Population and Economics

Population and Economics 9(3): 198–224
DOI 10.3897/popecon.9.e135768

RESEARCH ARTICLE

Factors in Intentions to Have a Third Child: What Machine Learning Analysis Reveals

Alla O. Makarentseva [1], Aleksandra Ya. Burdyak [1], Iaroslav I. Sheipak [2]

1 The Russian Presidential Academy of National Economy and Public Administration, Moscow, 119034, Russia

2 Moscow Institute of Physics and Technology (MIPT), Dolgoprudny, Moscow region, 141701, Russia

Received 28 August 2024 ♦ **Accepted** 27 December 2024 ♦ **Published** 24 October 2025

Citation: Makarentseva AO, Burdyak AYa, Sheipak IaI (2025) Factors in Intentions to Have a Third Child: What Machine Learning Analysis Reveals. Population and Economics 9(3):198-224. https://doi.org/10.3897/popecon.9.e135768

Abstract

An individual's reproductive trajectory can be represented as a chain of decisions about whether or not to have a child. Decision making is a time-consuming process and usually begins with the formation of intentions. Factors of intentions, as well as their subsequent implementation, differ for children of different orders. Third births are less frequently the focus of attention of researchers than first and second births. The aim of this work is to determine the factors of intentions to have a third child among men and women with two children. The empirical basis of the study is a population survey "People. Family. Society" 2023. We use machine learning methods (ensemble trees) and include in the analysis a very wide range of classifying features, consisting of demographic and socio-economic indicators, as well as value attitudes and assessments of subjective well-being. The peculiarities of the method are such that the presence and absence of a feature can have an asymmetric effect on the dependent variable.

The results of the study reveal a profile of conservative attitudes that determine higher chances of being included in the class of those intending to have a third child in the next three years: a negative attitude towards the admissibility of abortion, agreement with the statement that women would like to take care of the home and children, and moderate religiosity. In addition, factors that indicate the choice of the appropriate time for the birth of a child are important, namely the age of the youngest child from 3 to 6 years, positive expectations regarding improved housing conditions, as well as the desire for the spouse to have more children as part of the couple's agreed reproductive attitude. The class of those who express a desire to have more children, but do not express specific plans for having a child (this could be either a postponement or a complete refusal to realize the desire), is of particular interest in the context of the possibilities of demographic policy. The economic profile of factors for falling into this class is determined by the agreement with the statement that in the absence of one's own home and stable work it is better to postpone having a child and – at the same time – the absence of positive expectations regarding improvement of housing conditions, the presence of some housing deprivation; the presence of consumer loans payments. Removing economic barriers may allow this category of respondents to achieve the desired number of children.

Keywords

fertility, Russia, large families, third births, population survey, machine learning, LightGBM

JEL codes: J13

Introduction

Since 2024, large families and the promotion of third births have been the focus of the state's socio-demographic policy: in the first two months of last year, the Decree of the President of the Russian Federation "On measures of social support for large families" was adopted (Decree of the President… 2024), and later the national project "Family" was announced (National Project… 2025). At the same time, the demographic foundations of modern large families remain insufficiently studied. The aim of this work[1] is to determine the factors of intentions to have a third child among men and women with two children.

We use the conceptual apparatus of the life course concept (Clausen 1972) and represent individual reproductive trajectories as chains of decisions about the birth or non-birth of each child. Factors of intentions regarding the birth of children of different orders vary, which is confirmed by numerous empirical studies (Sinyavskaya and Tyndik 2009). The concept of the "transition to parenthood" has become firmly established in studies of the birth of first children (Rossi 1968; Baxter et al. 2015), and one can also talk about the "transition to having many children". Such transitions are not instantaneous, they are extended over time and begin with the formation of intentions. As part of the system of reproductive attitudes, intentions are studied as an independent determinant of reproductive behaviour (Philipov et al. 2006; Billari et al. 2009; Makarentseva and Biryukova 2023). Following the current discussion on the relationship between fertility and individual subjective well-being (Vignoli et al. 2020), we expand the range of factors studied in the intention to have a third child, including, in addition to demographic and socio-economic factors, factors of subjective well-being.

The empirical study was carried out using data from the survey "People. Family. Society" (2023) of the Presidential Academy. The analysis was performed in the Python environment using LightGBM machine learning models and the DART procedure, and SHAP values (SHapley Additive Explanations) were used to interpret the results. Machine learning methods are gradually becoming part of the analytical toolkit of fertility researchers (Vakulenko et al. 2023; Li and Xu 2022).

Theoretical Framework for Research on Third Births

Theories of demographic development are dominated by the modernization paradigm (Caldwell 1976), which justifies the "evolution towards a smaller family" (Esping-Andersen and Billari 2015). A series of studies by R. Inglehart shows that as social modernization

1 The article was prepared as part of the research work of the state assignment of RANEPA.

proceeds, the emphasis on survival and economic achievement as a priority gives way to an emphasis on quality of life (Inglehart 1971, 1977, 1997). In the area of fertility, this means an intergenerational shift from "norms that support fertility" to "norms of individual choice" (supporting, among other things, self-realization in social and personal life). In the mid-1980s, D. van de Kaa and R. Lesthaeghe developed the idea of a second demographic transition, linking the demographic changes observed in Western Europe with the change from materialistic values to post-materialistic ones (Lesthaeghe and van de Kaa 1986). Thus, according to R. Lesthaeghe, the central place in the decline in fertility is occupied by "the shift of the focus of life to the achievement of individual goals, which is the free determination by a person of both the goals themselves and the means of achieving them" (Lesthaeghe 1983: 429). This process is fueled by the secularization of values and growing income aspirations, and also gives rise to the avoidance of long-term commitments, including those such as raising children.

What factors in this paradigm can explain the maintenance and even some observed increase in fertility of third and subsequent children? Firstly, motherhood can be chosen as a way of self-realization. Evidence for this assertion among educated, postmodernist young women from Austria is presented in the work of K. Zulener (cited in Tomić-Koludrović et al. (2015)). The transition of parenthood into the category of ways of self-expression is a potential mechanism for stabilizing post-transition fertility. Tomić-Koludrović with coauthors also make an interesting observation, stating that in Croatia traditional and even pre-modern values and attitudes coexist side by side with modern and post-modern ones. According to their observations, the same respondent is capable of choosing traditional or premodern values and attitudes regarding raising children, but modern values and attitudes regarding paid work outside the home and postmodern values in organizing leisure time.

Secondly, the ongoing transformation of gender roles and distribution of responsibilities in the family becomes a central factor (Esping-Andersen and Billari 2015). G. Espin-Andersen and F. Billari allow for a return to the "extended family" as intra-family egalitarianism becomes both the normative model and the predominant practice in most families.

What individual-level factors can influence the transition to having many children? The birth of third children is rarely a separate subject of empirical research. In the classic book by C. Westoff et al. (1963) religiosity is named as one of the leading factors in having many children, but this study took place in the socio-demographic reality of the mid-20th century.

Based on the results of a qualitative study in Australia, A. Evans et al. (2009) argue that in a society with a mass norm of one- or two-child families, reasons for not having a third child are "rationalized" and articulated by respondents in the same language that permeates public discussions about maternal age, health risks, financial constraints, child care, and work-family balance. At the same time, the rhetoric of respondents who decided to have a third child is much less specific. The authors of the study highlight "parents' ideas about what a family is", which may include the desire to have children of different sexes. They point out that "the desire to have a child is not represented in public discourse, but it is one of the most compelling reasons for having children, including a third" (Evans et al. 2009). This discourse is related to the concept of "need for children" used by Russian researchers A.I. Antonov (1980), F.N. Ilyasov (2013), etc. The gender preference factor is consistently associated with the likelihood of family enlargement, as noted in a number of other studies (Berinde 1999).

In the work of J. Schröder et al. (2016) it is indicated that higher rates of third-births are found among women with two or more siblings, which the authors interpret as a transmission of family values. Women who have a new marital partner after the birth of their first or second child show higher rates of third childbearing, which may be due to the effect of union confirmation by common children.

The birth calendar can be a predictor of the occurrence of a third child: women who gave birth to their first child relatively early and had a short intergenetic interval between the first and second births have a higher probability of having a third birth (Hoem 1993). However, behind the choice of such a birth calendar there is the same wide range of factors as behind the choice of having many children. There is also evidence to the contrary: for example, a study conducted in Sweden showed that the later the marriage, the higher the likelihood of a third birth (Berinde 1999).

Quite a few works at the turn of the 20th and 21st centuries were devoted to the connection between high-order fertility and women's education. During this period, it was often hypothesized that, in some economic and institutional conditions, a high level of education might be positively associated with the chances of third births. By "conditions" was meant the advantage of reducing the direct and indirect costs of raising children due to greater state involvement in it (Hoem et al. 2001). Research results on this topic are contradictory. In the work of M. Callens and K. Croux (2005), with data for 15 countries, is shown that education has a negative effect on fertility; this effect also persists in Northern European countries, but the probability of third births is higher there.

A team of Russian researchers of large families (Pavlyutkin et al. 2021), based on qualitative research, highlights the importance of religious socialization and also examines the impact of social networks on reproductive trajectories (network effects). N. Balbo and M. Mills (2011) also write about the influence of the social environment, but from a different angle: "social pressure and social capital influence intentions to have a second or third child".

Thus, researchers usually focus on a limited range of determinants of the transition to having many children. We attempt to include as broad a set of factors as possible in our analysis so that we can use machine learning to select a subset of the best predictors of intentions to have a third child. Based on the capabilities of the database, which will be described in more detail below, the starting set of factors included gender, age, place of residence, age of children, presence of siblings, education, religiosity, several health characteristics, a number of indicators of financial situation, including subjective self-assessments, characteristics of housing conditions, involvement in lending, life satisfaction, happiness, trust, anxiety, involvement in communication, attitudes towards family and marriage, expectations of changes in financial situation, career and housing conditions, and others. One of the novelties of our study is the abandonment of manual selection of factors for analysis and the use of an approach in which the most suitable features are selected using machine learning algorithms.

Statistical estimates of the dynamics of third births in Russia

Russia is one of the countries with relatively low fertility, insufficient for the replacement of generations. However, after a long period of predominance of first births, their distribution by order changes, and against the background of the spread of childlessness, high-order fertility rates increase. The growth of heterogeneity among women in terms of the number of children born of different orders is noted in a number of countries and can be designated as the second stage of the "parity transition" or as one of the characteristics of the global demographic transition (Kalabikhina and Kuznetsova 2023). Let us briefly dwell on the dynamics of fertility of third children in Russia and how its current estimates compare with similar indicators in other countries.

The total fertility rate of third births in Russia in 2022 was 237 per 1000 women. Similar levels are demonstrated by countries such as Canada, Hungary, Denmark, and Finland (Figure 1). Developed countries with higher rates of third-order births include Ireland, Iceland, and the United States. Many developed countries show lower levels of third-birth fertility (southern countries with low fertility – Italy, Spain, Portugal, as well as Austria, Slovenia, the Czech Republic, etc.). A number of countries experienced a gradual decline in the intensity of this indicator in the period after 1990 (mainly the Scandinavian region: Norway, Sweden), while some experienced a moderate increase (Central and Eastern Europe: the Czech Republic, Bulgaria, Estonia, Slovenia). Russia stands out for its substantially more significant rise in the level of third-order births compared to other countries.

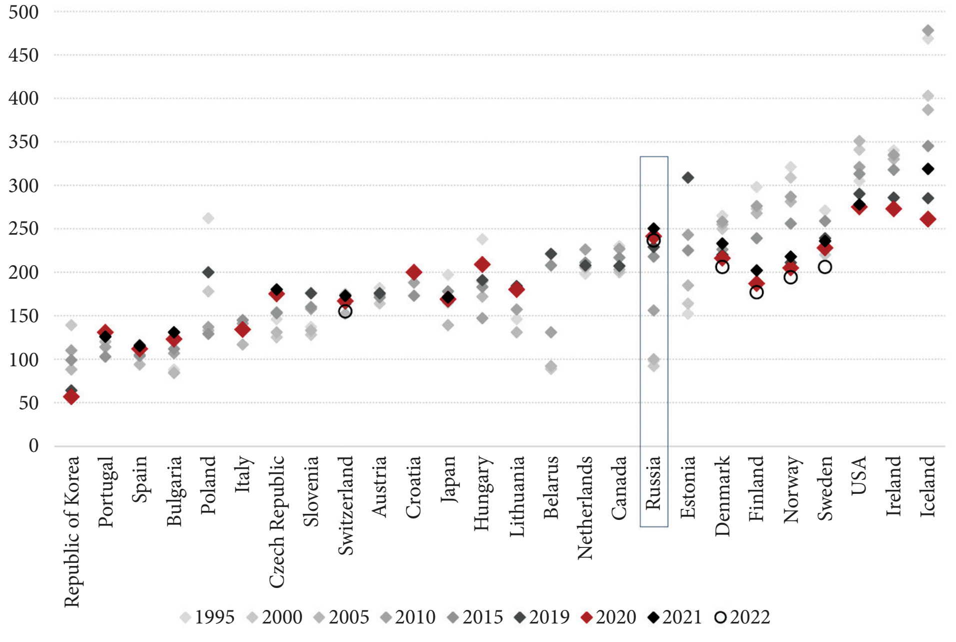


Figure 1. Total fertility rate of third and subsequent births in selected countries, per 1.000 women. *Source:* The Human Fertility Database 1990-2022, Rosstat, 2019-2022.

Since the second half of the 20th century, the dynamics of the total fertility rate of third births (TFR3) in Russia has demonstrated alternating stages of ups and downs. By the early 1960s it was at about 350 births per 1,000 women, quickly declining to 190 by 1970, and then to 146 by 1980. During the period of intensification of the Soviet family-demographic policy, the TFR3 increased to 257 (peaking in 1987), but against the backdrop of the subsequent socio-economic crisis it fell to 101 (1993). The low plateau lasted until 2006, after which the upward trend gained strength. In recent years, against the backdrop of a decline in first and second births, the contribution of third births has grown significantly – to 16.7%, which is the maximum value for the presented observation period. In the medium term, we can expect some reduction in the total fertility rate of third births – the lower number of second births after 2016 narrows the "reason" for having third children.

Cohort optics show a steady increase in the share of third births in the cumulative cohort fertility rate by the age of 35 in cohorts of women born in the 1980s (Figure 2). In the cohort of women born in 1987, the share of third births by age 35 was about 11%, which corresponds to the level of the cohort born in 1960; accumulated births at a later age will exceed it.

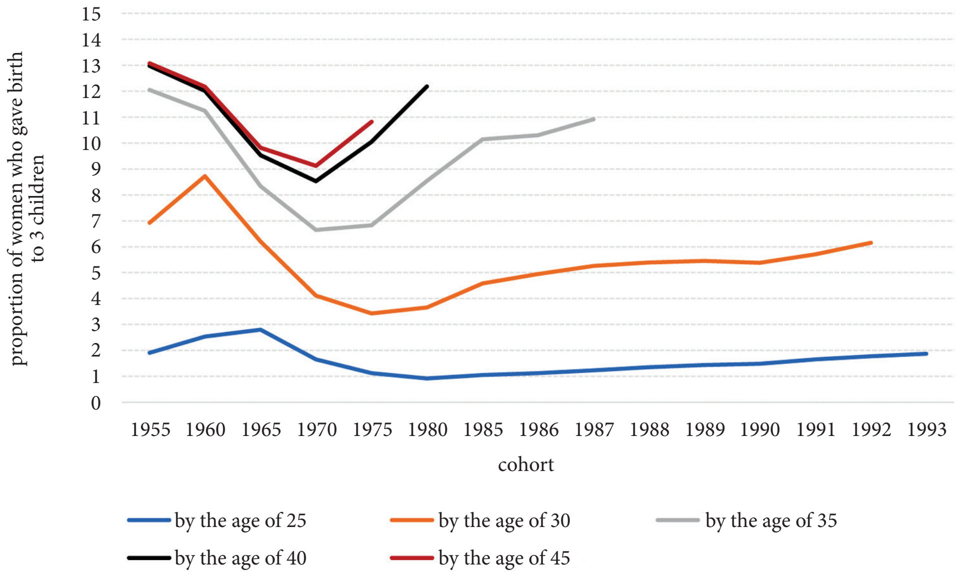


Figure 2. Cumulative cohort fertility rates of third births by woman's age and year of birth (born 1955-1993) in Russia. *Source:* The Human Fertility Database for data up to 2018; calculations based on Rosstat data for 2019-2022 were performed by P.O. Kuznetsova, Senior Researcher at the Center for InSAF, IAER RANEPA. *Note:* the denominator of the indicator is women with two children.

Thus, at present, Russia is experiencing a relatively high rate of third-child births by average European standards.

Model data and description

The empirical basis of the study is a population survey "People. Family. Society" (PFS), which covered 9,500 people and took place in the spring of 2023. The survey was conducted by telephone on a random sample of numbers. It represents the population of Russia aged 18-72. For the purposes of this study, two subsamples were formed: an extended one, which included respondents aged 20-44 who have a spouse[2] and two children (914 observations), and the main one, which included respondents from a married couple with two children, in which both partners were over 25 years old, and the woman was no more than 40 years old (686 observations). In the main sample, the proportion of the target group – those in-

2 From here on all marriages are taken into account, including unregistered ones.

tending to have many children in the coming years – is higher, which improves the modeling both statistically and substantively. Scheme of construction of dependent variable[3] is shown in Figure 3.

Data analysis was performed in the Python environment using the LightGBM (Light Gradient Boosting Machine) and DART (Dropouts meet Multiple Additive Regression Trees) machine learning models using confusion matrices comparing predicted and actual class labels.

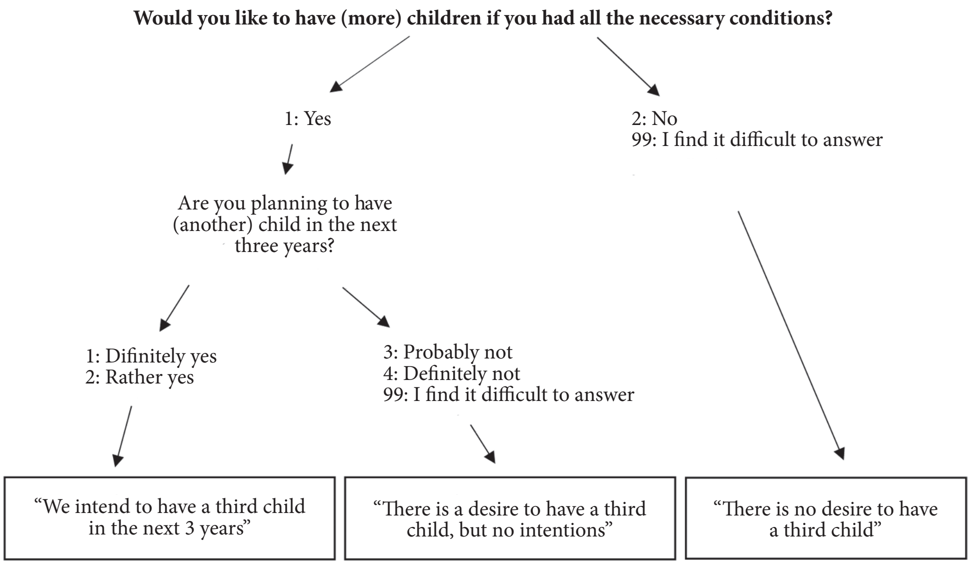


Figure 3. Scheme for constructing a dependent variable of the model. *Source:* questionnaire of the PFS survey.

The model performs the classification task. The dependent variable takes three values: "Intend to have a third child"; "There is a desire, but no intentions"; "No desire", with a frequency of 18, 42 and 40%, respectively, in the extended sample, and 22, 42 and 37% in the main sample. In fact, the classifier divides the training into three binary classification tasks: "Intend to have a third child in the next three years" versus the other two classes (Model 1); "There is a desire to have a third child, but there are no specific intentions" versus the other two groups (Model 2) and, accordingly, "No desire to have a third child" (Model 3) versus the rest. Errors are estimated using the One-vs-All multiclass classification objective (MultiClassOneVsAll).

The LightGBM algorithm is an ensemble of decision trees that is constructed using gradient boosting (LightGBM Documentation). Trees are calculated sequentially, each subsequent tree is constructed according to the principle of maximum reduction of the error of the previous stages of calculation (gradient improvement). Thus, when constructing multi-level trees, computing power is saved and the model training time is reduced, while maintaining

3 For more information on measuring and interpreting questions about reproductive attitudes in mass surveys, see (Makarentseva et al. 2021; Makarentseva and Biryukova 2023).

high quality of predictions (Ke et al. 2017; Hossain 2024). The algorithm selects features based on histograms and supports parallel and distributed computing. LightGBM takes into account the correlation between features and efficiently handles mutually exclusive or sparse features via histogram splits and feature bundling, which improves the performance and accuracy of the model. Mutually exclusive features (for example, the presence of a loan and the absence of a loan) can still appear within the same tree at different nodes — there is no hard prohibition. As a result, the model builds multi-level optimized trees that can capture complex relationships. Disadvantages of the LightGBM algorithm include an increased risk of overfitting, especially on small and noisy data, and the need for more careful tuning of hyperparameters such as the maximum number of leaves (max_leaves) and the minimum sum of weights in the leaf (min_sum_hessian_in_leaf). Both nuances were worked out when writing the calculation algorithm.

To ensure the stability of the modeling results, the DART technique, borrowed from neural networks, is used: during the model training process, some randomly selected trees are excluded from the ensemble, which prevents overtraining of the model, when the model produces "remembered" correct answers instead of a generalized solution. The excluded trees are used in model updating, which helps to improve the performance of the algorithm (Rashmi and Gilad-Bachrach 2015).

The quality of the classification model is presented in the form of a Confusion Matrix – a table that displays the correspondence between the actual and predicted class labels. If the target groups (classes) differ significantly in size in the database, then the highest modeling accuracy is achieved on the most numerous group. The results of machine learning models are illustrated by Shapley values (SHAP – SHapley Additive exPlanations) (Brownlee). SHAP assigns an additive contribution value to each feature for each individual prediction; violin plots show the distribution and sign of these contributions across all respondents.

During the modeling process, the data sample is divided into a training (on which the model is trained) and a test part in a ratio of 0.9 and 0.1, respectively. The proportions of classes in the training and test samples correspond to the proportions of classes in the original data, which allows us to maintain the representativeness of classes and avoid problems associated with class imbalance during training and testing of the model.

The dataset consists of 914 rows for the extended or 686 rows for the main subsamples, respectively (one row for one respondent), and 91 features, of which two features are numerical (age of the respondent and his/her spouse), the remaining features are categorical (the list and distributions are given in Table A1 of Appendix). Although LightGBM natively supports categorical variables, for SHAP analysis we applied one-hot encoding (splitting into dummy variables), which provides more interpretable feature attributions for individual categories and their effect on a particular choice of intentions regarding a third child (Machine learning textbook).

Some limitations of our analysis should be acknowledged before presenting the results. One of the most important concerns is the capabilities of the database. The PFS does not have a sufficient number of meso-level fertility factors (social offline and online networks). In general, the possibilities of measuring the quality of the social environment in the context of fertility in quantitative studies are limited, since obtaining comprehensive information about networks and their effects requires long-term and difficult for the respondent data collection (Rossier and Bernardi 2009). However, at the level of qualitative research, the influence of network effects on the transition to having many children is observed (Pavlyutkin et al. 2021; Makarentseva et al. 2021). Furthermore, not being a specialized survey on reproductive behaviour, the PFS does not have indicators of the quality of previous parenting experience, which is associated with subsequent intentions (Cassé et al. 2018). Sometimes, subjective well-being, measured through general and specific indicators of "happiness" (Billari 2009), serves as such an indicator; we also include these indicators in the analysis.

Modeling results

The accuracy of the model's prediction on the extended sample is 63.0% – the relatively small size of the group of those who intend to have a third child in the next three years leads to frequent errors in the classification of respondents in this group. An additional restriction on the age of the woman in a married couple and the transition to modeling on the main sample, on the one hand, reduces the number of observations. On the other hand, the share of the target group among respondents increases, and the model works much better: in most cases, it correctly classifies those who are aiming to have many children, and the prediction accuracy increases to 72.5%. Given the class imbalance, balanced accuracy and macro-F1 were also considered to ensure robustness of the evaluation. The matrices of predicted and actual indicators of target classes are presented in Figure A1 of the Appendix.

All further graphs in the article represent calculations based on the main sample of 686 respondents (Figure 4, Appendix Figures A2-A4). They demonstrate the effect of various features on falling into one or another class of intentions for a third child.

How to read and interpret charts:

- In the Appendix, Figure A4 shows three "ribbons" of classification results constructed by the model for one "signal" respondent. The model predicts that this respondent with 86% probability belongs to class 2 "There is a desire to have a third child, but there are no specific intentions", which is consistent with the actual data. The predicted probabilities of falling into classes 1 and 3 are small, amounting to 2% and 10%, so the classifier does not include him there. The strongest influence on the predicted probability of a given respondent being in class 1 is the respondent's answer that the spouse has no desire to have more children (reduces the probability), while this same factor increases the predicted probability of being in class 3. The factors on the ribbons are arranged in order of decreasing strength of their influence on the predicted probability of classification for a given respondent: to the left of the vertical line (marked in red) are those that increase, and to the right (marked in blue) are those that decrease this probability.

Then, such "ribbons" are summed up for all respondents for each of the three models, and the results of such summation are presented in the form of graphs in Figures 4 and A2-A3 of the Appendix:

- **The Y-axis** is a list of the first twenty features most frequently used by the classification model.
- **The X-axis** shows the significance of features, expressed as SHAP values (the contribution of each feature), which can be either positive or negative. For example, the signal respondent is represented on the graph of the first model (Figure 4) by a blue dot, plotted to the left of the vertical along the partn_wants_d axis at a distance taken from the first ribbon for the value partn_wants_d=0 (Figure A4). Or, for example, in Model 2 (Figure A2 of the Appendix), a light blue point is plotted on the partn_age axis (the color from blue to red reflects age values in the range of 25-44 years, in this case

Name of the feature
partn_wants_d: Spouse would like more children
housing_3yd1: Housing conditions are expected to improve within 3 years
partn.age: Age of the respondent's spouse
abort_mat: Abortion is considered acceptable in the event of serious financial difficulties in the family
abort_ever: Abortion is considered acceptable in any situation at the request of the woman
ageR_c: Age of the respondent
relig0_nprac: Non-practicing religiosity (believer, but does not attend services)
high_edu: Has a higher education
SSN_healthP_otvet: Completely agree that a person is fully responsible for his own health
credits_potr: Paying off consumer loans
yngch_age_36d: Age of the youngest child is 3-6 years
SSN_affN_lifecontrol: Weakly influence their life, 1-5 points out of 10
sibl_2plus: The respondent has two or more siblings
tp_nas_stolitsa: Type of settlement – regional capital
abort_med: Abortion for medical reasons is considered acceptable
conservative_yes: Agree with the statement "Yes, work is good, but women would like to take care of the house and children"
hlth_1y_const: Health has not changed over the past year
SSN_socialN_dover: Cautious, don't trust most people
alko_week: The respondent consumes alcohol once a week or more often
first_marr: First official marriage

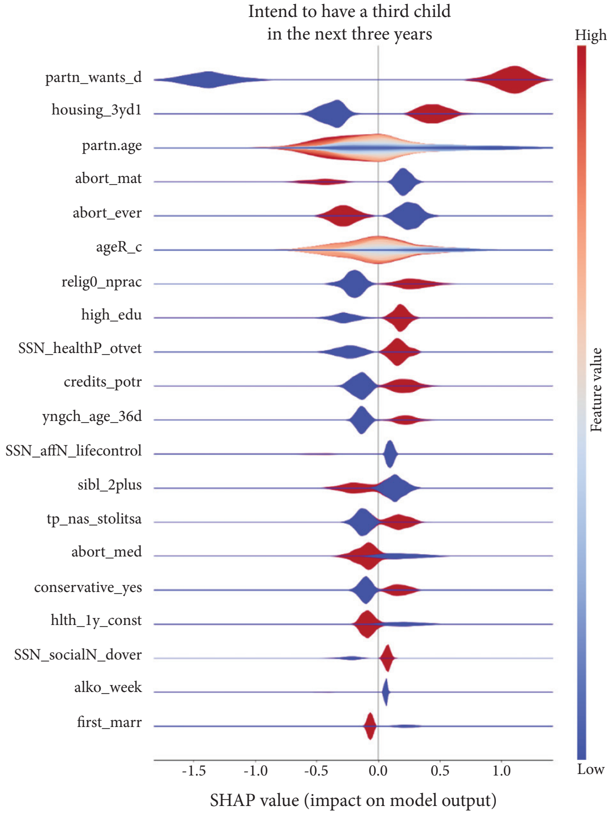


Figure 4. Top features working in classification model 1, SHAP value graph for the class "intend in the next three years". *Source:* authors' calculations based on PFS data.

partn_age=31). Since the factor works to increase the probability, the signal respondent point is located to the right of the vertical. In a similar way, the values of other factors are plotted for this and the other respondents; the points form "violins" due to accumulation.

- **Colors**: Red areas show the influence of high values of the trait on the decision to have a child, and blue areas show the influence of low values. For dummy variables, red areas correspond to the value 1, and blue areas correspond to the value 0.
- **Color scale**: If a variable is continuous and takes values over a range, its values are represented using shades of red and blue. For example, if the age of respondents is set in the range 25-44 years, then 25-year-old respondents will have dark blue dots, 44-year-olds will have deep red dots, and, for example, for a 35-year-old it will be a medium pale color.
- **Distribution density**: Each "violin" shows the distribution of SHAP values for a given feature. The width of the violin at each point reflects the frequency of a given SHAP

value, and wider parts of the violin indicate that a given feature functions as a classifier for a larger number of respondents.
- **Center Line**: The center line on each "violin" shows the median SHAP value for that feature. The values to the left of the central vertical line correspond the negative class, and to the right represent the positive class (for example, the left part in Figure 4 corresponds to the "no intention to have a third child" categories, and the right represents presence of the intention).

Thus, the wider the "violin", the higher the influence of a particular feature on the choice of a given class of intentions. Narrow violins indicate features with less influence. Important for interpretation is the fact that the presence and absence of a feature can have an asymmetric effect on the dependent variable (which distinguishes classification models from, for example, multinomial econometric models).

The greatest contribution to getting into class 1 "intend to have a third child in the next three years" is made by the indicator of the spouse's desire to have more children[4]. The age of the partner and the age of the respondent also have a significant influence, but it can be either positive or negative depending on the specific values – this feature is assigned a role of a control variable.

A significant classifying contribution is made by the "housing optimism" feature, namely the respondent's opinion that the family's housing conditions will improve in the next 3 years. The lack of housing optimism almost symmetrically reduces the chances of getting into the class of those intending to have many children. It is interesting to note that housing optimism often functions as a classifying feature for those who have a desire but no specific intentions regarding a third birth (Figure A2 in the Appendix) and for those who have no desire to have three children. In those models, it plays a negative role – respondents who believe that housing conditions will improve in the next three years are less likely to fall into those two classes. However, the contribution is smaller than for class 1, and for a number of representatives of the class "There is a desire, but no intentions" the effect is indistinguishable from zero.

In Models 2 and 3, the frequently used features included an indicator of agreement with the statement: "if you don't have your own home and a stable job, it's better to postpone having a child." For the class "There is a desire, but no intentions", this feature is included in the top five classifiers, working in a positive direction – supporters of this point of view with a high probability want a third child, but do not express specific intentions (Figure A2). In Model 3, this feature works less often, and there are cases when its effect is multidirectional – for some respondents it works positively, and for others negatively (Figure A3). Let us add that agreement with the thesis that "work is good, but women would like to take care of the house and children" distinguishes those who intend to have a third child soon (Figure 4).

Respondent's attitude towards the acceptability of abortions[5] acts as a powerful classifying factor in all three models. Those who do not consider abortion acceptable "in any situation at the woman's request" are highly likely to fall into the class of potential large families, while supporters of abortion's acceptability, on the contrary, do not fall into this class (Figure 4).

4 Strictly speaking, this is the respondent's reflected opinion about whether the spouse wants to have more children.

5 The detailed methodology for questions on abortions in the PFS survey is described in the article (Makarentseva 2024).

Disagreement with the opinion that abortion is acceptable in the event of serious financial difficulties in the family also distinguishes those who intend to have a third child in the next three years. The prohibition of abortion for medical reasons in Model 1 has a mixed effect, with a small proportion of respondents classifying in the opposite direction to the previous two. In Model 3, attitudes toward abortion are also among the top five classifying features: support of the acceptability of abortion in any situation at the request of a woman and the acceptability of abortion for medical reasons predict with high accuracy the respondent's reluctance to have a third child (Figure A3). The dividing line between those who do not want to have a third child and those who are not against a third child in principle, but have no intentions in the near future, can be drawn according to the criterion of attitude towards abortions for medical reasons: supporters of this point of view are highly likely not to fall into the class of "there is a desire, but no intentions" (Figure A2) and fall into the class of "no desire".

The indicators of the quality of current housing conditions worked in Model 2: the absence of belonging to a family with the worst housing conditions when there are two or more indicators of housing deprivation[6], classifies the respondent as not included in the "there is a desire, but no intentions" class (Figure A2). The feature "there are no deprivations, but there are fewer rooms than family members" works in the same Model 2 in the opposite direction: the absence of this feature increases the likelihood that there is a desire to have a third child, but there are no specific intentions. Also, the absence of this feature in Model 3 increases the likelihood of not wanting to have a third child, all other things being equal (Figure A3).

Financial status in the form of a synthetic indicator combining income level, consumer status (what money is enough for), share of food expenses, cash savings and prospects for financial situation for the next three years conditionally divides our respondents into four large groups in the ratio of 20% of the poorest (strata 1-2), 30% with average financial status (strata 3-4), 30% above average (stratum 5) and 20% of the wealthiest (stratum 6) (Table A1 of the Appendix). Financial status served as a classifier for two groups: "there is a desire to have a third child, but no intentions" and "no desire". Not belonging to stratum 5 (i.e. belonging either to the wealthiest or to the middle- or low-income groups) indicates that the respondent would like to have a third child but has no intentions (Figure A2). In Model 3, this classifier works in the opposite direction, reducing the probability that the respondent does not want more children (Figure A3). In Model 1, the financial situation did not make it into the top twenty factors for intentions to have a third child.

Living in the regional capital works in Model 1 and is not included in the group of influential features in other cases – all other things being equal, residents of capital cities are more likely to have many children compared to residents of other cities and rural settlements (Figure 4).

Considering the three levels of religiosity of respondents – high, moderate (believer, but hardly ever attends services) and absent – it can be noted that the second of the listed characteristics is the classifying feature. Believers who do not attend services are more likely than others to aim to have many children (Figure 4), while representatives of the other two groups are more likely to represent the "there is a desire, but no intentions" category (Figures A2-A3).

6 Five signs of housing deprivation are the need for cosmetic or major repairs, small living space, poor quality of utilities, and the expectation that housing conditions will become worse in the next three years.

Respondents with higher education are more likely to intend to have a third child in the next three years (Figure 4), while the lack of higher education according to Model 2 predicts belonging to the class "there is a desire, but no intentions" (Figure A2). The lack of higher education of the partner acts as a classifier in Model 3 – this feature distinguishes those who do not want to have a third child (Figure A3).

The presence of consumer loans is the second most frequently triggered classification criterion for belonging to the class of those who want to have a third child but do not have specific intentions (Figure A2). The absence of consumer loans with a high probability classifies the respondent into the category of those who do not want to have a third child (Figure A3).

If the respondent himself is not from a large family – there is no sign of two or more brothers and sisters – Model 1 classifies him, all other things being equal, as one of those who intend to have a third child in the coming years (Figure 4). If the respondent's younger of two children is between 3 and 6 years old, he or she is more likely to express an intention to have more children (Figure 4). In Models 2 and 3, the gender of the respondent is a significant classifying feature. Taking into account the partner's desire and other characteristics, men more often fall into the "there is desire, but no intentions" class, while women, all other things being equal, fall into the "no desire" class (Figures A2-A3). In Model 1, respondent gender was not included in the top twenty classifying features.

Three dimensions of subjective well-being serve as classifiers in Model 1. Those who "completely agree" that a person is fully responsible for their health; those who rate their control over their lives as higher than 5 points; and those who are not inclined to trust most people ("generalized trust") fall into the category of those who intend to have a third child in the next three years (Figure 4). Full responsibility for one's health also acts as a classifying characteristic in Model 2, working in the opposite direction: those who do not think so are highly likely to want to have a third child, but have no intentions (Figure A2). Model 2 also shows that those who do not feel lonely fall into the "desire but no intention" category. Those who sometimes or often experience feelings of loneliness are definitely among those who do not want to have a third child (Figure A3). The absence of a sign of frequent alcohol consumption (once a week or more) highly likely indicates the intention to have a third child (Figure 4).

Conclusions and discussion

The interpretation of the factors in the transition to having many children should differ somewhat from that in relation to children of the first two orders. The birth of a second and especially a first child is within the social norm of having children. The demographic and socio-economic factors identified for such intentions determine not so much the fact itself, but the choice of time, stage of life for the birth of a child. The transition to having many children is an individual choice outside the mass social norm of having many children.

When interpreting the results of models, we are talking not so much about specific factors as about a certain field of respondent characteristics. A profile of conservative attitudes emerges that determine higher chances of being included in the class of those intending to have a third child in the next three years: a negative attitude towards the acceptability of abortion (according to two indicators at once), agreement with the statement that women would like to take care of the house and children, moderate religiosity, when a person considers himself a believer, but almost never attends services. In addition, factors that indicate the choice of the appropriate time for the birth of a child are important, namely the age of the youngest child from 3 to 6 years, positive expectations regarding improved housing conditions, as well as the desire for the spouse to have more children as part of the couple's agreed reproductive attitude. The latter indicator cannot be interpreted as an independent one; it may be part of an already agreed reproductive attitude of the couple.

The class of those who talk about the desire to have more children, but do not express specific plans for having a child (this could be either a postponement or a complete refusal to realize the desire), is of particular interest in the context of the possibilities of demographic policy. The economic profile of factors includes: agreement with the statement that in the absence of one's own home and stable work, it is better to postpone having a child; and – at the same time – the absence of positive expectations regarding improvement of housing conditions, the presence of some housing deprivation; if the respondent is a consumer loan borrower. Subjective factors have a weak impact on getting into this class; the only significant factor is whether both spouses desire to have children. If we interpret the respondents' answers literally, then the removal of economic barriers could allow this category of respondents to have the desired number of children. However, as noted in the review of studies above (Evans et al. 2009), such responses can be interpreted as articulating rational reasons for abandoning intentions to have another child among those whose desire is vague and mixed.

Such indicators of subjective well-being as anxiety, happiness ratings and life satisfaction did not show a significant contribution, although happiness ratings are conceptually used as indicators of children's "usefulness" in the "happiness economics" approach (Billari 2009). In previous studies, including the 2020 PFS survey data, they turned out to be significant (Vakulenko et al. 2023). Perhaps measuring and including specific indicators (e.g., "happiness of being a parent" and the like) would show a different contribution.

In general, it is quite difficult to compare our results with previous studies, since almost all of them were carried out using econometric modeling methods. The approach we used was intended to limit the impact of manual selection of factors for analysis by allowing the model to select from the largest available set of indicators. It is too early to draw conclusions about the advantages of this method over traditional models. At the very least, it places more demand on the completeness of the database.

For further research on the factors of both reproductive intentions and actual fertility, the possibilities of empirical data will be of great importance. Firstly, verification of meso-level factors (network effects) and the quality of previous parental experience and satisfaction with it using quantitative surveys would significantly expand our understanding of the transition to having many children. Secondly, reproductive intentions are not fully implemented, and the factors of intentions and actual births differ – to study the latter, specialized longitudinal surveys are needed.

Reference list

Antonov AI (1980) Sociologiya rozhdaemosti: teoreticheskie i metodologicheskie problemy [Sociology of fertility: theoretical and methodological issues]. M.: Statistics. (In Russ.)

Vakulenko ES, Mitrofanova ES, Gorskiy DI (2023) Reproductive intentions of Russians with children at the beginning of COVID-19 pandemic. Voprosy Ekonomiki (4): 85–102. (In Russ.) https://doi.org/10.32609/0042-8736-2023-4-85-102

Ilyasov FN (2013) Potrebnost' v detyakh i reproduktivnoe povedenie [The need for children and reproductive behaviour]. Monitoring of public opinion: economic and social changes (1): 168–77. (In Russ.) URL: https://elibrary.ru/item.asp?id=18886414

Kalabikhina IE, Kuznetsova PO (2023) Population Heterogeneity in the Number of Children Born: Is There a "Parity Transition"? Monitoring of Public Opinion: Economic and Social Changes 2(174): 57–81. (In Russ.) https://doi.org/10.14515/monitoring.2023.2.2362

Makarentseva AO (2024) Judgments about the Acceptability of Abortion: Role of the Questions Order in the Survey Questionnaire and Influence of Other Factors. Monitoring of Public Opinion: Economic and Social Changes (3): 67-87. (In Russ.) https://doi.org/10.14515/monitoring.2024.3.2530

Makarentseva AO, Biryukova SS (2023) Factors, Consistency, and Realization of Reproductive Intentions in Russia. Monitoring of Public Opinion: Economic and Social Changes (2): 31–56. (In Russ.) https://doi.org/10.14515/monitoring.2023.2.2379

Makarentseva AO, Galieva NI, Rogozin DM (2021) Desire (Not) To Have Children in the Population Surveys. Monitoring of Public Opinion: Economic and Social Changes (4): 492–515. (In Russ.) https://doi.org/10.14515/monitoring.2021.4.1871

Pavlyutkin IV, Goleva MA, Borisova ON (2021) The sea of life: How large families are born in modern Russia. M.: Publishing house of PSTGU. (In Russ.)

Sinyavskaya OV, Tyndik AO (2009) Rozhdaemost' v sovremennoj Rossii: ot planov k dejstviyam? [Fertility in modern Russia: from plans to actions?] In: Zakharov SV, Maleva TM, Sinyavskaya OV (Eds) Parents and children, men and women in family and society. Issue 2. NISP, Moscow, 9–44. (In Russ.) URL: https://www.demoscope.ru/weekly/knigi/socpol/socpol_v2.html

Balbo N, Mills M (2011) The effects of social capital and social pressure on the intention to have a second or third child in France, Germany, and Bulgaria, 2004–05. Population Studies 65(3): 335-51. https://doi.org/10.1080/00324728.2011.579148

Baxter J, Buchler S, Perales F, Western M (2015) A life-changing event: First births and men's and women's attitudes to mothering and gender divisions of labor. Social Forces 93(3): 989–1014. https://doi.org/10.1093/sf/sou103

Berinde D (1999) Pathways to a third child in Sweden. European Journal of Population / Revue Européenne de Démographie 15: 349–78. https://doi.org/10.1023/a:1006287630064

Billari FC (2009) The happiness commonality: Fertility decisions in low-fertility settings. In: How generations and gender shape demographic change. Chapter 1. UNECE. Generations and Gender Programme. New York and Geneva: 6–32. URL: https://unece.org/sites/default/files/datastore/fileadmin/DAM/pau/_docs/ggp/2008/GGP_2008_GGConf_Publ_1_Chapter-1.pdf

Billari FC, Philipov D, Testa MR (2009) Attitudes, norms and perceived behavioural control: Explaining fertility intentions in Bulgaria. European Journal of Population / Revue Europénne de Démographie 25: 439–65. https://doi.org/10.1007/s10680-009-9187-9

Caldwell JC (1976) Toward a restatement of demographic transition theory. Population and Development Review 2(3/4): 321–66. https://doi.org/10.2307/1971615

Callens M, Croux C (2005) The impact of education on third births. A multilevel discrete-time hazard analysis. Journal of Applied Statistics 32(10): 1035–50. https://doi.org/10.1080/02664760500165040.

Cassé JF, Oosterman M, Volling BL, Schuengel C (2018) First-time motherhood experiences and the decision to have a second child. Journal of Family Issues 39(3): 555–78. https://doi.org/10.1177/0192513X16662100

Clausen J (1972) The life course of individuals. In: Riley MW, Johnson ME, Foner A (Eds) A sociology of age stratification. Russell Sage Foundation, New York, 457-514.

Esping-Andersen G, Billari FC (2015) Re-theorizing family demographics. Population and Development Review 41(1): 1–31. https://doi.org/10.1111/j.1728-4457.2015.00024.x

Evans A, Barbato C, Bettini E, Gray E, Kippen R (2009) Taking stock: parents' reasons for and against having a third child. Community, Work & Family 12(4): 437–54. https://doi.org/10.1080/13668800902823599

Hoem B (1993) The compatibility of employment and childbearing in contemporary Sweden. Acta Sociologica 36(2): 101–20. https://doi.org/10.1177/000169939303600202

Hoem JM, Prskawetz A, Neyer G (2001) Autonomy or conservative adjustment? The effect of public policies and educational attainment on third births in Austria, 1975–96. Population Studies 55(3): 249–61. https://doi.org/10.1080/00324720127700

Inglehart R (1971) The silent revolution in Europe: Intergenerational change in post-industrial societies. American Political Science Review 65(4): 991–1017. https://doi.org/10.2307/1953494

Inglehart R (1977) Values, objective needs, and subjective satisfaction among Western publics. Comparative Political Studies 9(4): 429–58. https://doi.org/10.1177/001041407700900403

Inglehart R (1997) Modernization, postmodernization and changing perceptions of risk. International Review of Sociology 7(3): 449–59. https://doi.org/10.1080/03906701.1997.9971250

Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q, Liu T-Y (2017) LightGBM: A highly efficient gradient boosting decision tree. In: Guyon I, von Luxburg U, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R (Eds) Advances in Neural Information Processing Systems, vol.30. Long Beach, CA, USA, 3149-57. URL: https://papers.nips.cc/paper/6907-lightgbm-a-highly-efficient-gradient-boosting-decision-tree

Lesthaeghe R (1983) A century of demographic and cultural change in Western Europe: An exploration of underlying dimensions. Population and Development Review 9(3): 411-35. http://dx.doi.org/10.2307/1973316

Lesthaeghe R, van de Kaa DJ (1986) Two demographic transitions? In: van de Kaa DJ, Lesthaeghe R (Eds) Population: Growth and Decline. Deventer, 9–24.

Li M, Xu X (2022) Fertility intentions for a second child and their influencing factors in contemporary China. Frontiers in Psychology 13: 883317. https://doi.org/10.3389/fpsyg.2022.883317

Philipov D, Spéder Z, Billari FC (2006) Soon, later, or ever? The impact of anomie and social capital on fertility intentions in Bulgaria (2002) and Hungary (2001). Population Studies 60(3): 289–308. https://doi.org/10.1080/00324720600896080

Rashmi KV, Gilad-Bachrach R (2015) DART: Dropouts meet Multiple Additive Regression Trees. In: Proceedings of the 18th International Conference on Artificial Intelligence and Statistics (AISTATS). San Diego, CA, USA. JMLR: W&CP volume 38. URL: https://arxiv.org/pdf/1505.01866

Rossi AS (1968) Transition to Parenthood. Journal of Marriage and Family 30(1): 26–39. https://doi.org/10.2307/350219

Rossier C, Bernardi L (2009) Social interaction effects on fertility: Intentions and behaviors. European Journal of Population / Revue Européenne de Démographie 25(4): 467-85. https://doi.org/10.1007/s10680-009-9203-0

Schröder J, Schmiedeberg C, Brüderl J (2016) Beyond the two-child family: Factors affecting second and third birth rates in West Germany. Journal of Family Research 28(1): 3–18. https://doi.org/10.3224/zff.v28i1.22918

Tomić-Koludrović I, Petrić M, Zdravković Ž (2015) Retraditionalization or Reflexive Modernity: A Sociological Explanation of Fertility Trends in Mature Transitional Croatia. In: Cvetičanin P, Mangova I, Markovikj N (Eds) A Life for Tomorrow: Social Transformations in South-East Europe. Institute for democracy "Societas Civilis", Skopje, 173–98.

Vignoli D, Mencarini L, Alderotti G (2020) Is the effect of job uncertainty on fertility intentions channeled by subjective well-being? Advances in Life Course Research 46: 100343. https://doi.org/10.1016/j.alcr.2020.100343

Westoff CF, Potter RG Jr, Sagi PC (1963) The Third Child: A Study in the Prediction of Fertility. Princeton University Press, Princeton, NJ.

Other sources of information

National project "Family" URL: nationalprojects.rf/new-projects/semya/

Decree of the President of the Russian Federation (2024) No. 63 of 23.01.2024 "On measures of social support for large families". URL: https://www.consultant.ru/document/cons_doc_LAW_467710/

Machine learning textbook / School of data analysis (SDA). Yandex. URL: https://education.yandex.ru/handbook/ml (date of reference: 07.08.24).

Brownlee J Machine Learning Mastery with Python. URL: https://machinelearningmastery.com/ (date of reference: 07.08.24).

Hossain MM (2024) Mastering LightGBM: An In-Depth Guide to Efficient Gradient Boosting / Medium. Towards Data Science. URL: https://medium.com/@mohtasim.hossain2000/mastering-lightgbm-an-in-depth-guide-to-efficient-gradient-boosting-8bfeff15ee17 (date of reference: 07.08.24).

LightGBM Documentation. Microsoft Corporation. 2023. URL: https://lightgbm.readthedocs.io/en/stable/index.html (date of reference: 07.08.24).

Information about the authors

- Alla Olegovna Makarentseva – Candidate of Economics, Leading Researcher, Center «Institute of Social Analysis and Prediction» IAER RANEPA. Moscow, 119034, Russia. E-mail: makarentseva-ao@ranepa.ru
- Aleksandra Yaroslavovna Burdyak – Senior Researcher, Center «Institute of Social Analysis and Prediction» IAER RANEPA. Moscow, 119034, Russia. E-mail: burdyak-ay@ranepa.ru
- Iaroslav Igorevich Sheipak – Master's Graduate in Machine Learning and Data Analysis, Moscow Institute of Physics and Technology (MIPT). Dolgoprudny, Moscow region, 141701, Russia. E-mail: iaroslav.sheipak@gmail.com

Appendix

Table A1. The share of respondents with this characteristic among respondents under 44 years of age with two children, including in households with a woman aged 25-40 years, %

Name of the feature	Families with two children, respondent under 44 years of age, 914 respondents				Families with two children, women 25-40 years old, 686 respondents			
	There is no desire to have a third	Intend to have a third	There is a desire, but no intentions	Total	There is no desire to have a third	Intend to have a third	There is a desire, but no intentions	Total
partn_wants_d: Spouse would like more children	25.2	91.0	64.0	**53.6**	28.2	91.3	62.5	**56.1**
Male: Men	33.5	55.7	58.5	**48.1**	33.7	53.7	58.6	**48.4**
ageR_c: Age of the respondent (mean, years)	37.7	35.2	37.0	**36.9**	36.4	34.9	36.2	**36.0**
partn.age: Age of the respondent's spouse (mean, years)	39.3	35.3	37.7	**37.9**	37.5	34.6	36.2	**36.4**
first_marr: First official marriage	75.6	68.9	76.9	**74.9**	76.2	68.5	79.6	**75.9**
high_edu: Has a higher education	63.4	65.3	60.1	**62.4**	66.3	65.1	59.3	**63.1**
partn_highedu: Spouse has a higher education	54.8	60.5	58.0	**57.2**	56.3	61.1	57.2	**57.7**
sibl_2plus: The respondent has two or more siblings	34.1	34.7	40.7	**37.0**	34.5	36.9	41.8	**38.0**
sibl_1plus: The respondent has one or more siblings	87.3	83.2	87.3	**86.5**	86.9	86.6	89.5	**87.9**
Smoke: The respondent smokes	32.4	34.1	37.6	**34.9**	31.7	34.2	37.2	**34.5**
alko_week: The respondent consumes alcohol once a week or more often	8.9	9.0	15.8	**11.8**	9.9	7.4	16.5	**12.1**
relat_meet: The respondent meets with friends less than once a month	18.0	18.6	21.2	**19.5**	18.3	16.1	20.0	**18.5**
relat_tfin: The respondent talks to friends on the phone once a month or less	6.4	6.0	8.0	**7.0**	6.0	6.0	6.7	**6.3**

Name of the feature	Families with two children, respondent under 44 years of age, 914 respondents				Families with two children, women 25-40 years old, 686 respondents			
	There is no desire to have a third	Intend to have a third	There is a desire, but no intentions	Total	There is no desire to have a third	Intend to have a third	There is a desire, but no intentions	Total
dom: Housing type – house	33.2	28.1	28.2	**30.2**	34.9	30.2	28.4	**31.2**
zh_owners: Live in their own house/apartment	90.9	83.2	87.0	**87.9**	89.3	83.9	88.1	**87.6**
zh_strat4_d1: Poor housing, two or more housing deprivations	18.0	17.4	27.5	**21.9**	15.1	16.8	27.0	**20.4**
zh_strat4_d2: One housing deprivation	45.4	46.7	48.4	**46.9**	47.6	47.7	51.2	**49.1**
zh_strat4_d3: No housing deprivations, but fewer rooms than family members	28.3	30.5	21.2	**25.7**	30.2	29.5	18.9	**25.4**
zh_strat4_d4: Good or very good housing, enough rooms, satisfaction with housing	8.3	5.4	2.8	**5.5**	7.1	6.0	2.8	**5.1**
credits_ipo: Mortgage borrowers	31.3	29.3	33.4	**31.8**	33.3	29.5	35.8	**33.5**
credits_potr: Consumer loan borrowers	36.8	44.3	46.9	**42.5**	34.9	45.6	49.1	**43.1**
credits_ovrdu: The respondent has overdue loan payments	1.7	4.2	4.4	**3.3**	1.6	4.7	4.9	**3.6**
str1: Three of five signs of poverty: income below the subsistence minimum; difficult to buy clothes or pay for housing and utilities; spend more than 1/2 of their income on food; no savings; expect their financial situation to worsen in the next three years	9.4	8.4	18.7	**13.1**	7.5	8.1	18.2	**12.1**
str2: One or two signs of poverty and no signs of wealth	6.4	6.0	11.4	**8.4**	4.8	6.0	9.8	**7.1**
str3: There are signs of both wealth and poverty	26.9	29.3	26.9	**27.4**	28.2	27.5	27.7	**27.8**
str4: No signs of either wealth or poverty	1.9	0.6	3.6	**2.4**	2.4	0.7	3.5	**2.5**
str5: One or two signs of poverty and no signs of wealth	32.4	26.9	21.0	**26.6**	35.7	28.9	20.4	**27.8**

Name of the feature	Families with two children, respondent under 44 years of age, 914 respondents				Families with two children, women 25-40 years old, 686 respondents			
	There is no desire to have a third	Intend to have a third	There is a desire, but no intentions	Total	There is no desire to have a third	Intend to have a third	There is a desire, but no intentions	Total
str6: Three of five signs of wealth: top 20% by income; can buy durable goods; spend less than 1/3 of their income on food; have enough savings for a year of consumption; expect their financial situation to improve in the next three years	23.0	28.7	18.4	**22.1**	21.4	28.9	20.4	**22.6**
hlth_1y_better: Health has improved over the past year	6.4	12.0	7.0	**7.7**	6.7	12.1	7.4	**8.2**
hlth_1y_worse: Health has deteriorated over the past year	15.5	21.6	23.1	**19.8**	13.1	22.8	23.9	**19.7**
hlth_1y_const: Health has not changed over the past year	77.3	65.9	69.9	**72.1**	79.4	64.4	68.8	**71.7**
marr_regstd: Registered marriage	91.7	85.0	90.9	**90.2**	90.9	84.6	91.6	**89.8**
housing_matters: Agree with the statement that if you don't have your own home and a stable job, it's better to postpone having a child	47.1	52.1	61.1	**53.9**	50.4	53.0	63.9	**56.6**
babded: Grandparents help with children	46.5	48.5	46.1	**46.7**	49.2	48.3	46.7	**48.0**
uhod: The respondent cares for elderly relatives	11.1	15.0	12.2	**12.3**	9.9	12.1	10.9	**10.8**
abort_med: Abortion for medical reasons is considered acceptable	78.1	62.3	63.0	**68.8**	79.0	62.4	62.5	**68.5**
abort_mat: Abortion is considered acceptable in case of serious financial difficulties in the family	37.4	17.4	30.3	**30.7**	38.5	16.8	30.5	**30.5**
abort_ever: Abortion is considered acceptable in any situation at the request of the woman	56.8	35.3	44.6	**47.7**	56.3	33.6	42.1	**45.5**
yngch_age_02d: Age of the youngest child is 0-2 years	14.7	25.7	23.1	**20.2**	19.0	25.5	27.0	**23.8**
yngch_age_36d: Age of the youngest child is 3-6 years	27.4	42.5	27.5	**30.2**	33.7	44.3	33.0	**35.7**

Name of the feature	Families with two children, respondent under 44 years of age, 914 respondents				Families with two children, women 25-40 years old, 686 respondents			
	There is no desire to have a third	Intend to have a third	There is a desire, but no intentions	Total	There is no desire to have a third	Intend to have a third	There is a desire, but no intentions	Total
yngch_age_7pd: Age of the youngest child is 7 years and older	51.0	26.3	41.2	**42.3**	44.8	26.2	35.8	**37.0**
yngch_age_nd: Age of the youngest child is not specified	6.9	5.4	8.3	**7.2**	2.4	4.0	4.2	**3.5**
career_3yd1: Career is expected to improve within 3 years	26.9	38.3	28.0	**29.4**	29.0	38.9	28.4	**30.9**
housing_3yd1: Housing is expected to improve within 3 years	38.8	63.5	39.6	**43.7**	42.5	65.8	41.4	**47.1**
financ_3yd1: Financial situation is expected to improve within 3 years	49.9	63.5	45.9	**50.7**	50.8	63.1	47.0	**51.9**
career_3yd3: Career is expected to worsen within 3 years	3.0	2.4	4.4	**3.5**	1.2	2.7	4.9	**3.1**
housing_3yd3: Housing is expected to worsen within 3 years	1.1	2.4	2.8	**2.1**	0.8	2.7	3.2	**2.2**
financ_3yd3: Financial situation is expected to worsen within 3 years	6.1	7.2	10.4	**8.1**	3.6	7.4	9.5	**6.9**
career_3yd5: The respondent is not working	28.3	25.7	24.9	**26.4**	32.5	27.5	26.3	**28.9**
conservative_yes: Agree with the statement "work is good, but women would like to take care of the house and children"	38.0	48.5	44.0	**42.5**	35.7	47.7	43.5	**41.5**
conservative_no: Disagree that "work is good, but women would like to take care of the house and children"	55.1	48.5	46.9	**50.4**	58.7	49.0	47.4	**51.9**
relig0_no: Not religious	21.1	17.4	26.7	**22.8**	20.6	16.8	27.0	**22.4**
relig0_high: Religiosity is high	41.0	40.7	37.0	**39.3**	40.1	40.3	38.2	**39.4**
relig0_nprac: Non-practicing religiosity (believer, but does not attend services)	38.0	41.9	36.3	**38.0**	39.3	43.0	34.7	**38.2**
tp_nas_stolitsa: Type of settlement – regional capital	46.5	46.7	43.3	**45.2**	46.0	47.0	39.6	**43.6**

Name of the feature	Families with two children, respondent under 44 years of age, 914 respondents				Families with two children, women 25-40 years old, 686 respondents			
	There is no desire to have a third	Intend to have a third	There is a desire, but no intentions	Total	There is no desire to have a third	Intend to have a third	There is a desire, but no intentions	Total
tp_nas_selo: Type of settlement – rural area	16.9	20.4	18.9	**18.4**	16.7	19.5	20.4	**18.8**
tp_nas_drgorod: Type of settlement – other city	36.6	32.9	37.8	**36.4**	37.3	33.6	40.0	**37.6**
SSN_healthm1: Subjective well-being, health domain, negative pole	1.9	1.2	5.4	**3.3**	2.0	1.3	6.0	**3.5**
SSN_gedom1: Subjective well-being, hedonism domain, negative pole	2.8	1.2	3.9	**3.0**	2.4	1.3	2.8	**2.3**
SSN_affm1: Subjective well-being, affect domain, negative pole	2.2	1.2	3.1	**2.4**	1.6	1.3	2.1	**1.7**
SSN_socintegrm1: Subjective well-being, social integration domain, negative pole	3.9	3.0	2.3	**3.1**	2.8	2.7	2.5	**2.6**
SSN_socialm1: Subjective well-being, domain of social reality assessment, negative pole	11.1	13.8	17.9	**14.4**	8.7	12.8	17.9	**13.4**
SSN_healthp1: Subjective well-being, health domain, positive pole	44.0	48.5	39.9	**43.1**	47.6	51.0	40.0	**45.2**
SSN_gedop1: Subjective well-being, hedonism domain, positive pole	31.3	41.9	25.9	**31.0**	33.3	40.9	26.3	**32.1**
SSN_affp1: Subjective well-being, affect domain, positive pole	16.6	24.0	14.5	**17.1**	16.3	24.8	12.6	**16.6**
SSN_socintegrp1: Subjective well-being, social integration domain, positive pole	50.4	61.1	50.8	**52.5**	49.2	61.1	48.8	**51.6**
SSN_socialp1: Subjective well-being, domain of social reality assessment, positive pole	16.3	16.8	15.3	**16.0**	15.9	16.1	15.4	**15.7**
SSN_healthN_bad: Rate their health as poor or very poor	1.4	3.0	4.4	**3.0**	1.6	3.4	4.2	**3.1**
SSN_healthP_better: Rate health as “better than peers”	25.8	31.7	26.4	**27.1**	25.8	30.9	25.6	**26.8**
SSN_healthP_better: Rate health as “worse than peers”	10.0	11.4	17.9	**13.6**	12.3	12.1	20.7	**15.7**

Name of the feature	Families with two children, respondent under 44 years of age, 914 respondents				Families with two children, women 25-40 years old, 686 respondents			
	There is no desire to have a third	Intend to have a third	There is a desire, but no intentions	Total	There is no desire to have a third	Intend to have a third	There is a desire, but no intentions	Total
SSN_healthP_otvet: Completely agree that a person is fully responsible for his own health	58.7	64.7	53.6	**57.7**	59.9	64.4	53.3	**58.2**
SSN_healthN_otvet: Do not agree that «a person is fully responsible for their own health”	7.2	10.8	14.0	**10.7**	6.0	10.1	14.4	**10.3**
SSN_healthP_good: Rate their health as good or very good	58.4	59.3	52.6	**56.1**	61.5	62.4	52.6	**58.0**
SSN_gedoN_udovl: Rather or completely dissatisfied with life	8.3	7.8	17.1	**11.9**	7.5	7.4	16.1	**11.1**
SSN_gedoP_udovl: Completely satisfied with life	19.7	24.6	12.2	**17.4**	20.2	25.5	11.9	**17.9**
SSN_gedoN_happy: Rather or not at all happy	3.3	1.2	3.6	**3.1**	2.8	1.3	3.2	**2.6**
SSN_gedoP_happy: Very happy	23.3	37.7	23.6	**26.0**	25.0	36.2	23.9	**27.0**
SSN_affN_trevog: Experience severe anxiety	5.0	6.0	8.3	**6.6**	2.8	5.4	7.7	**5.4**
SSN_affP_trevog: Do not experience anxiety	40.7	41.9	37.0	**39.4**	40.1	43.6	37.9	**39.9**
SSN_affN_lifecontrol: Weakly influence on their own life, 1-5 points out of 10	18.6	8.4	19.7	**17.2**	17.1	6.7	20.0	**16.0**
SSN_affP_lifecontrol: Have very strong influence on their own life, 10 out of 10	33.5	40.1	34.5	**35.1**	34.9	39.6	33.0	**35.1**
SSN_integrP_obsh: No lack of communication	74.2	73.7	74.6	**74.3**	73.0	72.5	74.7	**73.6**
SSN_integrP_vnim: No lack of attention	77.3	82.0	82.6	**80.4**	77.4	81.2	83.2	**80.6**
SSN_integrP_izol: Don’t feel isolated	85.0	89.8	89.6	**87.9**	85.7	89.9	90.2	**88.5**
SSN_integrP_odinok: Don’t feel lonely	78.7	79.0	81.3	**79.9**	77.8	77.2	81.8	**79.3**
SSN_integrN_obsh: Almost always feel a lack of communication	0.8	1.2	2.1	**1.4**	0.4	1.3	2.1	**1.3**

Name of the feature	Families with two children, respondent under 44 years of age, 914 respondents				Families with two children, women 25-40 years old, 686 respondents			
	There is no desire to have a third	Intend to have a third	There is a desire, but no intentions	Total	There is no desire to have a third	Intend to have a third	There is a desire, but no intentions	Total
SSN_integrN_vnim: Almost always feel a lack of attention	1.7	1.2	1.0	**1.3**	0.8	1.3	1.4	**1.2**
SSN_integrN_izol: Almost always feel isolated	0.6	1.2	1.0	**0.9**	0.8	0.7	1.4	**1.0**
SSN_integrN_odinok: Almost always feel lonely	0.6	1.8	0.3	**0.7**	0.0	1.3	0.4	**0.4**
SSN_integrP_gorod: Feel close to their city, town	74.8	83.2	78.0	**77.7**	74.6	83.9	76.5	**77.4**
SSN_integrP_RF: Feel close to Russia	89.2	94.0	90.4	**90.6**	89.3	94.0	90.5	**90.8**
SSN_integrN_gorod: Do not feel close to their city, town	21.9	16.2	21.2	**20.6**	22.6	16.1	22.8	**21.3**
SSN_integrN_RF: Don't feel close to Russia	9.7	5.4	7.0	**7.8**	10.3	5.4	6.7	**7.7**
SSN_integrN_nofrends: No friends	7.8	7.2	10.1	**8.6**	6.0	6.7	8.8	**7.3**
SSN_integrP_5frends: Many, 5 or more friends	14.1	24.6	19.9	**18.5**	13.9	23.5	18.2	**17.8**
SSN_socialP_strana: Believe that things in the country are going in the right direction	68.4	67.1	62.4	**65.6**	70.2	67.8	62.8	**66.6**
SSN_socialN_strana: Think the country is moving down the wrong path	16.3	19.2	21.8	**19.1**	13.9	18.8	23.2	**18.8**
SSN_socialP_bezop: Feel completely safe	15.5	17.4	11.9	**14.3**	15.9	18.1	10.9	**14.3**
SSN_socialN_bezop: Feel completely unsafe	3.6	6.6	8.8	**6.3**	2.8	6.0	8.1	**5.7**
SSN_socialP_dover: Believe that most people can be trusted	21.9	23.4	21.5	**22.0**	22.2	22.8	22.5	**22.4**
SSN_socialN_dover: Cautious, don't trust most people	75.1	76.0	77.5	**76.3**	75.4	77.2	76.1	**76.1**

Source: authors' calculations based on PFS data.

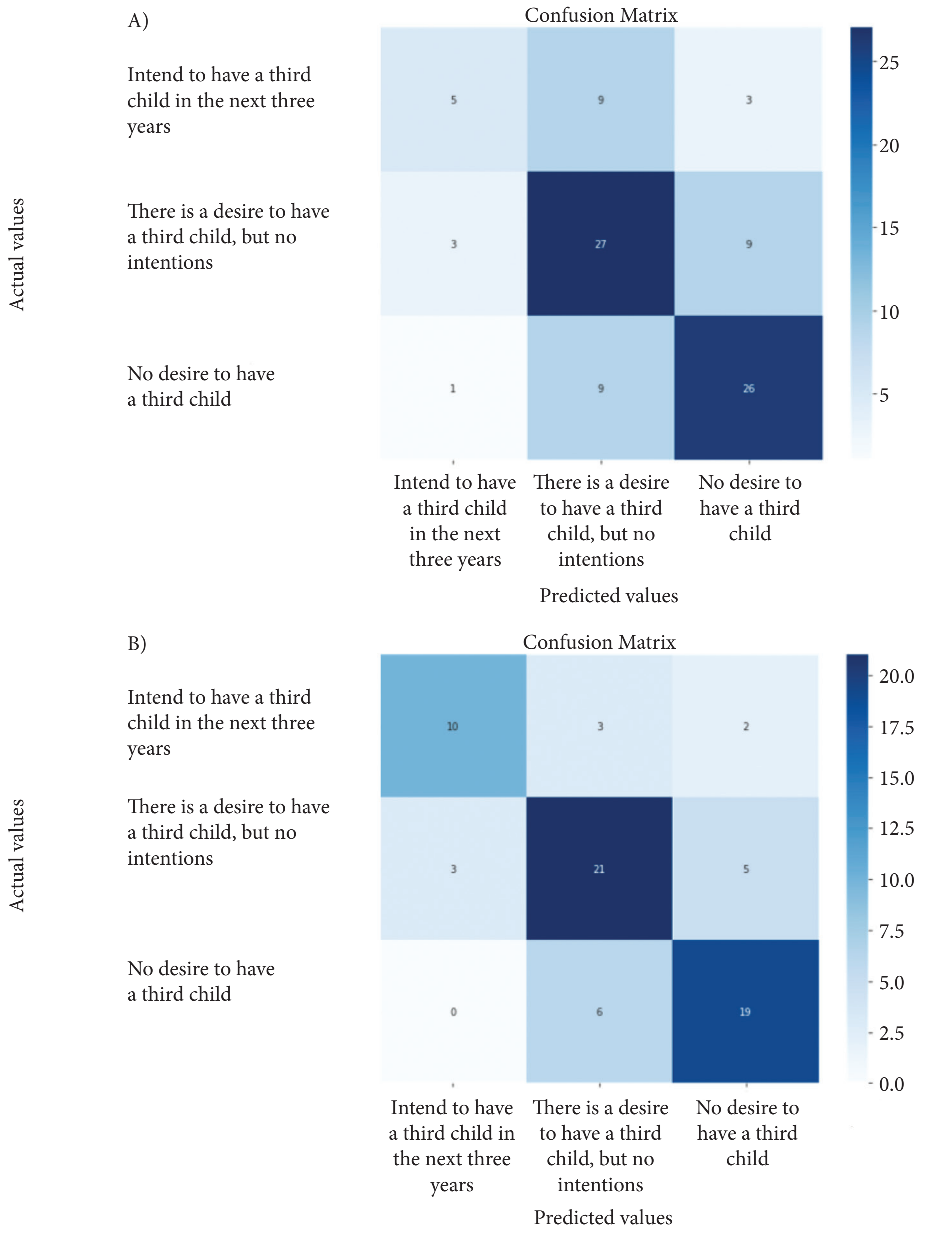


Figure A1. Matrices of predicted and actual indicators of target classes when modeling (A) on the extended and (B) on the main dataset, observations count divided by 10. *Source:* authors' calculations based on PFS data.

Name of the feature

male: Gender male
credits_potr: Paying off consumer loans
partn_wants_d: Spouse would like more children
housing_matters: believe that if you don't have your own home and a stable job, it's better to postpone having a child
SSN_integrP_odinok: Don't feel lonely
partn.age: age of spouse or partner
zh_strat4_d3: L3: No deprivations, but fewer rooms than family members
str5: Matpol5: One or two signs of wealth and no signs of poverty
first_marr: First official marriage, 1=yes
ageR_c: Age of respondent
housing_3yd1: Housing conditions will improve
abort_med: Abortion for medical reasons is considered acceptable
zh_strat4_d1: L1: Poor housing, two or more deprivations
yngch_age_7pd: Age of youngest child 7 years and older
yngch_age_36d: Age of the youngest child is 3-6 years
dom: Housing type - house
high_edu: Has a higher education
SSN_healthP_otvet: We completely agree that a person is fully responsible for his own health
relig0_nprac: Religiosity: non-practicing
babded: Grandparents help with children

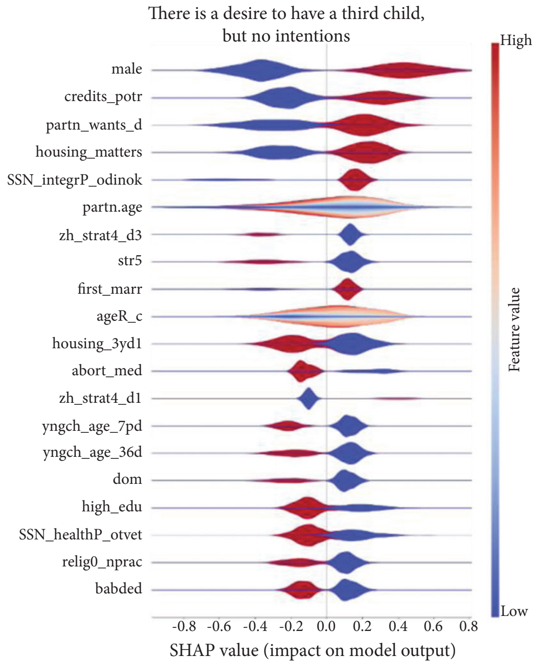


Figure A2. Factors operating in classification model 2, graph of SHAP values for the feature "there is a desire to have a third child, but no intentions". *Source:* authors' calculations based on PFS data.

Name of the feature

partn_wants_d: Spouse would like more children
male: Gender male
abort_med: Abortion for medical reasons is considered acceptable
abort_ever: consider abortion acceptable in any situation at the request of the woman
str5: One or two signs of poverty and no signs of wealth
partn.age: Age of spouse
dom: Housing type – house
credits_potr: Paying off consumer loans
yngch_age_7pd: Age of youngest child 7 years and older
ageR_c: Age of respondent
smoke: They smoke
housing_3yd1: Housing conditions will improve
zh_strat4_d3: L3: No deprivations, but fewer rooms than family members
hlth_1y_const: Health has not changed over the past year
SSN_socialN_strana: think the country is moving down the wrong path
SSN_integrP_odinok: Don't feel lonely
housing_matters: believe that if you don't have your own home and a stable job, it's better to postpone having a child
partn_highedu: Spouse has a higher education
babded: Grandparents help with children
career_3yd1: Career will improve

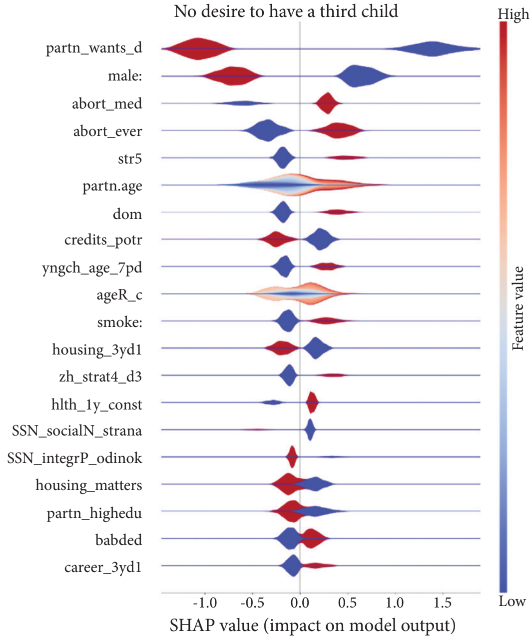
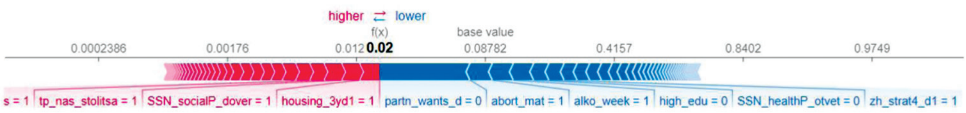
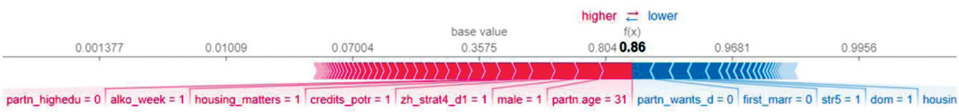


Figure A3. Factors operating in classification model 3, graph of SHAP values for the feature "no desire to have a third child". *Source:* authors' calculations based on PFS data.

1. Intend to have a third child in the next three years



2. There is a desire to have a third child, but no intentions



3. No desire to have a third child

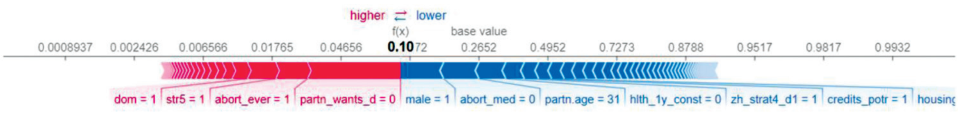


Figure A4. Numerical influence of factors on the probability of falling into the three classes predicted by the model for one of the respondents. *Source:* authors' calculations based on PFS data.