

Dimensions of individual well-being in post-crisis Russia: income and labour market status vs. happiness

Ekaterina Selezneva

Affiliation: Graduate student, PhD in Economics Programme (DSE), University of Turin (IT)

Address: via Real Collegio 30, Moncalieri (TO), Italy

E-mail: eka.selez@gmail.com

Abstract

Financial crisis of 1998 could be considered as a starting point of the new stage of the transformation of both, Russian economy and individual perception of well-being. Rather successful macroeconomic performance of the country in the beginning of the 2000s and recognition of the market economy status by the end of 2002 coincided with the growth of average satisfaction level of population. At the same time, satisfaction of the employed subgroup of the population was growing not faster, and sometimes even slower, than for all the other labour market categories. In the paper, it is made an attempt to figure out whether this phenomenon could be explained by the particular importance of inter-personal/family/time period comparisons of her relative position performed for the subjective well-being evaluation, especially if employed and non-employed sub-samples of population are considered.

Updated: 12 June 2007

1. Introduction

There exists a bunch of papers investigating a satisfaction of the Russian population since the beginning of 1990s (e.g. Ravallion and Lokshin (2001), Graham et al. (2004) and so on). Both, similarity and uniqueness of the Russian experience are noted, in particularity, when a role of subjectivity of happiness evaluation is investigated. Most of the studies are not going further than the year of 2001, and explore the issues related to the first decade of post-soviet transition of the country. Present research takes into consideration a period of 2000 – 2004, shortly after the default of 1998, which became an impulse to a new stage of the Russian economy transformation. The financial crisis developed into a new redistribution of economic resources and roles in the society, and as one can assume, brought some up-to-date ideas about successfulness of an individual in the market economy context. These developments should have influenced a process of general satisfaction formation.

An attempt to explain a disparity between behaviour of general life satisfaction measures on macro and micro levels in Russia after the financial crisis of 1998 is made in the present study. Main attention is paid to the role of a labour market status and economic characteristics, including both, absolute and relative indicators' role in a process of subjective satisfaction evaluation and explanation of the contradictory phenomena found.

Life-dimension(s) other than income that could provides individuals (and hence, a society) with utility (and hence, increase well-being) are widely searched for in happiness research during the last decades. At the same time, utility, earlier thought as a latent variable behind the satisfaction levels reported, nowadays tends to be considered as having several manifested forms and being distinguished from happiness. Data investigated in this study provide an evidence of particular importance of subjective position in the society evaluation, which is done thought through a prism of personality. Objective and subjective well-being are not identical, especially in tumultuous economic environment, and hence could result in presence of the happiness-utility duality.

The rest of the paper is organized as follows. Section 2 contains short literature overview of satisfaction and happiness studies in general and, in particularity, in Russia. Section 3 contains stylized facts about happiness, income and labour market status found from the RLMS data for the period of 2000-2004. Further, in Section 4, model and estimation strategies are stated. Section 5 includes the Russian Longitudinal Monitoring Survey dataset description, variables selected for the analysis and comment on the comparative characteristics of different sub-sets of observations actually involved in estimation (drops out and missing values analysis). Section 6

discusses the estimation results under different econometric specifications and composition of explanatory variables.

2. Literature overview

1) Short review of satisfaction and happiness studies

Subjective well-being research in economics is pioneered by Easterlin (1974)¹, who found that individual happiness differences among the population of some country can be explained by difference in economic conditions. In 1970s, main focus of interest was placed on the issue of financial satisfaction or ‘individual welfare of income’ investigated, in particularity, with help of Income Evaluation Question formulated by Van Praag (1971) (Leyden School²: van Praag, Kapteyn, Hageaars). Nowadays group of satisfaction questions is widely included in surveys’ questionnaires aiming to evaluate personal satisfaction with different (not only financial) life-dimensions. Answers are coded with help of some discrete response scale, and individuals (supposed to) give similar responses, namely, the same level of happiness achieved³, in similar circumstances. Such life-satisfaction subjective measures are proved to have biological foundations, (e.g. review in Ng (1997) and being able to contribute for economics research and even serve as guidelines for socio-economic policies⁴ (e.g., reviews in Frey and Stutzer (2002, 2005), van Praag (2007)).

Nowadays, the following terms are widely used as synonymous: happiness, satisfaction, utility, subjective well-being, welfare. But historically, concept of satisfaction/happiness is based on utility theory. Classical utilitarianism tradition (Bentham) was assuming any experience to be measurable on the “pain-pleasure scale”, and utility considered as a metric to asses an individual well-being. Pleasure as the ultimate goal pushes a rational individual to make choices providing her with maximum of utility (which is used as a synonymous for “pleasure”). Further theory developments raise discussions related to a possibility of utility/happiness measurement itself, its comparability across people and other ‘technical’ aspects⁵. Moreover, source of utility was seen mainly in income and/or basket of goods and opportunities faced by consumer (e.g. Edgeworth) and concept of welfare coincided with satisfaction of rational individual preferences through

¹ Before Easterlin’s work, empirical happiness studies were mainly considered as a topic of interest for sociologists, linking happiness and social indicators research (e.g. Andrews and Withey, 1976), or psychologists exploiting satisfaction as an indicator of mental health (e.g. Jahoda (1958).

² Main ideas and approaches of the Leyden school could be found in van Praag (2007).

³ This is a ‘common background’ assumption of interpersonal comparability of satisfaction levels.

⁴ For example, policies related to such questions as wage gaps between genders or races, taxation of goods that create addiction and so on.

⁵ For example, Pareto doubted about a possibility to estimate ‘happiness functions’ as well as Robbins (1932): (psychology can help in explanation of economic behaviour, but happiness is not measurable), Samuelson (1947): (axiom of immeasurability of utility (only ordinal utility functions are possible) etc

purchasing of objects needed or desired by an individual. That is why the relatively recent Sen's critique should be particularly noted as it has shifted attention from the commodity basis of utility to individual-specific 'capabilities'. They represent all the possible well-being state(s), and hence what, actually, person could do and who could become. The second layer of the approach appeals to the well-being level reached, thus a set of functionings achieved. Talking about *subjective* well-being of an individual one needs to aggregate across different interrelated 'functionings', or "beings and doings" (real conditions of life). It is important, that "being happy" is included by Sen in a list of functionings along with other life aspects like "being adequately nourished" and so on. Individuals not only have different conditions for realization of the individual abilities (e.g. level of education, ecological situation in an area etc). Moreover, Sen emphasizes that even if having the same income and consumption individuals are heterogeneous in terms of abilities and capabilities for effective use of resources available for implementation of their "life-projects".

Other significant theory development of the recent years is a new utilitarianism. It is based on the theory of experienced utility by Kahneman and distinguishes between several types of utility, emphasizing that decisional utility reflected in choices is not the same as experienced utility. Moreover, it is argued that retrospective conclusions about utility received can differ substantially from the one actually experienced. Life satisfaction and happiness assessments are often comprised of retrospective judgments. They are done through a prism of personality. "*Subjective well-being measures features of individuals' perceptions of their experiences, not their utility as economists typically conceive of it.*" (Kahnemann and Krueger, 2006, p. 4).

Search for life-dimensions that could provide individuals (and in its turn, a society⁶) with utility (and hence, increase well-being/welfare) constitutes one of the most important directions of happiness research. Income-related issues are recognized to play important but far from exhaustive role in a process of satisfaction evaluation (list of happiness covariates is shown to include, for example, 'big seven' in Layard (2005): family relationships, financial situation, work, community and friends, health, personal freedom and personal values). Recently, a particular attention is paid to the role of subjective and objective individual comparisons with some reference group⁷, individual relative position in different spheres of life (for example, individual rank in income distribution) and such phenomenon as aspirations and habituation⁸. Other important aspect of research work is related to heterogeneity of individuals that should be, if possible, taken into account (which agrees with theoretic postulates proposed by Sen and

⁶ As it is supposed that aggregation of individual well-beings gives a well-being of a society

⁷ E.g. works of Clark and Oswald (1994), McBride (2001), Frijters (2006), Kahnemann and Krueger (2006) etc. but at the same time, relative utility by Veblen (1899) could be also remembered.

⁸ e.g. Brickman and Campbell (1971) introduced the concept of a 'hedonic treadmill'

Kahnemann). “Personality” (partially influenced by a social background and/or culture traditions) expresses itself in different responses to the same life changes among individuals and affects not only level of satisfaction achieved but also strength of factors’ influence on happiness⁹. Moreover, possibility of culture based “default” answers should be taken into account when considering religiously, ethnically (and so on) complex and heterogeneous group of individuals¹⁰.

All abovementioned make us think that subjective and objective well-being is often different, and happiness-utility duality would be expected to be found in the Russian data concerning the tumultuous period after the crisis of 1998.

2) Main findings on satisfaction for Russia

There is number of papers attempting to shed light on the mechanisms underlying principles of subjective well-being evaluation in Russia during the transformation period followed dissolution of the Soviet Union. Main findings of some of these studies are summarized in a table A.1 in appendix. By consideration of different time sub-periods there were revealed a lot of similarities of the Russian experience to the typical world-wide patterns (for example, positive influence on the perceived well-being of a household (equivalised) income, being male, having higher education, good health, being employed, richer area/urban context; U-shaped form of the life-satisfaction – age relationship; negative impact of divorce or unemployment status).

At the same time, some phenomenon caused by country and culture specific and possibly by transformational character of the period is also found. Senik (2004) investigated a period from 1994 to 2000, noted the presence of unusual ‘optimistic expectations’ based on reference group income comparisons (similar to the “tunnel effect”, when individuals base their expectations about own further progression observing others, Hirschman (1973)). She also mentions that Ravallion and Lokshin (2000) had already upheld the presence of this phenomenon founding that positive attitude towards restriction of the income of the rich among a group with most uncertain perspectives. Frijters et al (2006), for 1994 – 2001, underlined importance of personal traits, showed consistence of main parameter estimates with those for the Great Britain and Germany. Although an uncommon detrimental effect of local unemployment rates on happiness level is found (higher the local unemployment rates, the more detrimental own unemployment). Eggers, Gaddy & Graham (2006) also turn to the role of individual self-comparisons with some reference group and investigated regional unemployment rates influence on self-satisfaction responses. They noted an important role of the peers groups for individual well-being evaluation and argued

⁹ e.g, Winkelmann and Winkelmann (1998), Clark et al. (2004)

¹⁰ Inter-countries/cultural comparisons in Tucker et al. (2006)

for a phenomenon of life standards lowering in transition (not favorable economic situation). These two leading, for example, to augmentation of the happiness level in the regions with higher unemployment rates.

Such uncommon directions of the factors influence were, rather probably, caused by the transitional character of the period in consideration. One can also argue that new period of changes that started after 1998 was different from the transformation of the planning economy into a market one in the beginning of 1990s. Moreover, it would be expected that not only structure of economy was changed (by 2002, when market economy status was granted to the country by EC) but similarly to ‘modernization of incomes’¹¹ perception of (‘subjective concept’) of welfare was reviewed in Russia.

3. Stylized facts: Russia after the crisis of 1998

1) General life satisfaction vs. economic situation

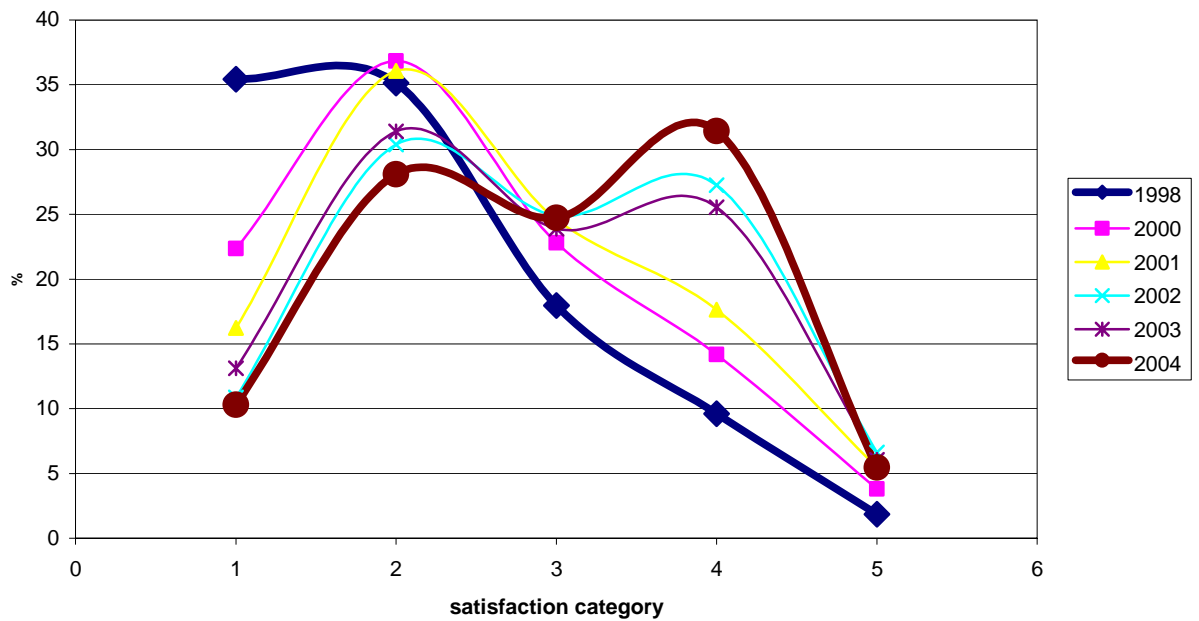
It is noted that ‘V-shaped’ behaviour of the main macro-economic indicators since the beginning of the 1990s till 2004 was mirrored quite well by the dynamic of aggregate satisfaction measures (e.g. Frijters et al. (2006)). Figures of the official statistics by Central Statistical Agency (GosComStat) for 2000 – 2004 demonstrate improvement of economic situation after the crisis, average real incomes were growing, labour market situation ameliorated¹².

Making rather standard (for happiness studies) assumptions that individual life satisfaction levels are comparable and additive across individuals, due their cardinal nature, we estimate an “average” satisfaction level for the sample in consideration. During the period from 1998 till 2004, this average level was increasing but always staid slightly lower than ‘neutral happiness level’ (if satisfaction is coded by five levels from 1 to 5. 1998: ‘mean happiness of the sample (weighted)’ equals to 2.07, in 2004 equals to 2.94). Median satisfaction increases from “rather unsatisfied” till the neutral “yes and no”. If talking about the distribution of the satisfaction levels reported, over the six years in consideration, it shifted from the strongly right-skewed in 1998 to rather symmetric one (which reminds the developed countries happiness distribution pattern).

Figure 3.1 Life satisfaction answers distribution

¹¹ It was analyzed by Rose and McAllister (1996). They and showed that non-market (membership in the communist party etc) influence on income shrunk greatly and has given way to market economy determinants of incomes and welfare as, for example, education level (human capital development).

¹² See appendix for a short review of labour market and macroeconomic situation.



Notes: Question posed: “As a whole, how satisfied you are with your life in general?” Answer coded as: (1) = not at all satisfied; (2) = rather unsatisfied; (3) = yes and no; (4) = rather satisfied; (5) = absolutely satisfied.
 Sample: 15-72 years old.

It is quite expectable from the facts cited above that the period in consideration is characterized by a high mobility of happiness level reported. The most unstable group is “fully satisfied”, while the most stable ones are ‘rather satisfied’ and ‘rather unsatisfied’ categories. Transitions, if happen, are mainly to one level higher or lower (conjunctive satisfaction) category.

Table 3.1. Transitional matrix for life satisfaction levels reported (2000-2004)

From t to $t+1$	Not at all satisfied	Rather unsatisfied	Yes and no	Rather satisfied	Absolutely satisfied	Total, %
Not at all satisfied	36.68	35.97	16.03	9.06	2.26	100.00
Rather unsatisfied	12.86	43.8	23.78	17.25	2.31	100.00
Yes and no	6.93	28.44	35.95	25.05	3.62	100.00
Rather satisfied	4.05	19.1	23.29	45.55	8.01	100.00
Absolutely satisfied	4.53	13.07	16.72	38.85	26.84	100.00
Total	12.66	31.85	25.06	25.22	5.21	100.00

Sub-sample: 15-72 years old.

Concluding, one can say that average level and distribution pattern, representing macro level, are changing in accordance with predictions of the standard welfare theory.

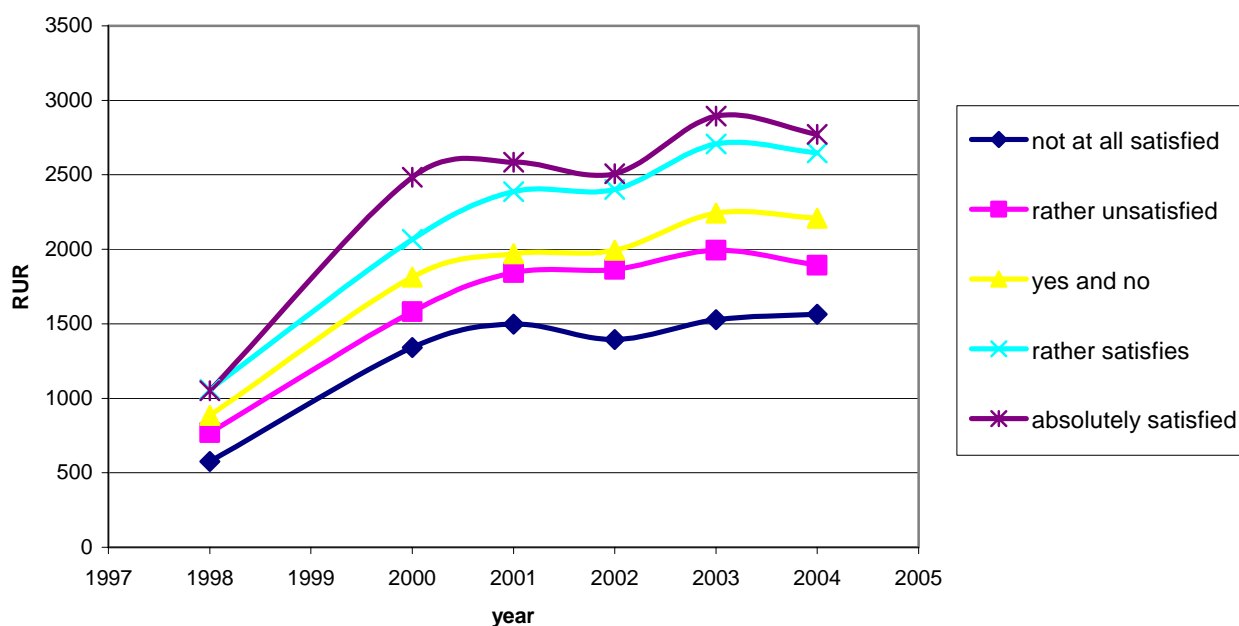
2) General life satisfaction vs Household income

Financial crisis of 1998 swallowed a great part of the population’s savings and financial investments. New distribution of resources pushed a new wave of changes in economy structure

to start. Some researchers note that it was ‘the most well-provided part of population’, who suffered most from the crisis of 1998 (e.g. Kislitsyna (2003)) although an increased inequality and number of people under poverty line cannot be lived behind. As analytics of the Demoscope Weekly note¹³, in 1999 only in 4% of regions poor was forming less than 20% of population, while by 2005 it was already true for 40% of regions. Moreover, only by the end of 2002 per capita incomes of population achieved the level of 1997. Hence, several years after the crisis called primary needs to come again to the fore. Thus one could expect an increasing role of income/expenditures categories role in affecting individual satisfaction. Nevertheless, different strength of the effect (for other variables also) can be expected for different income groups.

RLMS reported incomes are noted to be highly underrepresented and containing a high portion of ‘zeroes’ (for example, due to informal earnings or wage arrears). That is why the World Bank recommendations for income evaluation for countries with transitional economies are applied further in this paper. Moreover, period in consideration was characterized by a quite high level of inflation. Thus total household expenditures deflated back to the level of 2000 are used as a proxy of living standards instead of income. Figure below illustrate dynamics of average household (equivalised) expenditures over the period of 2000 - 2004.

Figure 3.2 Median per capita equalized (household) expenditures by satisfaction levels



Notes: household total per capita expenditures are in prices of 2000, equivalised by OECD scale.

¹³ Weekly bulletin of the Center of Human Demography and Ecology of the Russian Academy of Sciences, N.273-274, available in Russian on <http://demoscope.ru/weekly/2007/0273/tema02.php>

One can note some slight increase of median per capita equivalized expenditures over the period, although for the vast year some decline for nearly all categories could be noted. Nevertheless, “happier” categories are always associated with higher average per capita expenditures.

Table 3.2 Transition matrix for quintiles of per capita total expenditure, OECD scale, in terms of 2000

Quintiles	1 (poorest)	2	3	4	5 (richest)	Total, %
1 (poorest)	45.38	29.79	14.55	6.77	3.51	100.00
2	17	32.69	26.97	16.78	6.56	100.00
3	6.44	20.62	30.14	29.01	13.78	100.00
4	3.8	9.56	23.93	32.99	29.71	100.00
5 (richest)	1.77	5.04	10.78	26.47	55.94	100.00
Total	12.58	18.71	21.9	23.72	23.09	100.00

Notes: period of 2000 – 2004.

And if the happiest and least happy categories are rather unstable, the less mobile income/expenditure quintiles are the poorest and the richest ones (was also noted, for example, by Bogomolova and Tapilina, 1999). One more fact could be cited, that happier categories have on average, greater number of household members with an increasing proportion of household members older than 15 having job. While in other quintiles such a household “employment rate” is decreasing.

3) General life satisfaction vs. Labour market status

This paper considers individuals in the age range from 15 till 72 years old. It covers those who are considered by the official statistics as potentially economically active in Russia. Moreover, the lower bound of 15 years old is linked to an age when the obligatory education is completed. “Employment” group, including those working and on paid/unpaid leave, or self-employment, is found to be rather stable (about 90% of those who had job, report being employed one year after) in contrast to “Unemployment” category (just about 19% stays in this category after one year). Those in long-term unemployment are, on average, the least happy¹⁴ (it reflects stressful situation of uncertainty and/or institutional settings with poor unemployment benefits, lack of incentives to be registered etc). Rather predictably, transition into employment increases personal satisfaction level. Average satisfaction levels in ‘employed’ and ‘others’ (including both unemployed and economically inactive) categories are presented in a table 3.3.

Table 3.3. Average satisfaction level and its growth rate by labour forced categories

year	1998	2000	2001	2002	2003	2004
<i>Average satisfaction level by categories</i>						
Out of labour	1.09	1.40	1.63	1.92	1.78	1.90

¹⁴ This is matching with a strong detrimental impact of unemployment on happiness found by a number of researchers, for example, Winkelmann & Winkelmann (1998), Blanchflower and Oswald (2004).

force						
Unemployed	0.82	1.07	1.21	1.50	1.44	1.52
Employed	1.09	1.43	1.61	1.89	1.84	1.99
<i>Growth rate of average satisfaction by categories (relative to previous year)</i>						
Out of labour force	---	1.29	1.17	1.17	0.93	1.07
Unemployed	---	1.30	1.13	1.24	0.96	1.06
Employed	---	1.31	1.12	1.17	0.97	1.08

Sample: 15-72 years old. Satisfaction is coded on the scale from (0) not at all satisfied, (1) rather unsatisfied, (2) yes and no, (3) rather satisfied (4) absolutely satisfied. Individual weights applied.

It is noticeable that average satisfaction level for employed is higher than for all the others (those who have no job due unemployment or inactivity). Considering growth rates for these two categories it can be seen that happiness level for members of employment group is growing slower than for others during the years right after the crisis. It means that having a job doesn't always mean to be satisfied with life. This could be possibly explained by a concept of "good" and "bad" jobs (the former meaning more stable and well-paid employment), or characteristically for transitional economies phenomenon of "escaping" into informality¹⁵, or 'combination of economies' to work in, due tax evasion or survival strategy (especially in cases of labor underutilization, arrears (which were mainly paid off by the 2000) and forced leaves particularly widespread in mid-1990s, e.g Kim (2002)). Other negative tendency is 'working poor' phenomenon. One more, but less possible explanation is that in the period under consideration job was not among the first-importance spheres of life for satisfaction evaluation.

Summarizing the phenomena described in the stylized facts section, it can be seen that even if the 'social (aggregated) satisfaction' distribution was changing its shape from a strongly skewed right 1998 to a 'nearly symmetric' form of 2004 (which reflects general world-wide tendency) due to prediction of welfare theory, happiness among working people was not growing significantly faster than for the whole population (and sometimes even slower than for unemployed ones). It makes us arguing that, happiness on the aggregated and individual levels behaved differently. Moreover, in the period considered objective and subjective well-being were not the same, the later growing slower than the former one. This tendency could be induced by the increased role of interpersonal comparisons, which is quite characteristic for unstable environments (e.g. Hirschman's tunnel effect).

¹⁵ E.g. Kalugina and Najman (2003) found negative correlation between poverty and working in informal sector.

4. Model and estimation strategy

1) Model

In this paper we pose a question of whether the gap between subjective and objective well-being during the transformation period could be explained by a combination of expectations/aspirations and ‘indicators’ of comparisons performed by individual (where as a benchmark can be considered or some reference group, or past conditions/situation for the person herself).

We investigate, primarily, the corner-stones of the “Happiness¹⁶ – Labour Market – Income” well-being triangle, assuming that satisfaction is influenced by the two other dimensions (interrelated, but not mirroring each other due to transformational character of economy structure). Hence individual labour market status and relative poverty household position should be considered, using both, subjective and objective ‘position’ measures. For example, objective measure reflecting a relative position of individual in some sphere of life could be relative poverty of a household¹⁷, or some indicator for the conditions of a local labour market (e.g. unemployment rate); than among the subjective measures one would place future expectations or an economical rank of the family reported by individual. When possible, regional heterogeneity should be accounted for.

Subjective individual life satisfaction, the latent one, is than considered to be a function of the characteristics cited above with socio-demographic characteristics as controls.

$LS^* = LS^*(Income_characteristics, Labour_market_characteristics, soc.demogr_controls; \beta)$
We are assuming that:

- (a) Absolute level of income (or some proxy of it) should have positive but decreasing effect on subjective general life satisfaction. Labour status is supposed to have significant positive (if employed) and negative (if unemployed) effect.
- (b) Comparison factors should play significant role in a process of individual satisfaction evaluation. Moreover, subjective evaluation of the situation could be more important for the current personal happiness evaluation than ‘relative subjective’ (factors due to inability of individuals to evaluate them in a correct way).
- (c) For the local labour market characteristics no clear pattern can be expected as one of the following scenario is possible: (1) high level of unemployment would push down satisfaction level of unemployed as they feel they have less chances to find a job, while satisfaction of employed is pushed up as even a bad job can be considered as a good one. (2) high level of

¹⁶ In the present text the following terms are used as interchangeable: happiness, subjective well-being, general life satisfaction.

¹⁷ For example, relative expenditure poverty could be estimated as $REP = \frac{expenditures}{mean(expenditures)}$

unemployment makes unemployed and employed less certain to find a new job, but the former could feel happier as it is easier to justify their not working status, while latter would be less happy being concerned with loosing their jobs.

(d) Good aspirations for close future economical conditions for a household should push the life satisfaction up.

(e) Moreover, the effect of different groups of the explanatory variables could differ for two subsamples, of employed and ‘others’.

2) Modeling strategies

Making a standard assumption that all respondents understand a satisfaction related question in the same (or at least, similar) manner and that answers can be interpersonally comparable at least in the ordinal sense, we impose that level of satisfaction reported by individual in a moment t (LS_{it}) is based on discretization ($f(\cdot)$) of some latent continuous well-being function LS_{it}^* :

$$LifeSatisfaction_{reported} = LS_{it} = f\left(LS_{it}^*(t, X_{it}, \dots)\right) + v_{it}$$

Individuals are indexed by $i = 1, \dots, N$, and time point by $t = 1, \dots, T$. v_{it} is an error term.

Unobservable life satisfaction is a function of different characteristics, such as time-moment, vector of observable potential explanatory variables (X_{it}) which can be time- and/or individual (in)variant and so on (moreover, they could be of the personal, family and regional etc. level).

Let us suppose that observed LS_{it} (ranging from 0 to 4 in our case) is reported by individual according to the following rule ($f(\cdot)$):

$$LS_{it} = \begin{cases} 0, & \text{if } LS_{it}^* \leq \tau_1 \\ 1, & \text{if } \tau_1 < LS_{it}^* \leq \tau_2 \\ \dots & \\ 4, & \text{if } \tau_4 < LS_{it}^* \end{cases}, \text{ where } \tau_1 < \tau_2 < \tau_3 < \tau_4 \text{ are unknown threshold parameters.}$$

Two general approaches to model the satisfaction level reported are following:

- 1) Linear models¹⁸ treat well-being as being cardinal and directly comparable between individuals. Identical distance between latent variable thresholds is also implied. Dependent variable is treated as a continuous one and not assumed to be bounded.
- 2) Ordered Logit/Probit models¹⁹ assume the distances between latent variable thresholds to be variable. Individuals interpret satisfaction-scale in the same way and indicating the same level of happiness would locate two individuals on the same indifference curve. Such feature of the model helps to overcome the possible ordinality and scaling problems

¹⁸ For example, exploited in Di Tella et al., 2001; Senik, 2004; Ravallion and Lokshin, 2001

¹⁹ For example, Senik, 2004; Graham et al., 2004, Case, 2001.

in the satisfaction data (namely, the fact that different individuals could have diverse happiness scales, which in its turn leads to errors of estimation and ‘cardinal’-inconsistency, in a sense of direct incomparability of the satisfaction levels across individuals).

Given a dataset is composed of several *cross-sectional repeated samples*, one of the possible strategies would be an estimation of separate models for each round of the survey available. Another option is to investigate the pooled dataset. While in a cross-sectional case one could treat respondents as independent, different time-point observations for the same individual are not independent due to personality trait, standard errors of a model coefficients estimates are lower than the true ones, so such an intrinsic heterogeneity of respondents should be, indeed, taken into account. A way to overcome the problem is in obtaining a robust variance estimate (Huber/Whit-sandwich robust estimator of variance) adjusting for within cluster correlations (where, in our case, clusters are composed of observation for the same individual). Another solution, as dataset allows following a group of individuals over several points of time (*panel* subset of the data), unobserved heterogeneity of the population and personal traits could be not only corrected but to controlled for.

Paying attention to the personal traits is of great importance, partially, because of not random character of distribution of some significant life events across population. Marriage, childbirth, job character etc could be mentioned as examples (Frijters et al., 2006). Unobserved heterogeneity not included into model gives unreliable estimates similarly to omitting variables charged with individuals self-selection into the groups/statuses cited above. Way of treating individual heterogeneity (as latent variable or parameter to be estimated) leads to Random or Fixed effects model specification.

As noted by Ferrer-i-Carbonel and Frijters (2002), assumption done about the nature of happiness answers (cardinality or ordinality) makes no great difference for empirical findings. At the same time, shifting from random effects to fixed effects model and hence, controlling for personal traits would change the results significantly.

It is worth noting that while FE-specification for panel linear models (considering happiness as a cardinal category) is easily accessible via a variety of methods (e.g. first differencing) such approach is not directly applicable to ordered probit/logit models giving biased results. The situation with estimation procedures and ready-to-use tools of econometric packages is even worse when multiple or ordered choice models are addressed. In fact, standard econometric packages include fixed effects specification for binary choice models proposed by Chamberlain

(1980). In applied research, when dealing with ordered choice panel models, the following strategies are usually exploited:

- Random effects ordered probit model (user-written Stata procedure, `-reopro-`, by Guillaume R. Frechette, estimating the model parameters via Gauss-Hermite quadrature approximation).
- Collapsing the dependent ordinal-scale happiness variable into a dichotomic one applying the same threshold for all the individuals (for example, recoding satisfaction level higher than “neutral” into “happy” category, while all the others are considered as “unhappy”) and applying binary choice fixed effects model²⁰. All the individuals haven’t changed their “binary” satisfaction level during the period are dropped from the analysis.
- Fixed effects ordered logit model, proposed by Ferrer-i-Carbonel and Frijters (2004). Methodology also assumes that data are still collapsing to binary, enabling simple fixed effects logit estimation. At the same time, ordinality assumption is relaxed via application of individual-specific thresholds (individual specific interpretation of the happiness question). Method allows inclusion into estimating sample of all the individuals whose satisfaction scores changed (at least once)²¹.

5. Data, variables

1) Dataset

Dataset exploited is a part of the second wave of the Russian Longitudinal Monitoring Survey (RLMS)²². The second wave of the project started in 1994 with a sample of about 4000 households, with at most possible interviewing all the members of the households selected. These data form a national representative sample (not representative on a single region level) covering about 10000 individuals each round. Modules of questions repeating from round to round contain information about personal characteristics, labor market, income, health etc (individual level), consumption, living conditions etc (household level). Some groups of questions are not appearing from round to round.

²⁰ For example, Winkelmann and Winkelmann, 1998; Clark et al., 2001; Clark, 2003

²¹ Approximation of the estimation procedure is possible through creation of an indicator variable which equals to unity, if level reported is higher than individual mean. Than binary fixed effects logit procedure is applied. Such approach is exploited, in particular, also in Jones and Schurer (2007).

²² First 13 rounds of the dataset were publicly available online until 2006. Since October 2006 data are available for payment. Rounds 9-11 are in free access after registration on the web-site [Russia Longitudinal Monitoring Survey - UNC Carolina Population Center: http://www.cpc.unc.edu/rlms/](http://www.cpc.unc.edu/rlms/)
Description of the Survey, for example, in Mroz et al., 1999

Present paper consider a sub-sample of individuals reached an age of at least 15 Period covered is of 2000-2004, so mainly just five rounds (IX – XIII) of this cross-section repeated survey are exploited. Some information based on the 1998 data (Round VIII) is cited in the stylized facts section above.

2) Selected variables

Dependent variable:

The variable explored represents a discrete scale coding a subjectively reported happiness. Question asked in the RLMS survey is “To what extent are you satisfied with your life in general at the present moment?” Answer is coded as (0) not at all satisfied, (1) rather unsatisfied, (2) yes and no, (3) rather satisfied (4) absolutely satisfied²³.

Explanatory variables:

Along with labour market and economic situation characteristics, a list of some (rather standard for the happiness research) socio-demographic variables should be also controlled for. As it was already mentioned above, several life domains are generally considered as affecting a level of general life satisfaction apart of the economic aspects, for example, health or family life. Importance order of the domains can be different for different individuals (Easterlin, 2005).

Variables possibly explaining the differences in the life satisfaction levels include not only personal characteristics. The following variables were selected:

- Household’s economic situation:
 - As income proxy, household total expenditures are introduced²⁴: namely, logarithm of a household total expenditures²⁵ in prices of 2000
 - Perceived current family economic/financial situation reported as a subjective valuation of a family economical rank on the nine-step ladder from poorest to richest people (01=lowest step; 09=highest step)
 - Expectations/aspirations concerning future family financial situation next year, namely, whether the family will live better of worse (5=much better, 4=somewhat

²³ Originally coded from ‘1’ (absolutely satisfied) to ‘5’ (not at all satisfied).

²⁴ Introduced on the household level as in previous research on the Russian data it is found that household income or expenditure is a good prediction of satisfaction, better than individual income (e.g. Senik, 2002)

²⁵ RLMS constructed ‘monthly total household expenditures’ is an indicator consisting of both, food and non-food household spending (includes nominal additions to savings; payments for insurance etc). Procedure of the indicator construction assumed replacing missing amounts of purchases (if purchase is indicated) with medians (by settlement type and family size; or of entire sample).

Together with the log of expenditures log of household size is introduced to account for economy of scale in big families (logarithm is used to avoid choice of equalizing scale). Household size is estimated as a sum of members in six age-gender categories, presented in the data. If data on some of the categories were missing, they are replaced with zeroes.

better, 3=nothing will change, 2=will live somewhat worse, 1=will live much worse)

- ‘Relative household (expenditure) poverty’: as index representing the household expenditures divided by the mean expenditures among households with the same number of household members.
- Labour market related:
 - Labour market status: 2 = “Employment”²⁶ (means working including on paid/unpaid leave, or engages in individual economic activity); 1 = “Unemployment”(BLS definition); 0 = “Out of Labour Force”
 - economic region’s unemployment rate²⁷: (constructed as proportion of those unemployed relatively to the working age individuals in regional sub-sample)

Socio-demographic controls:

- Gender (dummy, 1, if female, 0 if male)
- Age and age squared (using the most probable year of birth)
- Self-perceived health (5=very good; 4=good; 3=average, not good, but not bad; 2=bad; 1=very bad)
- Education (the highest diploma obtained). Three levels (1=Primary (here as less than 8-9 grade of school completed, 2=Secondary (general obligatory education; high school; initial vocational education and training), 3=Tertiary (university, institute, academy + postgraduate)²⁸)
- Marital status. Dummy: 0 (base) = “never married”, "divorced and not remarried", "widower (widow)", 1 "in a registered marriage" "living together, not registered"
- Number of children under the school age (number of children of age 0 - 6 in household)
- Settlement status, level of urbanization: (3=urban (town, city); 2="PGT", urban village; 1=rural)
- Time dummies

3) Attrition and missing values analysis

It should be noted that RLMS dataset is not a truly panel as it was designed, primarily, as a cross-sectional repeating survey. High level of attrition is reported (which can happen both,

²⁶ Note, that among employed there are some percentage of pensioners as Retirement doesn’t mean an obligatory renunciation of working (such a law existed from 1997 till 2000, permitting to a person of pension age to chose among working and receiving pension). Moreover, pension age population is showed to be rather economically active. Variable is included to compensate for declining probability to be economically active with age. Generally, retirement age for women is 55 years old, for men is 60 years old.

²⁷ List of the geographic regions can be found in the appendix.

²⁸ If answer to a question was missing, it was replaced by the highest educational level reached by the previous round (starting for the round XIII back to round VIII).

through moving to another location and/or non-responding/participation decision). To maintain sample size new entrants are added up each round. Consequently, extraction of a balanced panel of individuals over several time points leads to significant shrinkage of the sample in consideration. In such settings attrition, if being non-random, could create serious problems for the statistical inference as sample extracted would not represent accurately the characteristics of underlying population any more.

Tables 5.1 below give an idea about a number of individual respondents in each of the five rounds covering 2000-2004, number of answers to the question of interest (satisfaction level) and number of individual observations containing information about all the explanatory variables included into analysis²⁹. As it can be seen, there is just a quite small number of missing answers to the satisfaction question (on average, 0.5% per round). At the same time, each year there are more than 1300 individuals, with this number increasing from round to round, that are not presented in the panel a year after (dropped out) and hence replaced with new entrants.

Table 5.1. Number of observations by rounds, age of respondents is 15 and older

Phase II, Round number	Year of survey ³⁰	Adults in a sample	Number of answers to general satisfaction question	Individuals, with no missing values in any of explanatory variables
Round IX	2000	8249	8179	6300
Round X	2001	9152	9095	6959
Round XI	2002	9492	9473	7456
Round XII	2003	9626	9593	7222
Round XIII	2004	9651	9606	7475
Total, R IX-XII	---	46170	45946	35412

Pooled sample of the five rounds gives about 46000 individuals answered the satisfaction question, and of being 15 - 72 years old. As Stata performs in the 'complete case analysis' style, actual number of observations in estimation is smaller by around 23% due to list-wise deletion of cases with missing values (this number is reported in the last column of the table 5.1). Pooling these complete observations together we get an unbalanced sample of 35412 observations with average 'individual participation rate' of 2.7. It is worth noting that this sample contains also

²⁹ Before these numbers were calculated, some cleaning procedures were applied to the merged rounds of the dataset: checking for entering data accuracy, range of variables' values in range prescribed by questioners; consistency of two main demographic characteristics (year of birth and gender) across the rounds, with the 'most probable year of birth' variables constructed. 5 observations which had gender changes over the period were dropped.

³⁰ For all the rounds data were collected in a period of September-December.

those people who participated in the survey at least once. Only about one fourth of these individuals is presented in all the five rounds considered. Balanced sample without missing values in any explanatory variable is composed of 2352 individuals (which gives 11760 observations for the five rounds).

Let us describe in more details the two phenomena mentioned above, namely, attrition and presence of some missing information for variables in the dataset. In both cases, nonrandom pattern of the drop outs/ missing data can lead to biased (and less efficient) estimation, potentially influencing conclusions consistency.

Drops out

Table 1 of the Appendix 4 contains information about means (and frequencies) of the main variables in the consideration for all the cross-sectional samples and sub-samples formed by the individuals dropped out of the sample one round after.

Individuals *dropped out* each round are, on average, happier (but farer from the crisis of 1998 lesser is the difference); 3.5 years younger than average individual in the original sample; they are healthier (which could be an age related issue, together with some other characteristics that have rather stable patterns of changes over the life cycle). It is more likely to drop out for men, habitants of urbanized areas (towns/cities); those with higher level of education. All these factors are (in common sense) related to the higher actual per capita household expenditures of the drops out, although they have more or less the same subjective evaluation of their economic rank in income distribution as those in the original sample. At the same time, share of employed is lower among those leaving the survey. Lower participation rate in the drops out subsample, as well as higher share of unemployed respective to the original sample, could be one of the reasons for lower desire to participate in a survey.

Missing values

At it was already noted, when estimating some model, apart of the drops out, we have one other group of individuals excluded from the sample. It happens not due their decision not to participate or move out, but due to estimation procedure itself requiring *no missing values in any of the variables in consideration*³¹ (complete case analysis style, used by Stata software).

³¹ In case when regression coefficients do not varying across groups, based on different sample characteristics, dropping the cases with missing values will not introduce bias in estimation results (property of robustness of predictor variables to Non Missing At Random assumption). But this would be true only in a case when the model of interest is correctly specified with no interaction between the predictor variables omitted. At the same time,

High percentage of missing values in a variable cast doubts about its representativeness. In fact, the most wide-spread receipt for the large samples is to drop out the cases that contain missing values if such missings' percentage is not much higher than 5%. In our case, the most "suffering" variables (total household expenditures, urbanization level, family composition and economic rank of the household) have, in fact no more than 6% of missing values included. The only 'troublesome' variable is that one coding aspirations for a next year's financial situation of a family (it includes more that 16% of missing values). These missing values cannot be replaced by some previous values and one of the possible solutions would be to fill the gaps with help of missing values handling procedures³². As a procedure, chosen for missing data handling could significantly affect the estimation results, this issue is not addressed in the present paper and list-wise deletion of cases is accepted further.

If MCAR (Missing Completely At Random) assumption is satisfied³³ than missing values contained in variables do not create inferential problems. Otherwise, biased estimates on the reduced sample, or even reverse direction of effects is possible. In our case Little's test, which has as null hypothesis of MCAR in the data, is rejected for both, pooled³⁴ and round by round tests. In real life settings, subgroups of people with different characteristics may differs in probability of answering groups of questions, so MAR (Missing At Random) assumption is usually implied³⁵. In such case patterns of missing values are supposed to be conditionally random. Hence exists a group of variables that enable "prediction" of whether person will answer or not, but not the level of variable if wouldn't be missing (for example, reach are less disposed to reveal their incomes). Separate variance t-tests allow to asses the plausibility of MAR assumption. In essence, two groups of observations are defined, with and without missing values for some variable. Difference between means of these two groups is tested for series of variables of interest. Characteristics of main variables in the analysis depending on whether there are missing values presented in one of two variables with the highest percentage of missing values, aspirations and total household expenditures are considered. If aspiration variable contains missing values, subsample formed by those with missing values in it, have significantly higher number of household members and, which could be a related issue, higher total

³² In case of high share of missing values in a dataset, usage of some of the imputation procedure is recommended. There exists a branch of statistical methods allowing for imputing of some values instead of the missing values (e.g. Multiple Imputation with maximum likelihood estimation ecc). In this paper they are not implemented (apart of the lagged values of the educational variable), as such methods exploitation can also lead to biased estimates. For some ideas about missing values handling/treatment, see for example von Hippel (2004)

³³ Missing Values Analysis Modul of SPSS v.13 contains *Little's MCAR test*, which is a chi-square test for missing completely at random. If H0 is not rejected, missing data can be deleted, in the opposite case, imputing is recommended.

³⁴ Chi-Square (178) = 1543,121, Sig. = ,000

³⁵ SPSS v.13 also allows for separate variance t-tests to asses MAR assumption validity

expenditures. At the same time they have lower subjective economic rank evaluation than those who reported their aspirations.

If subdivision is based on the missing values presence in household total expenditures variable, than ‘missing values group’ is formed by significantly happier and younger individuals, who have reported higher economic rank and better expectations than others in the sample.

It is also noticeable the higher percentage of missing expenditure values among men and those living in urbanized areas. At the same time, expenditures are not presented more often among people with lower levels of education as well as the answer to the aspirations question.

Original sample vs. sub-samples in estimation

Given in mind attrition and missing values problem, let us compare characteristics of the original representative sample and sub-samples that could be exploited for model estimation. There will be checked whether there are some discrepancies in pooled and separate rounds sub-samples.

There are several classical tests³⁶ used for assessing the representativeness of a sample. General idea is to test (for set of the variables chosen for model estimation) whether

H₀: Sub-sample matches characteristics the original sample for a variable

H₁: Otherwise

Table 2 of Appendix 4 summarizes the test findings for a list of variables which are planned to be included into the model. Having in mind the descriptive characteristics of the individuals groups excluded from the analysis through non participation or incomplete information, it is not surprisingly that shifting the analysis from the original sample to the unbalanced (complete cases) and than to balanced (complete cases) one can notice that the group in consideration is becoming less happy, older and contains higher proportion of females.

It can be seen also, that shifting from the original dataset to a sub-sample containing no missing values in explanatory variables (UP, unbalanced panel) preserves the main characteristics of the original dataset. The main differences between these two datasets is in significantly higher average economic self perceived rank of individuals and lower proportion of economically inactive sample members, with increased “employment rate”. Complete cases subset (UP) also consist of households with slightly higher number of household members. Average age of the unbalanced panel is also slightly higher although insignificantly. It is worth noting that cross-sectional comparisons (round by round) show that an average sample member herself is ageing.

³⁶ Depending on the nature of variable in consideration:

- Continuous variable: mean comparison t-test
- Categorical variable: chi-square goodness of fit test (Pearson)
- Proportions: tests on the equality of proportions using large-sample statistics; Z-statistics

As balanced panel analysis leads to implication of somehow simpler estimation procedures, let us check for discrepancies between the original dataset and BP data. One can see that only the educational levels composition and the mean of subjective economic situation of a family is not significantly different between the original and balanced panel sets of data. All the other variables have the means that are significantly different.

Having in mind that the pooled data comparison assumes averaging over a period in consideration, and could lead to a bias, we also checked the dataset characteristics on the round by round basis also. Comparison of the original dataset with unbalanced panel sub-set gives similar results as for the pooled dataset. The main concern is, again, about the labour market statuses disproportion. Round by round comparisons between the original dataset and BP show significant discrepancies between two datasets, as it was in a case of pooled samples. Moreover, characteristics of two samples are less similar if moving from the moment of the crisis (it could be an issue of the fact that balanced subsample contains, first of all, those individuals who survived the attrition, which is showed to be not a random event).

Resuming the facts discussed above, one should keep in mind the discrepancies found between the original sample and balanced/unbalanced panels in estimation. Exploiting the balanced subsample of the original dataset would be a good idea in simpler estimation sense, but BP extracted has significant differences from the original dataset, having a distorted sample composition for the key socio-demographic and economic characteristics (which can be partially, a consequence of the higher proportion of females and older average age of the population in the BP).

Complete cases subset comparison with the original dataset, round by round, reveals greater share of employed and lower share of out of labour force population in the sub-sample, while preserving other characteristics. It also argues for an idea of separate models estimation for sub-groups of individuals based on labour market participation status.

6. Estimation results

1) Main findings: basic cross-section regressions

At the beginning, multiple and ordered logit regressions can be estimated on the cross-sectional (round by round)³⁷ and pooled samples. In such a way a question of the coefficients stability over the five-year time period is to be addressed, as importance of different factors could change while transformation period were proceeding. Marginal effects calculated after ordered logit models

³⁷ Tables in Appendix 5 contain some estimation results: coefficients, and as an example, standardized coefficients, Odds Ratio, marginal effects for 2002

are mirroring the direction of factor effects on satisfaction level formed by linear models. This would be one of the reasons to prefer linear models further as they are simpler in estimation and interpretation.

When talking about the relative significance of different groups of factors, one could notice, that in case of inclusion of household economic situation variables only (logarithms of the total household expenditures and household size) Ramsey RESET test doesn't allow to reject the null hypothesis of no omitted variables only for the first two years (2000, 2001)³⁸. An attempt to explain happiness level by socio-demographic variables only (gender, age demeaned and its square, self-assessed health level, educational level, marriage, little children under the school age, level of urbanization of the settlement³⁹) leads to non rejection of the null hypothesis of no omitted variables, coming to the same conclusion by the specification link test also.

Inclusion of all the list of explanatory variable gives H1 RESET result for linear regression model, while the null of correct specification is not rejected by the specification link test on round by round models but not for the pooled sample on 5% significance level (this result holds for the binary logit also). This could be one more argument against a pooled sample based model.

Some modifications of the covariates list were performed: as collinearity diagnostics reflected no potential problems for all the variables in explanatory list but very high level of the Variance Inflation Factors for age and age squared, age variable were centered (demeaned by the mean age of the group in consideration, 40 y.o). Urbanization level variable was collapsed to binary indicator as influence of the "PGT" (urban village) settlement status was unclear and generally insignificant.

Some findings based on cross-sectional regressions are the following.

Household economic conditions. Logarithm of total household expenditures was included together with the logarithm of household members' number to account for decreasing marginal utility of money for happiness level evaluation and economy of scale in households. As it was expected, increase of (logarithm of) total monthly expenditures of a household conveys a positive significant effect on satisfaction evaluation (also, household size increase implies sharing of the same income by greater number of household members, which is reflected in negative coefficient before the logarithm of corresponding variable).

³⁸ While specification link test endorse the set of variables chosen as the covariates.

³⁹ There were the following covariates list modification: Collinearity diagnostics results lead to the necessity of transformation of the age and age squared variables. Age is demeaned (mean age of the group in consideration is 40 y.o). Variance inflation factors of all the other variables don't exceed the value of 3.

Noticeable, that magnitude of the coefficient before expenditures-related variable, is strongly significant, growing for the OLS specification over 2000-2003 (from 0.168 up till 0.198) and than drops in 2004 (coefficient of 0.148). Effect of the (log of) household size was significantly negative for all the years, with a slight tendency to decrease.

Comparison categories. Subjective evaluation of the family income/economic rank, always has a positive and highly significant coefficient (which is around 0.2 for all the years in OLS specification, and decreasing for 2000-2003 for binary and ordered logit models) as well as, again subjective, indicator of next year economic situation indicator (this coefficient is rather stable, around 0.25 for OLS, meaning an average increase of satisfaction level by one point if individual hopes that her household economic conditions would improve next year. Coefficients are again decreasing for both logit models). Talk about relative importance of factors for the general satisfaction evaluation, for example, for the year of 2002, subjective economic rank and expectations stay (for linear and ordered logit models) more important than logarithm of the household, while “order of importance” among them is changing if rely on standardized coefficients, arguing for greater immediate importance of the subjective economic rank.

If added to the model, relative expenditure poverty indicator (ratio of a household income to a mean total expenditures of all the households with the same size in a particular year) is significantly positive in 2000 and than becomes negative (although not always significant). The latter variable inclusion makes the earlier years (2000-2002) expenditure coefficient smaller and than greater for two other years, standard errors are slightly higher. Other coefficients remain practically unchanged. Even if the VIF coefficient for this variable is not exceeding 2, pairwise correlation with logarithm of the total expenditures is rather high (of the magnitude of 0.60), which possibly could cause some collinearity related issues.

Labour market characteristics. “Out of Labour force” category is chosen as a base. Than “unemployment” status has significantly negative effect (which its magnitude changing from -0.25 up to -.39, for OLS. Employment effect is negative for the 2000-2002 and than becomes positive although not significant for all the rounds but 2003. (Proxy) or a regional (38 regions) unemployment rate has always negative effect significant for all the years but 2002 and 2004.

Socio-demographic and family characteristics. Gender effect on satisfaction level is negative and not significant only for the last year. At the same time, health effect is generally increasing over the five years (OLS: from 0.15 up till 0.32 for one-point increase of subjective health evaluation; for binary and ordinary logit models the coefficient increase is even more pronounced from 0.31/0.36 up till 0.69). U-shape of relation between satisfaction and age can be noted, with both the coefficients, before age(centered) and age-squared, being significant. Educational level (being primarily coded as “secondary”, “tertiary” and “less than secondary” as the base group

had negative insignificant coefficients for secondary and insignificantly positive for the tertiary level) When collapsed to the binary category, inclusion of the higher education variable only gives significant positive effect on life satisfaction. Marriage, official or de facto, namely being “partnership” (with “not married” as a base category) have positive (and generally increasing over the period) significant effect on happiness when separate round estimates are considered. High level of urbanization of the settlement where the household lives, namely, settlement being town or city, have strong positive significant effect on satisfaction level.

Now recalling the difference in happiness levels found for employed and “others” (including unemployed and out of labour force) groups it is worth to investigate whether the coefficients are different for these two subgroups. In case of non-linear, e.g. binary logit model, direct comparison of the coefficients estimates on two subsamples is not possible due the fact that difference can be induced the difference in unobserved heterogeneity between genders. Than tests developed by Hoetker (2004; 2007) basing on the approach proposed by Allison (1999) should be applied. Pooled sample OLS or logit model demonstrate better fit for the “others” category. In both cases test for equal variation is not rejection the null hypothesis, while LR test states that the coefficients for estimates are not the same across groups. Round by round results analysis provides us with the following observations: coefficient before a dummy coding gender (female) is negative but in most cases insignificant for ‘others’ group , as well as the coefficient before dummy coding higher education level diploma received is significantly positive only for employed subsample. Ratio of the coefficients between groups (“others” to employed) is more than unity for logarithm of total expenditures (2000-2001) and health status (2000-2002). As hypothesis of equal residual variation is rejected only for 2001, standard Wald chi-square test for testing the differences between coefficients of the two groups is applied. Surprisingly, no difference in strength of the economical factors influence is found. Coefficient before health is appeared to be significantly different for 2001-2002 (OLS and binary logit specification) being more than 1.5 times higher for the ‘others’ group. Also for the 2004, significantly different are the coefficients before logarithm of the household size (for employed groups increase of the household size is more detrimental).

2) Main findings: Fixed and Random Effects panel models

Several econometric specifications

Several econometric specification estimates, based on panel nature of the data are presented in the table 6.1 below. Dichotomization of the dependent variable is done assuming “neutral” satisfaction point to be the threshold (this definition considers as happy those who have

satisfaction level higher than average in pooled dataset). In case of FE ordered logit, an approximation procedure is exploited.

Table 6.1 several econometric specifications of the model

	Xtreg, re	Xtreg, fe	Xtlogit, re	Xtlogit, fe	RE ordered probit	FE, ordered logit ⁴⁰
Dependent variable:	satlif0	satlif0	satlif01 0-1, 2-4	satlif01 0-1, 2-4	satlif0	satbin_ol
Economic situation						
Log(total expenditures)	0.143*** (0.008)	0.066*** (0.011)	0.333*** (0.022)	0.155*** (0.033)	0.190*** (0.009)	0.178*** (0.027)
Log(hh size)	-0.109*** (0.016)	0.027 (0.035)	-0.238*** (0.045)	0.075 (0.101)	-0.139*** (0.020)	0.039 (0.083)
Comparison indicators						
Subj. econ. rank	0.186*** (0.004)	0.151*** (0.006)	0.426*** (0.012)	0.315*** (0.016)	0.249*** (0.005)	0.299*** (0.013)
econ. aspirations	0.229*** (0.007)	0.194*** (0.009)	0.520*** (0.021)	0.449*** (0.028)	0.315*** (0.009)	0.435*** (0.023)
Labour market characteristics						
Unemployment	-0.262*** (0.026)	-0.144*** (0.032)	-0.673*** (0.078)	-0.338*** (0.099)	-0.352*** (0.034)	-0.238*** (0.082)
Employment	0.029* (0.015)	0.062*** (0.023)	0.110** (0.044)	0.158** (0.068)	0.043** (0.020)	0.154*** (0.056)
Proxy: region unempl. rate	-1.916*** (0.352)	-0.537 (0.483)	-5.032*** (1.008)	-1.748 (1.395)	-2.636*** (0.450)	-2.115* (1.147)
Soc.-dem. controls						
Soc.-dem. controls	yes	yes	yes	yes	yes	yes
Time dummies	yes	yes	yes	no	yes	no
Constant	yes	yes	yes	no	no	no
Observations	35412	35412	35412	17843	35412	25298
Individuals	13321	13321	13321	4741	13321	
Pseudo R-squared				0.107		0.0912
R-squared		0.116				
r2_o	0.287	0.00256				
r2_w	0.113	0.116				
r2_b	0.367	0.0129				
rho	0.252	0.813	0.296			0.277

Notes: ***p<0.01, ** p<0.05, *p<0.1. Robust st.err. in parenthesis.

Sample: unbalanced panel, complete cases, participation at least twice

Dependent variable: *Satlif0*: general life satisfaction coded from 0 to 4; *Satlif01*, 0-1, 2-4: dichotomized happiness level (1 if *satlif0* ≥ 2, 0 otherwise); *Satbin_ol*: dichotomized happiness level (1 if *satlif0* > individual_mean(happiness))

Dummy variables, base categories: education level - “lower than secondary education”; marital status - base “never married”; year - 2000.

List of socio-demographic controls: *age_centered*, *age_centered*², *female*, *health*, *married*, *higher education*, *urban area*, *child 0 – 6 years old*.

⁴⁰ Approximation: FE conditional binary logit model is estimated for the dependent variable, which is equal to 1 if outcome is higher than individual mean of satisfaction levels reported during the period covered.

Income-expenditures categories. Logarithm of total household expenditures is highly significant and affects positively level of satisfaction. In all cases, if RE or FE specification is applied, respected coefficient drops by about a half in FE case. For example, increase of the total expenditures by 1000RUR (when the mean expenditure in a sample is about 6000RUR) would lead to nearly one unit increase of the satisfaction level reported (linear RE model), while for the FE specification happiness level augments by less than a half step).

Logarithm of household members' number is negative and significant for random effects panel specifications. For the FE specifications the coefficient is found to be insignificant, which is partially in concordance with the finding of Senik (2004) that Russian individuals are/were more concerned with the economic situation of the whole family but not of the individual one.

Economic comparison categories. Subjective economic evaluation of household position in income distribution has significant positive influence on happiness level in all specifications, as well as the aspiration index. From RE to FE specification coefficient for the subjective economic rank is lower for FE, at the same time, its relative importance with respect to the total expenditures rises. For the linear FE model, self-positioning on the 6th(of 9) step, namely, next after the mean-position) in economic/income distribution) would lead to nearly a unit happiness increase.

Labour market characteristics. With "out of labour force" as a base category, unemployment has negative significant coefficient in all specifications. Employment has positive and significant for all panel specifications effect on satisfaction. Proxy of geographic area unemployment rate has negative influence on satisfaction (insignificant for the FE linear and ordered logit specifications). So, when regional unemployment increases, individual has worse expectations of his/her own perspectives to find (another) job and it pushes his/her happiness down.

Socio-demographic factors used as controls: Good Health has strong positive (significant) effect for all specifications, as well as fact of living in partnership (married officially or de facto). Also positive effect of urbanization on the happiness level can be noted (possibly, through greater opportunities for both, employment and leisure time). Children under school age (under 7 years old) have negative effect although not always significant. Education level influence is found to be insignificant for the secondary education (with respect to the primary one), and positive influence of the tertiary level (significant for the RE specifications). Effect of age on life satisfaction was found to be of the 'U' shape for all the specification but binary FE logit, where it is just positive (leaving the coefficient before the age squared insignificant). Age about of 46 could be estimated to be a reverse point when age rising has positive influence on satisfaction. Average age of the sample in estimation is 40 years old, and age squared coefficient insignificance could be due to the fact that the right tail of the age curve is considered.

Gender effect can be evaluated only for the RE and pooled sample cases due its time-invariance. Being female has significantly negative effect for all the specifications, as it found in most international happiness studies.

Econometric specification choice

All econometric specifications produce significant estimation results. Coefficients cannot be compared directly among models, but their simple comparison, if signs and significance of the coefficients are considered as the criteria, reveals similar inference results. Tests show that panel random effects model is preferred to the pooled simple version⁴¹, and further comparisons argue for fixed effects specification⁴². Namely, Hausman test shows that the coefficients between RE and FE specifications differs systematically, meaning that or model is misspecified or/and these differences occurred due to violation of the uncorrelation assumption. Moreover, ρ , percentage of the total variance due to u_i in FE is very high attributing about 81% to unobserved individual heterogeneity, emphasizing an importance of panel-level variance component.

Believing the model to be correct, Random effects specification is to be abandoned for the fixed-effect version. Main disadvantage of FE style model is that the sample exploited for estimation shrinks drastically, leaving the number of individuals in consideration trice less that in the RE case. This fact also leads to increase of the coefficients' standard errors.

One more way to test Random Effects specification against the Fixed Effects one and get coefficient estimated for time invariant variables is given by so called "hybrid method" combining the virtues of both of the approaches. Time-varying predictors, but not dependent variable, are decomposed into within and between parts, and than included into the model, estimated as RE. Deviations from the individual mean regressors' coefficients are the same as those given by FE estimation methodology. When coefficients estimates are the same for two parts of time-varying regressors, RE specification can be accepted.

In linear case, tests show that FE specification is important for all the explanatory variables but for age and fact of living in partnership. And obviously, joint test confidently rejects the H0 of the coefficients' pairs identity ($\chi^2(13) = 473.43$; Prob > $\chi^2 = 0.0000$).

Time-invariant variables estimates received with this method could be cited. So being female significantly decreases satisfaction level but only by 0.05 points. Level of urbanization has positive impact by 0.08 points for "PGT" and 0.14 for towns.

⁴¹ Breusch and Pagan Lagrange-multiplier test for random effects : $\chi^2(1) = 2322.67$, P-value = 0.0000

⁴² Hausman test: $\chi^2(16)=298.71$, P-value = 0.0000

One should also remember to use standard robust errors of the coefficient estimates, as individuals composing sample are heterogeneous (it is also confirmed by the Wald test for groupwise heteroskedasticity in fixed effect regression model that strongly rejects the null hypothesis).

Let us estimate the linear FE model for the two subsamples of interest, employed and 'others' (non-employed). Results achieved are practically identical: for both groups, it could be found U-shaped age-effect, the significant positive role of the total household expenditures, of subjective household economic position and of increasing aspirations about the future; importance of health level and of the fact of living in partnership. Inclusion of interactions of economic and labour market characteristics with gender gives the only significant coefficient arguing for negative influence of increased household size for woman in group of non-employed.

Given in mind the year 2002 as a kind of breakpoint, when Russian economy was granted the status of the market one, we would expect to find some changes in explanatory factors influence before and after this threshold. Estimation of Fixed Effects linear and logit model on unbalanced subsample reveals two main differences between time-periods. Logarithm of household total expenditures becomes insignificant, while positive influence of being employed becomes significant (on 10% sig.level). These two facts supports an idea of economic situation and labour market stabilization reached by the end of 2002.

Further developments and caveats

(1) Endogeneity and state dependence issues

Both, Random and Fixed Effects estimators, assume that current level of the dependent variable doesn't influence the future values of any explanatory variable. When this property is violated by the data, coefficients' estimates will be biased, tests and statistical inference invalid (e.g., Hausman test of choice between RE or FE specifications).

Given in mind the transitional matrices cited in the Section 3 of stylized facts, one could also see one more possibility for the model development. Stability of the happiness levels reported over time and mobility mainly to conjunctive categories argue pro possible state dependence presence would suggest inclusion of the lagged dependent variable into the explanatory variables list. Such model could help in evaluation of habituation effect, and give an idea about increase of expenditures level (or in some other predicting variable) compensating for high levels of subjective well-being in previous period. This approach also gives rise to endogeneity problem.

Rather simple endogeneity test for Fixed Effects model is cited in Wooldridge (2002). It implies that leads of presumably endogenous variable will be insignificant if included into the original equation. In our case “aspirations about the future economic situation of the family” and “subjective economic rank” variables have highly insignificant for expectations and insignificant at 5% but significant at 1% sig. level for economic rank lead coefficients in the modeled equation. Yet Durbin-Wu-Hausman test performed on the round by round subsamples (in the augmented original model, the significant coefficient found before the predicted-residuals term from the auxiliary regression for expectation (and economic rank) variable on all the (assumably) exogenous regressors and instruments) is rejecting the hypothesis of the variables exogeneity. Instrumental Variables technique is to be exploited in such cases.

(2) Attrition issues

It was figured out in section 5 that characteristics of the people, moving out of the survey (through non-response or changing of dwelling the household lives in) does not look random. Attriters, on average, were reporting higher satisfaction levels before leaving the sample. This suggests a possible influence of the same group of factors on both, variable of interest, happiness, and fact of moving out of the survey. Attrition bias in such the case could create serious problems with the validity of inference. The second caveat is the incidental truncation problem, appearing through the presence of missing values (answers) to a range of questions (in our case, as it was ascertained above, the most suffering variables are expenditures and aspirations about the future).

Testing for sample selection bias (attrition being an absorbing state) can be performed⁴³ by including in the model a lag (or lead) of the selection indicator, which is equal to unity if dependent variable (and the set of explanatory ones) is observed in a particular period. In our case, if linear Fixed Effects model is chosen as a base and indicator of participation in the next round is included into the model, its coefficient of 0.126 is found to be significant according to the t-test (with cluster robust standard errors). Unfortunately, inclusion of the inverse Mill’s ratio evaluated from a probit selection equation into the original one does not insure consistency of the FE estimator and Instrumental Variables methods are again coming into the fore. The main problem than would be to find some characteristics influencing the fact of quitting the panel and at the same time, not influencing the variable of interest⁴⁴.

(3) Causality issues

⁴³ See, for example, Wooldridge (2002), Chapter 17

⁴⁴ As the candidate for instruments in this case characteristics of the interviewer personality could be exploited, although rarely available. As Van Berg, Lindeboom and Lopez (2007) found in their study interviewers’ identifiers could be the valid instruments while correcting for the selection basis.

In this paper it is assumed that individual satisfaction is (one-sidedly) affected by household economic situation and labour market individual position. At the same time, it should be noted that a risk of causality reverse exists. So it could be controversial whether it is elevated happy mood that “enhances” an ability to gain money, or it is income received that increases happiness. Similar argumentation is applicable to the labour market individual position and to some characteristics among the controls into the model (e.g. health⁴⁵). Moreover, it is assumed that labour market status and income are not completely mirroring each other (that is roughly the case in transitional periods). At the same time, it could be reasonable to surmise that the fact of employment (or unemployment) influences the satisfaction level reported both, directly and indirectly (through income gained). So in this case, the problem of mediating variable is faced. Such issues lead to some more complex models and not addressed in the present paper.

7. Conclusions

Given a transformational character of the period in consideration, aggravated by the financial crisis of 1998, it would be expected that household economic situation and labour market position characteristics to be shifted on the top in the list of factors affecting individual happiness level evaluation. Moreover, instability of the (economy) environment is supposed to give an extreme importance to inter-personal comparisons while process of the subjective well-being evaluation and, possibly, explain the discrepancies between macro and micro level happiness measures.

Dataset even if originally constructed as a repeated cross-section study, enables two styles of analysis: on round-by-round cross-section or panel samples. The latter one suffers not only of the missing values in some variables but of the sample attrition. Comparison of the original sample characteristics and those of the balanced panel extracted, has shown inappropriateness of the balanced panel for the analysis, as most of the sample characteristics become distorted due non-random features peculiar for attriters. That is why an unbalanced panel sample without missing values in any of explanatory variables is chosen for model estimation. Given the interest in labour market status impact evaluation, only potentially economically active individuals were left in the sample, reducing attention to those of the age from 15 to 72 years old.

Subjective well-being evaluation model were estimated further both, in linear and non-linear (binary and ordered logit) versions on cross-section and (unbalanced) panel samples. Estimation results showed a particularly important role of the household current and expected economic conditions of a household on individual happiness level (particularly, in a sense of highest

⁴⁵ Frey and Stutzer (2005) cite some more situations when reverse causality is possible in relation between happiness and some events: marriage, schooling, volunteering (pro-social behaviour).

increase of model fit measures when these variables are included into the model, and magnitude of the standardized coefficients). In concordance with the world tendency, unemployment status has negative detrimental effect on satisfaction level, while the fact of employment not always significantly positive, which agrees with the results found in the preliminary analysis (that average happiness level was not growing faster for employed subsample of population). Employed group and 'others' (non-employed) were expected to give the same importance to the explanatory factors while subjective well-being is evaluated. Comparisons between coefficients of the model estimated on sub-samples of employed (52% of men) and non-employed labour market categories (59.6% of women) reveals the particular importance attached by "others" to such factors as personal health and living in partnership (which is not so much surprising having in mind prevalence of women in the group). Employed, in their turn, pay more attention to the subjective evaluation of their family economic rank, while a ratio of the household expenditures to the mean expenditures among the households with the same size is never significant when panel dimension is exploited. Possibly, this happens because of inability of individuals to evaluate such index properly. When talking about an indicator of labour market situation (proxy of unemployment rate for geographical area), it is found to be negatively significant, but only for Random Effects specifications that were abandoned for the Fixed Effects version. Fixed Panel estimates are practically identical for two subsamples. Inclusion of gender interactions with variables of interest gives the only significantly negative coefficient for increase of household size (and hence, amount of house keeping and care activities) when non-working women are addressed.

Round-by-round estimation results for the whole sample were arguing that effect of some independent variables was changing over the time (e.g. coefficient by the logarithm of total household expenditures was growing up till the year 2003 and than dropped down). Such tendency could be linked with the fact of the market economy status achieved by Russia by the 2002. In fact, estimation of the Fixed Effects model for two different time periods 2000-2002 and 2003-2004 reports a slightly higher coefficient for the comparison categories for the second subsample (for both, linear and logit specifications), while the coefficient before the logarithm of the total household expenditures becomes insignificant. At the same time, for linear model, fact of employment becomes significantly (at 10% sig.level) positive, which could be one of the arguments towards harmonization of labour market and economic situation reached by the end of 2002.

Some further extension of the explanatory variables set is to be done, as model fit measures are not very high (about 0.10 for within-R2 for linear Fixed Effects and 0.12 for pseudo-R2 for logit case). Among factors to include could be some more detailed job characteristics, while one cannot be sure that it was not uncertainly and general depressive further looking attitude to be among the most detrimental influencing factor within first years after the crisis. Hence, one of the further model developments could be subjective well-being state-dependence investigation.

References

- Andrews, F.M. & Withey, S. (1976). Social indicators of well-being. *American perceptions of life-quality*. New York: Plenum Press.
- Blanchflower and Oswald (2004), Well-being over time in Britain and the USA, *Journal of Public Economics*, v. 88, 1359-86.
- Bogomolova, Tatyana and Vera Tapilina, Income Mobility in Russia in the mid-1990s, Economic Education and Research Consortium, Russian Economic Research Program, WP N. 99/11, 1999.
- Clark, A. and Oswald, A. (1994), Unhappiness and unemployment, *Economic Journal*, 104, 648-59.
- Alesina, Di Tella, MacCulloch (2001), Inequality and Happiness: Are Europeans and Americans different?, NBER WP 8198
- Easterlin, Richard A. (2005), Is there an "Iron Law of Happiness?", Institute of Economic Policy Research, University of Southern Carolina, Working Paper 05.8.
- Eggers Andrew, Clifford Gaddy, Carol Graham, (2006) Well-being and unemployment in Russia in the 1990s: Can society's suffering be individuals' solace?, *The Journal of Socio-Economics*, 35, 209 – 242.
- Ferrer-i-Carbonell, Ada and Paul Frijters, How Important is Methodology for the Estimates of the Determinants of Happiness?, *Economic Journal*, Vol. 114, No. 497, pp. 641-659, July 2004
- Ferrer-i-Carbonell, Ada and Bernard van Praag (2001), Poverty in the Russian Federation, *Journal of Happiness Studies* 2: 147–172.
- Frey, B. S. and A. Stutzer (2002): What can Economists Learn from Happiness Research? *Journal of Economic Literature*, (40), 402-435.
- Frey and Stutzer (2005), Happiness Research: State and Prospects, *Review of Social Economy*, Vol. LXII, No. 2.
- Frijters, Paul, Ingo Geishecker, John Haisken-DeNew and Michael A. Shields, Can the large swings in Russian life satisfaction be explained by ups and downs in real incomes?, *Scandinavian Journal of Economics*, 2006.
- Frijters, Paul, Ingo Geishecker, John Haisken-DeNew and Michael A. Shields (year???), Income and life satisfaction in post-transition Russia: a new empirical methodology for panel data,
- Graham, C., Eggers, A. and Sukhtankar, S. (2004) Does happiness pay? An exploration based on panel data from Russia. *Journal of Economic Behavior and Organization*, 55, 319-342.
- Jones, Andrew M. and Schurer, Stefanie (May 2007), "How Does Heterogeneity Shape the Socioeconomic Gradient in Health Satisfaction?". Ruhr Economic Paper No. 8 Available at SSRN: <http://ssrn.com/abstract=989560>.

- Kahneman, Daniel and Alan B. Krueger, Developments in the Measurement of Subjective Well-being, *Journal of Economic Perspectives*, Vol. 20. N. 1, 2006, pp. 3-24.
- KALUGINA E. and NAJMAN B. (2003), Travail et pauvreté en Russie: évaluations objectives et perceptions subjectives, *Economie et Statistique* N° 367
- Kim, B., (2002) Russian household participation in the informal economy: evidence from VTsIOM Data, *Economics of transition*, vol. 10, n. 3, pp. 685-713.
- Kislitsyna, Olga, Income Inequality in Russia during Transition. How can it be explained?, Moscow: *EERC*, 2003, pp 1-51
- Layard, Richard, Happiness. Lessons from a new science, Penguin Books Ltd., 2005
- Ng (1997), A case for happiness, cardinalism and interpersonal comparability, *The Economic Journal*, 107, 1848 - 1858.
- Ravallion, M. and Loskin, M. (2001) Identifying welfare effects from subjective questions, *Economica*, 68, 335-357.
- Ravallion, Martin, Michael Lokshin, (2002) Self-rated economic welfare in Russia, *European Economic Review*, 46, 1453-1473.
- Rose, Richard and Ian McAllister, Is money the measure of welfare in Russia?, *Review of Income and Wealth*, Series 42, Number 1, March 1996, pp. 75-90.
- Schaffer, M.E., 2007. xtivreg2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models. <http://ideas.repec.org/c/boc/bocode/s456501.html>
- Surgen, Robert (1993), Welfare, Resources, and Capabilities: A Review of Inequality Reexamined by Amartya Sen, *Journal of Economic Literature*, Vol. 31, No. 4, 1947-1962.
- Sen A. (1985), *Commodities and Capabilities*, Amsterdam, North-Holland
- Senik, Claudia, When information dominates comparison. Learning from Russian subjective panel data, *Journal of Public Economics*, 2004, 88, 2099-2133
- Tucker, K., D. J. Ozer, S. Lyubomirsky and J. K. Boehm, (2006), Testing for measurement invariance in the satisfaction with life scale: a comparison of Russians and North Americans, *Social Indicators Research*, 78, 341-36
- Van den Berg, Gerard J., Maarten Lindeboom, Marta Lopes (2007), Interviewer identities as valid instruments for selective panel survey attrition – an evaluation with matched survey-register data, report at the BHPS(British Household Panel Survey) -2007 Conference, University of Essex, Colchester, UK, 5 – 7 July 2007.
- Van Praag, Bernard M.S. (1971), The welfare function of income in Belgium: an empirical investigation, *European Economic Review*, vol. 2, 337-369.
- Van Praag, Bernard M.S. (2007), Perspectives from the Happiness Literature and the Role of New instruments for Policy Analysis, IZA DP No. 2568.
- Veenhoven, R., Are Russians as unhappy as they say they are? (2001), *Journal of Happiness Studies*, 2, pp. 111-136.
- Winkelmann L. and Winkelmann R. (1998) Why are the Unemployed So Unhappy? Evidence from Panel Data, *Economica*, 65, 1-15.
- Wooldridge (2002), Jeffrey, *Econometric Analysis of Cross Section and Panel Data*, The MIT press.

Appendices

A1. Life satisfaction studies for Russia (mainly, RLMS based)

Short summary of some papers studying life satisfaction in Russia are presented below.

Authors, year of publication	Sample, RLMS	Econometric model	Main findings
Ravallion and Lokshin (2001)	Balanced panel, 1994 and 1996	First-difference model of life satisfaction	<u>Income</u> (particularly, the household one) is an important determinant of subjective economic welfare; Permanent welfare loss is formed by even a transient period of <u>unemployment</u>
Graham et al. (2004)	1995 and 2000	Ordered logit model of life-satisfaction	Life satisfaction is U-shaped in age; positively associated with being male, higher education, good health, household equivalised <u>income</u> , <u>employment</u>
	Balanced panel, 1995 and 2000	First-differenced (2000-1995) ordered logit model	Significant effect on life-satisfaction: increased log <u>income</u> (+), getting divorced (-), leaving school (-)
Senik (2004)	1994 to 2000	Ordered logit model with controls for individual-specific averages	Life satisfaction vs. Income distribution. Importance of social-economic and demographic characteristics; unusual : positive influence of reference group income on life satisfaction (optimistic expectations) (Hirschman, 1973, “tunnel effect”)
		Linear fixed effects regression	
Eggers et al. (2006)	1995, 1996, 1998, 2000, 2001	Ordered logit/probit with lagged dependent variable; Fixed effects	Lowering of standards in bad economic situation/transition. High <u>regional unemployment rates</u> associated with higher life satisfaction.
Frijters et al. (2006)	1994 – 2001, 18-65 y.o.	Fixed-effect ordered logit ⁴⁶	Personality traits as important predictors of general satisfaction. Models estimated separately for males and females. Decomposition of the total changes in latent life satisfaction into changes in (family, health, job, income, area income), (year), (FE). Even in the analyzed period of changes, the parameter estimates are consistent with British and German panel data results. Effects on life-satisfaction: getting married(+), divorced(-), spouse died(-), young children(-), older children(+);

⁴⁶ Proposed by Ferrer-i-Carbonel and Frijters (2004): an extension of binary conditional fixed-effect logit model of Chamberlain (1980) with person-specific thresholds

			unemployed (-) (unusual: higher the local unemployment rate, more detrimental own unemployment) Richer area/urban context(+)
--	--	--	-----------------------------------------------------------------------------------------------------------------------------------------

Other related studies

Authors	Sample, RLMS	Model	Findings
Bogomolova and Tapilina (1999)	1994 – 1996		Explored how changes in certain socio-economic variables over time affect income mobility.
Beuran and Kalugina (2006)	1994 – 2003, those who have only one, formal or informal job	Ordered probit model for social exclusion (individual objectively or subjectively poor); ML estimation for 2 simultaneous equations (social exclusion and working in informal sector)	“Informal employment vs. welfare”. Average salaries are higher in informal sector; and worse working conditions and lack of social benefits doesn’t increase the probability of being excluded. (confirmed the result of Kalugina and Najman (2003) about the existence of negative correlation between poverty and working in informal sector in Russia)
Kislitsyna(2003)	1992, 1994, 1996, 1998, 2000		Changes in income structure of population, income inequality
Ferrer-i-Carbonell and van Praag (2001)	RUSSET panel dataset, 1997, 1998	Leyden approach	Welfare and well-being poverty concepts, objective and subjective measures. Official poverty lines are discussed

A2. Some economic indicators, 2000-2004

Figures, presented below describe the economic situation development in Russia in 2000-2004.

Source: State Committee for Statistics <http://www.gks.ru>

GDP and CPI dynamics

	2000	2001	2002	2003	2004
Real Gross Domestic Product in year-averaged market prices of 2000, bln rubles	7305,6	7677,6	8041,8	8632,7	9249,4
Indices of real volume of GDP (in per cent to the corresponding period of the previous year)	110,0	105,1	104,7	107,3	107,1
CPI, in per cent to 2000 level	100	119	136	153	171

Living standards of population

	2000	2002	2003	2004
Average per capita money incomes of population, monthly, RUR (thou. RUR before 2000)	2281	3947	5170	6410
Real disposable money incomes, percentage as of the previous year	112	111	115	110
Subsistence minimum level (average per capita): RUR per month	1210	1808	2112	2376
in percent to previous year	120 ^{*)}	121	117	113
Population with money incomes below subsistence minimum level: mln. persons	42.3	35.6	29.3	25.2
percentage of the total population	29	24.6	20.3	17.6
percentage as of the previous year	84.9 ^{*)}	89	82.3	86
Minimum wages (annual average), RUR	107.8	400	487.5	600
Real minimum wages, as percentage of the previous year	106.9	138.2	107.2	111

^{*)}By comparable methodology of subsistence minimum compilation.

Labour force⁴⁷

Economically active (15-72), including:	1998	2002	2003	2005
Employed	86,8%	91.9%	91.4%	92.4%
Unemployed	13.2%	8.1%	8.6%	7.6%

In 1992-1998, a number of economically active in Russian economy declined from 70.3 to 61.1% of the working age (15-72) population. In 1999 just 65.5% of w.a.p. were economically active; in 2001: 64.3% of w.a.p. and only than the activity rate started to grow again since 2002 due to increase of employment category.

It is also noted that economical activity declined for all the age groups during the period of 1992-2005. At the same time, after the crisis of 1998, a portion of employed among aged 55 and older was growing in contrast to the youngest group (younger than 24), and even became greater (e.g. 2004: 60-72 y.o. employed 16.1%; 2005: less than 20 y.o. employed 11%.)

⁴⁷ According to "Demoscope weekly", № 277-278, 19.02-04.03.2007, <http://demoscope.ru/weekly/2007/0277/barom01.php>, based on the information provided by the Central Statistical Agency

A3. Economic regions of Russia

The following eight Geographical Regions are suggested by the RLMS web-site documentation:

- 1) Metropolitan areas: Moscow and St. Petersburg
- 2) Northern and North Western
- 3) Central and Central Black-Earth
- 4) Volga-Vaytski and Volga Basin
- 5) North Caucasian
- 6) Ural
- 7) Western Siberian
- 8) Eastern Siberian and Far Eastern

This subdivision is mainly in accordance with the administrative division of 1995 that divided territory of Russia into 12 economic regions:

- 1) Centralny (Central area) (including Moscow and Moscow area)
- 2) Centralno-Tshernoziom (Central Black Earth area)
- 3) Vostotchno-Sibirsky (East-Siberian area)
- 4) Dalnevostotchny (Far-Eastern area)
- 5) Severny (Northern area)
- 6) Severno-Kavkazsky (Nor-Caucasian)
- 7) Severo-Zapadny (Nor-Western) (including Saint-Petersburg and Leningrad oblast)
- 8) Privolzhsky (Volga-area)
- 9) Uralsky (Ural-area)
- 10) Volgo-Viatsky (Volga-Vyatka area)
- 11) Zapadno-Sibirsky (West Siberian area)
- 12) Kaliningrad (not presented in the dataset considered)

A4. Attrition and Missing Values Analysis

Table 1. Means and frequencies of the variables.

Cross-section sample and those who drops out the next period

	2000	Dropped between 2000 and 2001	2001	Dropped between 2001 and 2002	2002	Dropped between 2002 and 2003	2003	Dropped between 2003 and 2004	2004
satifaction	1.37	1.61	1.59	1.66	1.88	1.91	1.80	1.83	1.94
woman	0.54	0.45	0.54	0.46	0.54	0.46	0.54	0.47	0.53
age	40.70	36.58	40.48	36.98	40.35	37.60	40.35	36.45	40.18
health	3.18	3.29	3.19	3.29	3.21	3.24	3.22	3.27	3.24
urban	2.38	2.64	2.37	2.66	2.37	2.70	2.38	2.58	2.42

<i>1. rural</i>	27.99	17.58	26.64	17.3	26.48	16.9	26.85	22.35	26.01
<i>2. pgt</i>	6.88	5.92	6.3	5.21	5.92	3.77	6.44	5.64	6.31
<i>3. town/city</i>	65.13	76.5	67.06	77.49	67.6	79.33	66.71	72.01	67.68
married	0.64	0.58	0.64	0.57	0.62	0.56	0.62	0.56	0.60
education	2.07	2.12	2.10	2.15	2.11	2.15	2.11	2.16	2.12
<i>1. Primary</i>	9.4	7.47	7.66	5.49	7	6.23	6.18	5.2	5.38
<i>2. Secondary</i>	74.76	74.4	74.93	75.43	75.24	74	76.05	76.02	76.56
<i>3. Tertiary</i>	15.84	18.13	17.41	19.08	17.76	19.77	17.77	18.79	18.06
Children 0 – 6y.o.	0.22	0.18	0.21	0.21	0.20	0.18	0.21	0.15	0.20
Hh size	3.52	3.23	3.51	3.34	3.54	3.27	3.54	3.35	3.49
Econ.rank	3.67	3.78	3.87	3.88	3.84	3.83	3.83	3.84	4.01
aspirations	2.98	3.04	3.15	3.23	3.19	3.32	3.19	3.34	3.17
Total p.c. expenditures, prices of 2000	1795.24	2376.48	2025.00	2429.51	2087.72	2846.00	2361.46	2997.57	2469.31
Total expenditures, prices of 2000	5979.53	6937.54	6678.03	7050.71	6741.25	8952.80	7750.68	8628.32	7881.16
Labour status	1.23	1.15	1.22	1.18	1.23	1.14	1.24	1.20	1.20
<i>1. out of LF</i>	37.48	41.56	37.38	39.01	37.29	40.92	36.21	37.57	35.97
<i>2. unemployed</i>	4.95	5.01	4.62	5.84	4.42	6.03	3.91	4.39	4.16
<i>3. employed</i>	57.57	53.43	58	55.15	58.28	53.05	59.88	58.05	59.88
Dropped individuals		1138		1233		1310		1504	

Notes: individual weights applied. Sample: 15 – 72 years old.

Table 2 Pooled original, UP and BP without missing

	N 0	Mean 0	Mean UP	t1: mean0=meanUP	P, t1	Mean BP	t2: mean0=meanBP	P, t2
satlif0	45946	1.749	1.747	-0.179	0.858	1.648	-8.801	0.000
woman	46170	0.558	0.554	-1.255	0.210	0.585	5.237	0.000
age	46170	40.252	40.521	2.343	0.019	43.062	18.000	0.000
health	46036	3.198	3.203	0.949	0.342	3.161	-5.326	0.000
urb1	44983	0.268	0.274	2.007	0.045	0.323	11.476	0.000
urb2	44983	0.064	0.059	-2.790	0.005	0.051	-5.197	0.000
urb3	44983	0.669	0.667	-0.477	0.633	0.626	-8.590	0.000
married	46059	0.640	0.643	1.046	0.296	0.711	15.042	0.000
edu1	46000	0.070	0.071	0.462	0.644	0.069	-0.589	0.556
edu2	46000	0.755	0.754	-0.345	0.730	0.750	-1.297	0.195
edu3	46000	0.174	0.174	0.079	0.937	0.181	1.849	0.064
child06	46170	0.235	0.247	3.375	0.001	0.237	0.432	0.666
hhmemb	44945	3.411	3.369	-3.733	0.000	3.332	-4.887	0.000
Out of LF	46170	0.368	0.335	-9.813	0.000	0.314	-11.266	0.000
Unemployed	46170	0.044	0.046	1.296	0.195	0.038	-2.780	0.005
Employed	46170	0.588	0.619	9.023	0.000	0.648	12.092	0.000
econrk	45519	3.842	3.875	3.175	0.001	3.875	2.225	0.026
aspir	38153	3.151	3.148	-0.474	0.636	3.090	-6.896	0.000
texp2000	43737	7100.023	6971.010	-1.828	0.068	6213.364	-10.271	0.000

texp00pc	43737	2262.657	2251.435	-0.449	0.653	2022.760	-8.037	0.000
unem_geo	46170	0.073	0.074	1.778	0.075	0.074	5.985	0.000

Sample 0: original sample, 15 – 72 years old.

Sample 2 (UP): 35419 obs (5 rounds, unbalanced panel, complete cases; actually in estimation)

Sample 3 (BP): 11765 obs (balanced sample, no missing)

Note: categorical variables are substituted by groups of dummies, hence their means represent a share of individual category.

A5. Regressions

(Round-by-round)

Ordinary regressions and Ordered logit regressions, round by round, weighted

vl	2000		2001		2002		2003		2004	
COEFFICIENT	reg	logit	reg	logit	reg	logit	reg	logit	reg	logit
ltexp00	0.168** (0.018)	0.282** (0.044)	0.139** (0.018)	0.205** (0.044)	0.154** (0.019)	0.245** (0.043)	0.198** (0.019)	0.348** (0.046)	0.148** (0.017)	0.302** (0.047)
lhhmemb	-0.202** (0.034)	-0.249** (0.085)	-0.131** (0.033)	-0.101 (0.076)	-0.103** (0.034)	-0.181* (0.079)	-0.145** (0.034)	-0.220** (0.081)	-0.151** (0.032)	-0.325** (0.082)
econrk	0.211** (0.010)	0.381** (0.024)	0.202** (0.011)	0.374** (0.025)	0.206** (0.011)	0.377** (0.025)	0.182** (0.011)	0.332** (0.026)	0.215** (0.010)	0.409** (0.026)
aspir	0.210** (0.016)	0.395** (0.039)	0.235** (0.018)	0.412** (0.043)	0.247** (0.018)	0.436** (0.043)	0.247** (0.020)	0.423** (0.046)	0.247** (0.018)	0.468** (0.046)
unemploy	-0.251** (0.064)	-0.400* (0.161)	-0.386** (0.068)	-0.808** (0.162)	-0.300** (0.069)	-0.625** (0.158)	-0.280** (0.074)	-0.536** (0.170)	-0.375** (0.070)	-0.881** (0.172)
employed	-0.011 (0.038)	0.054 (0.089)	-0.061 (0.037)	-0.129 (0.084)	-0.062 (0.036)	-0.103 (0.083)	0.098** (0.038)	0.256** (0.086)	0.012 (0.034)	0.048 (0.088)
unem_geo	-4.600** (1.219)	-7.601** (2.937)	-1.400* (0.627)	-1.691 (1.439)	-1.722 (1.212)	-3.024 (2.757)	-3.290** (0.711)	-6.371** (1.647)	-0.134 (1.236)	-4.117 (3.167)
woman	-0.079** (0.028)	-0.162* (0.066)	-0.073** (0.028)	-0.113 (0.063)	-0.086** (0.027)	-0.203** (0.063)	-0.087** (0.028)	-0.158* (0.065)	-0.042 (0.026)	-0.069 (0.066)
agedem	-0.004** (0.001)	-0.007** (0.003)	-0.004** (0.001)	-0.010** (0.003)	-0.003* (0.001)	-0.009** (0.003)	-0.006** (0.001)	-0.014** (0.003)	-0.004** (0.001)	-0.011** (0.003)
agedem2	0.000** (0.000)	0.001** (0.000)	0.001** (0.000)	0.001** (0.000)	0.001** (0.000)	0.001** (0.000)	0.001** (0.000)	0.001** (0.000)	0.001** (0.000)	0.001** (0.000)
health	0.146** (0.022)	0.362** (0.053)	0.218** (0.023)	0.447** (0.052)	0.241** (0.023)	0.414** (0.052)	0.196** (0.023)	0.378** (0.054)	0.321** (0.022)	0.680** (0.057)
edu3	0.113** (0.039)	0.297** (0.087)	0.074* (0.037)	0.196* (0.084)	0.089* (0.035)	0.237** (0.084)	0.142** (0.038)	0.377** (0.089)	0.115** (0.034)	0.336** (0.088)
married	0.082* (0.033)	0.120 (0.083)	0.122** (0.033)	0.171* (0.077)	0.198** (0.033)	0.305** (0.076)	0.164** (0.034)	0.257** (0.078)	0.174** (0.031)	0.309** (0.079)
urban	0.082** (0.016)	0.111** (0.039)	0.124** (0.015)	0.235** (0.036)	0.082** (0.016)	0.158** (0.036)	0.085** (0.016)	0.217** (0.037)	0.069** (0.015)	0.132** (0.037)
Constant	-1.551** (0.178)	-6.185** (0.448)	-1.965** (0.163)	-6.606** (0.421)	-1.939** (0.181)	-6.088** (0.434)	-2.160** (0.167)	-7.065** (0.439)	-2.261** (0.176)	-7.466** (0.491)
Observations	5174	5174	5380	5380	5558	5558	5230	5230	5367	5367
R-squared	0.26		0.25		0.26		0.27		0.33	
R2_p		0.147		0.142		0.144		0.160		0.190

Notes: * significant at 5%; ** significant at 1%

Robust standard errors in parentheses.

Standardized coefficients and OR (ordered logit estimation), 2002 (cross-sectional weighted)

satlif0	Coef. Linear reg.	Std. Err. Linear reg.	Standardized coeff, Linear reg	Odds Ratio, Ologit	Std. Err., ologit	bStdXY, ologit
---------	-------------------	-----------------------	--------------------------------	--------------------	-------------------	----------------

ltxp00	.1535525	.0186162	.1165786	1.350627	0.0492743	0.1207
lhmemb	-.1032578	.0343895	-.0433381	0.834341	0.054506	-0.0402
econrk	.2058938	.0105179	.2665825	1.485502	0.0312535	0.2711
aspir	.2468964	.017831	.1801157	1.624195	0.0594848	0.1872
unemployed	-.3002874	.0687111	-.0559689	0.5690045	0.0764835	-0.0556
employed	-.0624962	.0360448	-.0270461	0.9019923	0.0633485	-0.0236
unem_geo	1.721995	1.212028	-.0173715	0.0500931	0.1162767	-0.0160
woman	-.0863855	.0270309	-.0384044	0.8543978	0.0442632	-0.0370
agedem	-.0025404	.0011485	-.0355476	0.9954965	0.0021926	-0.0334
agedem2	.0006361	.0000694	.1431623	1.001202	0.0001373	0.1431
health	.2409316	.0228802	.1501115	1.607034	0.0738628	0.1564
edu3	.0892736	.0353626	.030229	1.165798	0.0776192	0.0275
married	.1984361	.0333223	.0859747	1.46079	0.0941389	0.0869
Urban (PGT or town)	.0824567	.0157472	.0647644	1.165638	0.0349805	0.0637

Note: bStdXY regression coefficients with both the x-variable (the independent variable) and the y-variable (the dependent variable) in standard deviations.

Marginal effects after ordered logit estimation, 2002 (cross-sectional weighted)

	Mfx (satlif=0)	Mfx (satlif=1)	Mfx (satlif=2)	Mfx (satlif=3)	Mfx (satlif=4)
satlif0					
ltxp00	-.0213074	-.0505707	.0082688	.0520065	.0116029
lhmemb	.0128392	.0304723	-.0049825	-.0313375	-.0069915
econrk	-.028055	-.0665854	.0108873	.0684759	.0152773
aspir	-.0343827	-.0816034	.0133429	.0839202	.018723
unemployed	.0497058	.0893813	-.0327695	-.0889707	-.017347
employed	.0072402	.0173598	-.0026711	-.0179016	-.0040274
unem_geo	.2122367	.5037193	-.0823626	-.5180207	-.1155727
woman (d)	.0111305	.0264573	-.0042521	-.0272378	-.0060979
agedem	.00032	.0007594	-.0001242	-.000781	-.0001742
agedem2	-.0000852	-.0002022	.0000331	.0002079	.0000464
health	-.0336297	-.0798162	.0130507	.0820823	.0183129
edu3 (d)	-.0104307	-.02584	.0031933	.0268742	.0062032
married (d)	-.0279551	-.0632082	.0124535	.0645852	.0141246
urban	-.0108653	-.0257875	.0042165	.0265196	.0059166

(unbalanced panel)

Fixed Effects models, coefficient estimates for two time-periods

	Linear	Linear	Logit	Logit
--	--------	--------	-------	-------

	FE, 2000- 2002	FE, 2003- 2004	FE, 2000- 2002	FE, 2003- 2004
Ln(tot.exp)	0.077*** (0.016)	-0.018 (0.024)	0.137*** (0.049)	0.034 (0.078)
Ln(hh_size)	0.078 (0.054)	0.155* (0.082)	0.327* (0.169)	0.671** (0.262)
Econ. rank	0.142*** (0.008)	0.148*** (0.012)	0.284*** (0.024)	0.359*** (0.038)
Expectations	0.177*** (0.013)	0.197*** (0.019)	0.385*** (0.042)	0.415*** (0.065)
unemployed	- 0.154*** (0.046)	-0.156** (0.068)	- 0.417*** (0.147)	-0.324 (0.229)
employed	0.046 (0.035)	0.085* (0.050)	0.127 (0.106)	0.257 (0.166)
Unem. rate.reg.	-0.139 (0.670)	-1.223 (0.967)	2.059 (2.140)	-3.900 (3.153)
Age_demeaned	0.178*** (0.010)	0.090*** (0.017)	0.482*** (0.032)	0.355*** (0.058)
Age_dem^2	0.001*** (0.000)	0.000 (0.001)	0.002** (0.001)	0.000 (0.002)
health	0.124*** (0.019)	0.115*** (0.027)	0.285*** (0.057)	0.276*** (0.091)
edu3	0.026 (0.100)	-0.033 (0.096)	0.138 (0.292)	-0.051 (0.391)
married	0.167*** (0.053)	0.118 (0.082)	0.410** (0.172)	0.189 (0.235)
child06	-0.080** (0.036)	-0.004 (0.057)	-0.234** (0.117)	0.282 (0.190)
Observations	20715	14697	7726	3238
Number of id	10478	9450	2913	1619
R-squared within	0.13	0.08		
Log-likelihood	-18077	-9944	-2397	-992.7
r2_pseudo			0.141	0.115

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1