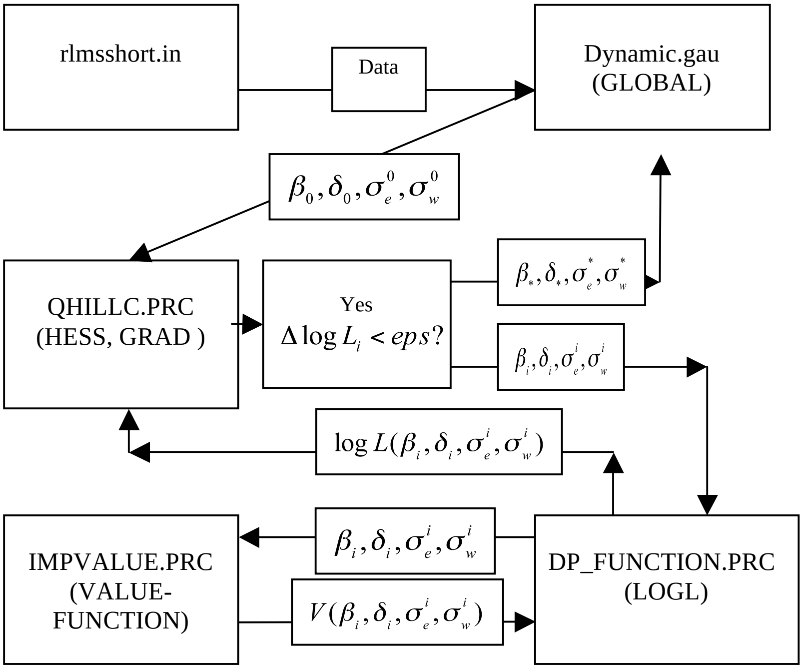

Figure 2. The NFPX algorithm.

5. Conclusions

Both structural and reduced form analysis support the hypothesis that probability of birth initially decreases with the increase of the number of children and then increases (quadratic nature). Structural form models help to understand the crucial importance of “taste for children”, which should be taken into account in the process of demographic policy formation, which should concentrate on propaganda of traditional values and big families.

Individual-specific effect was significant only in the structural form model, while probit random effect model has shown insignificant contribution of individual unobserved heterogeneity to the total variance and its coefficients are similar to those in pooled probit model. Moreover, high value of variance in random shock in the structural model demonstrates large proportion of non-planned births. These results don’t contradict economic intuition and results obtained in other studies.

References

1. Aguirregabiria V., Mira P. (2007) "Dynamic Discrete Choice Structural Models: A Survey". University of Toronto, Working Paper №297.
2. Arroyo C. R., Zhang J. (1997) "Dynamic microeconomic models of fertility choice: A survey". Journal of Population Economics, 10, pp. 23-65.
3. Avdeev A., Monnier A. (1995) "A Survey of Modern Russian Fertility". Population: An English Selection, 7, pp. 1-38.
4. Bingley P., Lanot G. (2004) "Employer pay policies, public transfers and the retirement decisions of men and women in Denmark". European Economic Review, 48, pp. 181-200.
5. Duan J.C., (2000) "Correction: Maximum Likelihood Estimation Using Price Data of the Derivative Contract". Mathematical Finance 10, pp.461-462.
6. Grogan L. (2003) "What Caused the Post-Transition Fertility Decline in Central and Eastern Europe and the Former Soviet Union". Mimeo University of Guelph.
7. Heleniak T. (1995) "Economic Transition and Demographic Change in Russia, 1989–1995". Post-Soviet Geography, 36, pp. 446–458.
8. Keane M.P., Wolpin K.I. (1994): "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence". Review of Economics and Statistics, 76, pp.648-72.
9. Kohler H., I. Kohler (2001) "Fertility Decline in Russia in the Early and Mid 1990s: The Role of Economic Uncertainty and Labour Market Crisis". European Journal of Population, 18, pp.233-262.
10. Leung S (1991) "Stochastic Dynamic Analysis of Parental Sex Preferences and Fertility". Quarterly Journal of Economics, 106, pp. 1063-1088.
11. Maleva T., Sinyavskaya O. (2006) "Social and Economic Factors of Fertility in Russia: Empiric Measurements and Challenges of Social Policy". SPERO, №5. pp. 70-97. (Rus)
12. Pakes A. (1986) "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks". Econometrica, 54, pp. 755-84.
13. Philipov D., Kohler H. (2001) "Tempo Effects in the Fertility Decline in Eastern Europe: Evidince from Bulgaria, the Czech Republic, Hungary, Poland, and Russia". European Journal of Population, 17, pp. 37-60.
14. Rimashevskaya N. (1997) "Social Results of Russian Economic Transformation". SOTSIS, 6. pp. 55-65. (Rus)
15. Roshina Y. (2006) "Modelling of Fertility Factors in Russian Families". SPERO, 5, pp. 98-133.
16. Roshina Y., Boykov A. (2005) "Factors of fertility in modern Russia". Moscow: EERC Working Paper Series.

17. Rust J. (1987) “Optimal replacement of GMS bus engines: an empirical model of Harold Zurcher” Econometrica, 55, pp. 999-1033.
18. Vishnevsky A.G. and others (2006). “Demographic modernization of Russia, 1900-2000”. New Publishing House.
19. Wolpin K. (1984) “An Estimable Dynamic Stochastic Model of Fertility and Child Mortality". Journal of Political Economy, 92, pp. 852-874.
20. Zaharov S., Ibragimova M., Kartseva M., Sinyavskaya M. (2007) “Behavior of women in the labor market and fertility in modern Russia”. Moscow: EERC Working Paper Series.

Appendix 1. Estimation Results

Variables	Total		Unmarried		Married	
	df/dx	SE	df/dx	SE	df/dx	SE
Age (omitted: 16-19): 20-24	0.0020	0.0050	0.0100	0.0080	**-0.0160**	0.0050
25-29	-0.0030	0.0060	0.0070	0.0110	**-0.0200**	0.0050
30-34	**-0.0170**	0.0050	-0.0020	0.0110	**-0.0290**	0.0060
35-39	**-0.0300**	0.0040	*-0.0190*	0.0080	**-0.0370**	0.0070
Married	**0.0350**	0.0040				
Use of Contraception	-0.0030	0.0040	0.0110	0.0070	-0.0070	0.0040
Education (omitted: 8 years): Secondary	-0.0010	0.0040	0.0040	0.0070	-0.0050	0.0060
Vocational education	-0.0020	0.0050	-0.0060	0.0080	-0.0050	0.0060
University degree	-0.0030	0.0060	-0.0070	0.0090	-0.0030	0.0060
Employment (omitted: none): Employed	**0.0130**	0.0040	*0.0150*	0.0070	**0.0110**	0.0040
On Maternity Leave	-0.0020	0.0070	0.0370	0.0460	-0.0090	0.0060
Number of Children	**-0.0290**	0.0040	0.0000	0.0140	**-0.0240**	0.0050
Number of Children^2	**0.0060**	0.0010	-0.0090	0.0080	**0.0050**	0.0010
Health: Average (omitted: bad)	0.0040	0.0080	-0.0010	0.0120	0.0050	0.0080
Good	0.0050	0.0080	-0.0020	0.0120	0.0080	0.0090
Drinking (omitted: abstain): 1 Time a Month	0.0090	0.0050	0.0130	0.0090	0.0050	0.0050
2-3 Times a Month	0.0070	0.0040	***0.0200***	0.0080	0.0010	0.0050
1 Time a Week	0.0080	0.0060	0.0220	0.0120	0.0010	0.0070
Over 2 Times a Week	0.0060	0.0080	0.0330	0.0190	-0.0050	0.0080
Real Income/100	0.0010	0.0010	0.0010	0.0020	0.0000	0.0020
Real Income=0	0.0030	0.0040	0.0130	0.0070	-0.0020	0.0040
Additional Income of The Household /100	0.0000	0.0000	-0.0010	0.0010	0.0010	0.0010
Additional Income of The Household=0	0.0030	0.0050	-0.0050	0.0070	0.0050	0.0040
Living space (omitted 1^{st}): 2^{nd} Quintile	0.0020	0.0050	0.0060	0.0090	-0.0020	0.0050
3^{rd} Quintile	0.0010	0.0050	0.0000	0.0080	0.0030	0.0060

Variables	Total		Unmarried		Married	
	df/dx	SE	df/dx	SE	df/dx	SE
4th Quintile	0.0060	0.0050	-0.0030	0.0080	<u>0.0120</u>	0.0070
5th Quintile	0.0100	0.0060	-0.0040	0.0080	<u>0.0150</u>	0.0080
Number of other members of the household	**0.0040**	0.0010	<u>0.0030</u>	0.0020	0.0020	0.0010
Number of desired additional children	**0.0440**	0.0050	**0.0220**	0.0070	**0.0530**	0.0070
Number of desirable additional children^2	**-0.0140**	0.0020	**-0.0070**	0.0030	**-0.0170**	0.0020
Lives in (omitted: city >1 mln): Regional Center	0.0010	0.0050	-0.0010	0.0080	-0.0020	0.0060
Small Town	*0.0120*	0.0060	0.0110	0.0090	0.0060	0.0060
Countryside	**0.0170**	0.0060	<u>0.0190</u>	0.0100	0.0060	0.0060
Region (omitted: central): North-West	-0.0000	0.0060	-0.0020	0.0090	0.0040	0.0070
South	-0.0010	0.0050	-0.0040	0.0080	0.0070	0.0070
Volga	0.0010	0.0040	0.0020	0.0080	0.0030	0.0050
Ural	0.0120	0.0080	0.0080	0.0110	0.0080	0.0090
Siberia	-0.0040	0.0050	0.0010	0.0090	-0.0060	0.0060
Far-East	**0.0300**	0.0110	<u>0.0310</u>	0.0170	*0.0290*	0.0150
Year (omitted '94): 1995	0.0060	0.0070	-0.0060	0.0090	0.0140	0.0100
1996	0.0030	0.0060	-0.0040	0.0090	0.0070	0.0090
1998	-0.0010	0.0060	-0.0090	0.0080	0.0050	0.0090
2000	-0.0010	0.0060	-0.0030	0.0090	0.0010	0.0080
2001	-0.0020	0.0060	-0.0040	0.0080	0.0030	0.0070
2002	0.0000	0.0060	-0.0030	0.0090	0.0030	0.0070
2003	0.0000	0.0060	*-0.0140*	0.0070	0.0090	0.0090
Husband's age					<u>-0.0010</u>	0.0000
Husband is employed					0.0020	0.0050
Husband's health (omitted: bad): Average					-0.0140	0.0110
Good					-0.0090	0.0110
Husband's drinking (omitted: abstain): 1 Time a Month					0.0040	0.0070
2-3 Times a Month					-0.0030	0.0050
1 Time a Week					0.0000	0.0050
2 Times a Week and More					<u>-0.0090</u>	0.0050
Husband's Real Income/100					0.0000	0.0010
Husband's Real Income=0					0.0010	0.0050
N		11768		4504		6361
Statistical Coefficients	Wald $\chi^2(46)$=471.5 Prob>χ^2= 0.000 Pseudo R^2=0.126		Wald $\chi^2(45)$=121.5 Prob>χ^2= 0.0000 Pseudo R^2=0.071		Wald $\chi^2(55)$=509.0 Prob>χ^2= 0.0000 Pseudo R^2=0.202	

Table 4. Pooled probit model, marginal effects. Significance of the coefficient under respective confidence level 1% – **bold**; 5% – *italics*, 10% – <u>underlined</u>.

Variables	Total		Unmarried		Married	
	df/dx	SE	df/dx	SE	df/dx	SE
Age (omitted: 16-19): 20-24	0.0270	0.0740	0.1450	0.1050	**-0.3430**	0.1270
25-29	-0.0320	0.0880	0.1040	0.1450	**-0.4110**	0.1430
30-34	**-0.2890**	0.1030	-0.0250	0.1800	**-0.6480**	0.1700
35-39	**-0.6070**	0.1240	-0.3730	0.2300	**-0.8670**	0.2070
Married	**0.5880**	0.0640				
Use of Contraception	-0.0360	0.0530	<u>0.1560</u>	0.0940	<u>-0.1220</u>	0.0690
Education (omitted: 8 years): Secondary	-0.0070	0.0670	0.0720	0.0980	-0.0840	0.1020
Vocational education	-0.0240	0.0790	-0.0910	0.1360	-0.1030	0.1110
University degree	-0.0490	0.0930	-0.1150	0.1690	-0.0680	0.1270
Employment (omitted: none): Employed	**0.1970**	0.0590	*0.2190*	0.1000	*0.2040*	0.0810
On Maternity Leave	-0.0420	0.1080	0.3860	0.3160	-0.2000	0.1350
Number of Children	**-0.4620**	0.0620	-0.0210	0.2400	**-0.4350**	0.0780
Number of Children 2	**0.0880**	0.0110	-0.1360	0.1430	**0.0880**	0.0130
Health: Average (omitted: bad) Average	0.0550	0.1150	-0.0170	0.1840	0.0980	0.1610
Good	0.0770	0.1170	-0.0240	0.1860	0.1420	0.1650
Drinking (omitted: abstain): 1 Time a Month	*0.1260*	0.0620	<u>0.1840</u>	0.1090	0.0920	0.0860
2-3 Times a Month	<u>0.1130</u>	0.0610	**0.2730**	0.1020	0.0140	0.0880
1 Time a Week	0.1150	0.0780	*0.2780*	0.1310	0.0180	0.1150
Over 2 Times a Week	0.0950	0.1070	*0.3870*	0.1650	-0.0830	0.1680
Real Income/100	0.0080	0.0210	0.0130	0.0370	0.0050	0.0300
Real Income=0	0.0400	0.0580	*0.1870*	0.0930	-0.0370	0.0840
Additional Income of The Household /100	0.0060	0.0060	-0.0080	.0110	<u>0.0180</u>	0.0100
Additional Income of The Household=0	0.0390	0.0780	-0.0780	0.1180	0.0800	0.0700
Living space (omitted 1st): 2nd Quintile	0.0350	0.0730	0.0910	0.1250	-0.0400	0.1020
3rd Quintile	0.0180	0.0760	0.0020	0.1300	0.0550	0.1030
4th Quintile	0.0970	0.0760	-0.0380	0.1310	<u>0.1890</u>	0.1030
5th Quintile	<u>0.1380</u>	0.0820	-0.0700	0.1430	*0.2280*	0.1100
Number of other members of the household	**0.0600**	0.0180	<u>0.0530</u>	0.0290	0.0390	0.0270
Number of desirable additional children	**0.6840**	0.0740	**0.3380**	0.1180	**0.9450**	0.1130
Number of desirable additional children ^2	**-0.2160**	0.0240	**-0.1100**	0.0420	**-0.2990**	0.0350
Lives in (omitted: city >1 mln): Regional Center	0.0160	0.0780	-0.0140	0.1280	-0.0240	0.1100
Small Town	*0.1790*	0.0760	0.1630	0.1260	0.1070	0.1060

Variables	Total		Unmarried		Married	
	df/dx	SE	df/dx	SE	df/dx	SE
Countryside	**0.2440**	0.0790	*0.2730*	0.1300	0.1150	0.1130
Region (omitted: central): North-West	-0.0040	0.0870	-0.0320	0.1510	0.0650	0.1210
South	-0.0100	0.0800	-0.0640	0.1360	0.1210	0.1100
Volga	0.0120	0.0710	0.0340	0.1170	0.0470	0.0990
Ural	<u>0.1670</u>	0.0880	0.1230	0.1500	0.1250	0.1270
Siberia	-0.0710	0.0860	0.0190	0.1400	-0.1090	0.1210
Far-East	**0.3470**	0.1030	*0.3610*	0.1630	*0.3760*	0.1480
Year (omitted '94): 1995	0.0860	0.0990	-0.0920	0.1690	0.2100	0.1370
1996	0.0430	0.0980	-0.0600	0.1590	0.1140	0.1400
1998	-0.0100	0.0980	-0.1440	0.1630	0.0830	0.1390
2000	-0.0150	0.0950	-0.0470	0.1520	0.0130	0.1360
2001	-0.0370	0.0920	-0.0590	0.1470	0.0410	0.1320
2002	0.0010	0.0900	-0.0370	0.1460	0.0470	0.1290
2003	0.0060	0.0900	-0.2400	0.1550	0.1490	0.1270
Husband's age					<u>-0.0140</u>	0.0080
Husband is employed					0.0380	0.0880
Husband's health (omitted: bad): Average					-0.2550	0.1750
Good					-0.1660	0.1750
Husband's drinking (omitted: abstain): 1 Time a Month					0.0700	0.1110
2-3 Times a Month					-0.0590	0.0950
1 Time a Week					-0.0020	0.0990
2 Times a Week and More					-0.1700	0.1100
Husband's Real Income/100					0.0030	0.0130
Husband's Real Income=0					0.0190	0.0910
N	11768		4504		4922	
Statistical Coefficients	Wald $\chi^2(46)$=444.9 Prob>χ^2= 0.000 LR ρ=0: $\chi^2(01)$=1.79 Prob>=χ^2=0.091		Wald $\chi^2(45)$=77.5 Prob>χ^2= 0.002 LR ρ=0: $\chi^2(01)$=0.27 Prob>=χ^2=0.301		Wald $\chi^2(55)$=396.5 Prob>χ^2= 0.000 LR ρ=0: $\chi^2(01)$=2.74 Prob>=χ^2=0.049	

Table 5. Panel probit random effect model, marginal effects.
Significance of the coefficient under respective confidence level 1% – **bold**; 5% – *italics*, 10% – <u>underlined</u>.