ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-6-271-287
UDC 349.232,330.43(045)
JEL E24, J31

Analysis of Household Income Dynamics in the Russia Based on the RLMS Database

E.V. Martyanova[a], A.V. Polbin[b]
[a, b] Institute of Applied Economic Research, RANEPA, Moscow, Russia;
[b] Gaidar Institute for Economic Policy, Moscow, Russia

ABSTRACT

The **goal** of the study is to estimate the parameters of the stochastic wage process using data from the Russian Longitudinal Monitoring Survey of Higher School of Economics (RLMS-HSE). The main **method** of analysis is econometric estimation, which includes two steps. In the first step, the authors estimated a Mincer-type regression. In the second step, they estimated the parameters of the stochastic wage process using the generalized method of moments. As a **result**, the autoregression coefficient turned out to be lower, and the variance of shocks was higher than in similar foreign studies. The results of the research allow to **conclude** that labor incomes in Russia are less stable over time and are marked by great uncertainty. The practical value of the work lies in the possibility of using the obtained estimates when calibrating general equilibrium models with heterogeneous agents, which is demonstrated in the framework of estimation of macroeconomic effects from hypothetical tax maneuvers based on the canonical model with heterogeneous agents.

Keywords: wages; labor income; labor income profile; RLMS; generalized method of moments; general equilibrium model; heterogeneous agents

For citation: **Martyanova E.V., Polbin A.V. Analysis of household income dynamics in the Russia based on the RLMS database. *Finance: Theory and Practice*. 2022;26(6):271-287. DOI: 10.26794/2587-5671-2022-26-6-271-287**

INTRODUCTION

Household income dynamics play an important role in decisions about consumption and savings, human capital accumulation. In the economic literature, household income trajectories are described as stochastic processes in which the key indicators are variance of shocks and autocorrelation. The first indicator characterises how big unexpected changes in income can be experienced by the household, the second indicator — how long income will be able to adjust to previous levels. With the risk of declining incomes in the future, the most important characteristic of consumer behavior is the precautionary motive in which households save for the rainy day. Of course, the precautionary motive is different for households with different levels of assets (which includes the entire history of income shocks), leading to a heterogeneous propensity to consume for households with different levels of accumulated wealth. This fact has stimulated the active development of dynamic stochastic models of general equilibrium with heterogeneous economic agents for the analysis of monetary and fiscal policy in business cycle analysis as opposed to traditional models with representative agents (see, for example, [1, 2]). General equilibrium models with heterogeneous agents, the key source of heterogeneity in which idiosyncratic uninsurable risk in income, were also built to analyse many other issues, namely social insurance [3], impact of economic policies on inequality and entrepreneurial sector [4, 5].

Despite the fact that income modelling is used in many areas of economic research, as far as we know, there are no such estimates on Russian data. In the [6] income decomposition was made only on permanent and uncorrelated transitive component, but the presence of autocorrelation was not considered. This work aims to fill this gap and give a wide range of estimates based on RMLS microdata.[1]

The results can be of practical value for calibrating general equilibrium models for the Russian economy with heterogeneous economic agents. The last section of the article describes the calibration and results of the general equilibrium model with heterogeneous agents under alternative tax policies: increase in VAT, increase in VAT with reduction of insurance premiums, increase in VAT with reduction of income tax.

The article is structured as follows. The second section provides an overview of the literature, the third section describes the data used, the fourth section presents the results of an empirical analysis. The last section presents a canonical general equilibrium model with heterogeneous agents as an example of the practical application of the estimates from the previous sections to calibrate models of a given class, on the basis of which the effects of hypothetical tax manoeuvres are Estimated.

REVIEW OF THE LITERATURE

The rejection of the assumption of full insurance, highlighted the idiosyncratic shocks that determine the economic choice. The assumption of incomplete markets where individuals are at risk of unemployment or changes in labor productivity allows the precautionary motive to be taken into account in general equilibrium models with heterogeneous agents [7–9]. The precautionary motive encourages agents to accumulate savings in the event of an unexpected loss of work or loss of productivity.

In models with heterogeneous agents, idiosyncratic risks of labor incomes follow the Markov chain of the first order. Taking this into account, the problem of a household with exogenous work in a heterogeneous agent economy can be formulated as follows:

$$V(\epsilon, a) = \max_{c,a'} \left[u(c) + \beta E \left\{ V(\epsilon', a') \mid \epsilon \right\} \right]$$

[1] Russian monitoring of the economic situation and health of the population of HSE Research Institute (RLMS HSE), conducted by the National Research University "Higher School of Economics" and LLC "Demoscope" with the participation Centre for Population of the University of North Carolina in Chapel Hill and the Institute of Sociology of the Federal Research and Sociology Center of the RAS. URL: https://rlms-hse.cpc.unc.edu and http://www.hse.ru/rlms (accessed on 11.11.2021).

$$\text{s.t. } a' = (1+r)a + w_\epsilon - c$$

$$a' \geq a_{\min}$$

$$\pi(\epsilon' \mid \epsilon) = \text{Prob}\{\epsilon_{t+1} = \epsilon' \mid \epsilon_t = \epsilon\},$$

where $V(\epsilon, a)$ — value function; $u(c)$ — utility of the agent dependent on consumption c; β — discount factor; ϵ — state variable for stochastic agent's income; a — asset stock of the agent, not lower than a_{min}; r — interest rates in the economy; w_ϵ — agent's labor income ϵ; $\pi(\epsilon' \mid \epsilon)$ — transition matrix showing the probability of transition from state ϵ to new state ϵ'.

The standard approach for calibrating data matrix $\pi(\epsilon' \mid \epsilon)$ used in heterogeneous agent models includes two steps: estimation of the continuous Markov process for the stochastic component of earnings and discretization of this process. The most commonly used discretization algorithms are presented in the works [10–12]. The labor income process is represented as fluctuations around the deterministic function of the observed variables, using the stochastic process of the residuals of the standard Mincer regression. Thus, the logarithm of income is generally described by the formula:

$$\ln W_{i,t} = \underbrace{g(t, X_{i,t} \ldots)}_{\text{characterization influence}} + \underbrace{[\alpha_i + \beta_i t]}_{\text{heterogeneous profile}} + \underbrace{[\int_{i,t} + \nu_{i,t}]}_{\text{stochastic component}}$$

$$\epsilon_{i,t} = \rho \epsilon_{i,t-1} + \eta_{i,t}, \quad (1)$$

where $\ln W_{i,t}$ — logarithm of the agent's labor; $g(\cdot)$ — function of deterministic variables (time period, demographic variables, etc.). Input parameters for discretization of a random process are estimates of autoregression coefficient ρ and variance σ_η^2.

Parameter estimates are obtained by minimizing the distance between elements of the empirical autocovariance matrix of the regression residuals and their theoretical counterparts, derived from the specification (1).

Single opinion on specification in the literature did not work out. In an earlier work [13] based on the Panel Study of Income Dynamics (PSID) the following specification was estimated:

$$y_{i,t} = \alpha_i + \epsilon_{i,t},$$

$$\epsilon_{i,t} \sim ARMA(p, q),$$

where $y_{i,t}$ — stochastic component of labor income.

The time series for both income and wages were stationary only in the first differences, indicating a random walk. The variance of individual effects α_i was insignificant. In addition, T. MaCurdy proposed a statistical test that pointed to a process without individual effects. The final specification did not include individual effects, and the shock $\epsilon_{i,t}$ описывался процессом ARMA (1, 2). was described by the ARMA process (1, 2). The autocorrelation estimates for income and wages were 0.974 and 0.975 respectively, which is quite close to one. In the future, the idea that the stochastic component of earnings is better described by the process of random walk was also supported by the conclusions of [14–16].

However, these findings are contested in the work [17], in which the author's two approaches to modeling the process of labor incomes are compared. The first is the restricted income profile (RIP). This approach assumes that individuals are exposed to strong and highly persistent income shocks. In doing so, they meet similar income profiles throughout their lives. The second is the heterogeneous income profile (HIP). This approach implies that individuals are exposed to less persistent income shocks and face individual income profiles throughout their lives. The theoretical motivation of the second approach is the notion of human capital, which implies differences in income levels among people with different abilities.

In this work, the logarithm of labor income described by the restricted income profile was specified by the following specification:

$$y_{i,j,t} = \alpha_i + \beta_i j + \epsilon_{i,j,t} + \mu_t \nu_{i,j,t},$$

$$\epsilon_{i,j,t} = \rho\epsilon_{i,j-1,t-1} + \phi_t \eta_{i,j,t}, \ \epsilon_{i,0,t} = 0,$$

where j — potential experience of an individual (age minus duration of education); β_i — slope coefficient with variance σ_β^2; and random heteroscedastic errors — $\mu_t \nu_{i,j,t}$ and $\phi_t \eta_{i,j,t}$.

The restricted income profile was described by specification without individual heterogeneity ($\sigma_\beta^2 = 0$). Individual income differentials are statistically and quantitatively significant according to PSID data for 1968–1993. The autoregression coefficient in the restricted income profile model was 0.99, and in heterogeneous model — 0.8. This means that at retirement age, 65 to 80% of income inequality is associated with heterogeneous individual effects. The article concludes that the estimation of the autocorrelation coefficient in the model with a restricted income profile is overestimated. The bias is caused by the fact that a person with too high or low income will systematically deviate from the average income profile. Then, in the proposed econometric model, these biases will be interpreted as the result of positive income shocks.

In addition, this article criticizes the test proposed in [13] for low power in relation to higher autocorrelation orders. Using the Monte Carlo method, the author showed that the test points to RIP for a process generated as HIP.

The paper [18] used administrative data from Germany on quarterly earnings of employees over the 27 years of their career. Estimates were given for those individuals who dropped out or had just graduated from school, as the covariance matrix for this educational group was comparable to data from the USA. T. Hoffmann suggested modification of the specification [17], which took into account the age effects of the covariance matrix. In a simple specification without heterogeneous feedback, i.e. when $\sigma_\beta^2 = 0$, autocorrelation coefficient ρ was equal to 0.98, in modified specifications estimates range from 0.8 to 0.9.

Most of the approaches described above are based on fairly simple time series models. Empirical studies suggest that household incomes are asymmetric and non-linear, so researchers also propose flexible models that take this evidence into account.

In [15] the authors rejected the assumptions about equally distributed independent income shocks in favor of models with conditional variance. In addition to ARCH effects, shocks depended on age, time, and unobserved heterogeneity. The authors also suggested that the incomes of people with different levels of education may be subject to different processes. According to the article conclusions, the unit root specification better describes the process of labor income, and the variance of shocks is state-dependent and heterogeneous among agents.

In the work [19] the heterogeneity of the whole process of income generation is assumed: in initial value, variance of shocks, MA and AR process parameters, deterministic trends, convergence rate, measurement error. The authors' results show that the variances of shocks and measurement errors differ significantly among employees and refute the hypothesis that the wage process has a unit root.

However, to be included in the structural model, it is important that the specification not only is flexible and good at describing the reality, but that its conclusions can be used without significantly complicating the general equilibrium model and increasing the calculation time. In [20] authors proposed a non-parametric model that could be used in macroeconomic models with heterogeneous agents. Instead of estimating and discretizing the Markov process, the authors proposed to evaluate the age-specific Markov chain directly from the data. The model was evaluated on PSID data and on synthetic data that correspond to tax data on income in the USA. The authors included the estimated age-specific matrix in the standard life cycle model and compared its conclusions with those of a similar model where the standard AR (1) process was assumed for wages. In comparison with the standard model, the method of estimation proposed by the authors results in greater inequality of consumption, which better corresponds to the real data. The left tail of the distribution looks more realistic than standard specification model. However, the proposed specification has the same disadvantage as the specification with AR (1): in a life cycle model without a bequest motive or entrepreneurial sector, the right tail of the generated distribution is still not sufficiently thick compared to the actual distribution.

In addition to the idiosyncratic risk, there may also be an aggregate risk in the general equilibrium model with heterogeneous agents. The problem is that panel data is usually not long enough. In [16] is offered method based on generalized method of moments, taking into account and addressing the macroeconomic history of family members. The stochastic earning process was modeled as ARMA (1, 1) with a regime-switching component in conditional variance. The results showed that the idiosyncratic risk of labor income is counter-cyclical. Autocorrelation estimates were high — from 0.94 to 0.96, and the variance of shocks during the recession is reduced by 75%: from 0.12 to 0.21.

Modern research also takes into account the nature of changes in income and distinguishes between external shocks and endogenous responses of individuals to them. So, in [21], the combined model is estimated, which takes into account wages, working hours, transitions between employment and unemployment, changing jobs. In addition to many conclusions about the reasons for the change of work, the authors find that the income of individuals, though stable, but still not characterized by random walk.

The following conclusions can be done from the literature review. Wage modeling is important for many areas of economic research. First, it is an important input for macroeconomic models with heterogeneous agents. Second, understanding the process and risks of income generation is important for household consumption research. In addition, some studies focus on the income process itself: finding the right specification, formulating stylized facts, identifying the main factors influencing household labor income.

EMPIRICAL PART

In the first step, the stochastic earnings process is disposes of the influence of demographic and geographical determinants. To do this, the Mincer regression of the logarithm of wages in real prices of a set of individual and geographical characteristics and binary variables for periods is estimated:

$$\ln W_{i,t} = X_{it}\beta + \theta\, Year_{it} + y_{i,t}\ , \qquad (2)$$

where $\ln W_{i,t}$ — logarithm of real labor income (hourly wage or wage for the last 30 days); X_{it} — set of individual (age, age squared, binary variables for level of education) and geographic characteristics; $Year_{it}$ — binary variable vector for each period from 2001 to 2019.

Residuals of this equation $y_{i,t}$ represent a stochastic component of labor productivity. It is assumed that it is the sum of three orthogonal components — individual effects α_i, autoregression shock $\epsilon_{i,t}$ and transitive shock $\nu_{i,t}$:

$$y_{i,t} = \alpha_i + \epsilon_{i,t} + \nu_{i,t}\,,$$
$$\epsilon_{i,t} = \rho\epsilon_{i,t-1} + \eta_{i,t}\,. \qquad (3)$$

This specification was used, for example, in the work [17] on the USA and [18] on data for Germany.

In this model $\eta_{i,t}$ and $\nu_{i,t}$ are taken from normal distributions with zero expectation and variances σ_η^2 and σ_ν^2 respectively. The initial autoregression shock $\epsilon_{i,0}$ — is a random variable distributed with zero mathematical expectation and variance σ_{ϵ_0}. Individual effects α_i have zero mathematical expectation and variance σ_α^2. In a model without individual effects, it is assumed that $\sigma_\alpha^2 = 0$.

The true value measurement errors cannot be directly estimated from the data, so researchers usually assume that it does not depend on time. Then the variance of measurement error enters the variance of transitory shocks $\nu_{i,t}$.

It can be shown that for specification (3) the theoretical second moments of the distribution are given by formula:

$$E\left[y_{i,t}y_{i,t+h}\right]=\begin{cases}\sigma_\alpha^2+E\left[\varepsilon_{i,t}^2\right]+\sigma_\nu^2, & if\ h=0\\ \sigma_\alpha^2+\rho^h E\left[\varepsilon_{i,t}^2\right], & if\ h>0,\end{cases} \quad (4)$$

where h — integers from 0 to $T-t$; T — maximum period under consideration; $E\left[\epsilon_{i,t}^2\right]=\rho^{2t}\sigma_{\epsilon_0}^2+\sum_{k=1}^{t}\rho^{2(t-k)}\sigma_\eta^2$.

Parameter estimation is made by minimizing the distance between empirical moments $\widehat{M}$ and their theoretical counterparts $M(\theta)$, given by (4):

$$\min_\theta\left[M(\theta)-\widehat{M}\right]^T W\left[M(\theta)-\widehat{M}\right].$$

Nelder-Mead algorithm was used to minimize. Newton, Broyden-Fletcher-Goldfarb-Shanno (BFGS), algorithms, the conjugate path method converged over more time, and the differences in the numerical optimum values were negligible. Since the confidence intervals for the parameters were based on a bootstrap, which requires multiple finding of the optimum, it was decided to favor the Nelder-Mead algorithm. The valuation suggested that matrix W — is the identity matrix because in [22] it is shown that the optimal matrix of weights W leads to bias of estimates.

Standard errors were calculated using a bootstrap with 500 replications. Each pseudo-sample included as many households (or individuals) as the original sample. First step regression was evaluated for each pseudo-sample. Stochastic process parameters (3) were estimated based on regression residuals (2). Confidence intervals were calculated from the derived distribution of estimates.

In general equilibrium models with heterogeneous agents, the labor supply can be either inelastic (exogenous) or elastic (endogenous). Depending on this, estimates should be made based on different time series: monthly or hourly wages. To complete the analysis, we will provide estimates based on several alternative specifications.

In the case of a household, it is not clear what is meant by its earned income. In some works, labor income refers to the income of the head of household, which can be defined differently. For example, in a work [23] also based on the RLMS, a household was sampled if it had at least one individual aged 25–60. The head of household was the oldest man of working age or the oldest woman if there were no men. This was the approach used in this work.

Thus, eight specifications are estimated that differ in the following characteristics:

- model specification: with or without individual effects;
- unit of observation: individual or head of household;
- dependent variable: hourly wages or wages for the last 30 days.

The estimations were based on RLMS data from 2001 to 2019. In order to form a representative sample for panel analysis, the latest 28th wave was selected as the baseline. Household data from representative sample of this wave were used for further analysis.

Two samples were selected for analysis: a household sample and an individual sample. Employed individuals were included in the sample of individuals. The lower age limit was 25 years, and the upper age limit was equal to the age of retirement: 60 years for men and 55 years for women. Households in which at least one individual is over 25 years of age and below retirement age. The head of household was the oldest man of working age. If not, the oldest woman of working age was appointed head. This approach was used in similar works on both the USA [24] and Russian [23] data.

The individual's wage was calculated as the sum of payments for the last 30 days of the main and additional work. Household data and individual data were used to construct a sample of households. Wages were converted to 2019 prices using the consumer price index (CPI). Hourly wage calculated as ratio of wages to hours worked over the last 30 days.

Data pre-processing and model estimation were carried out using the Python programming

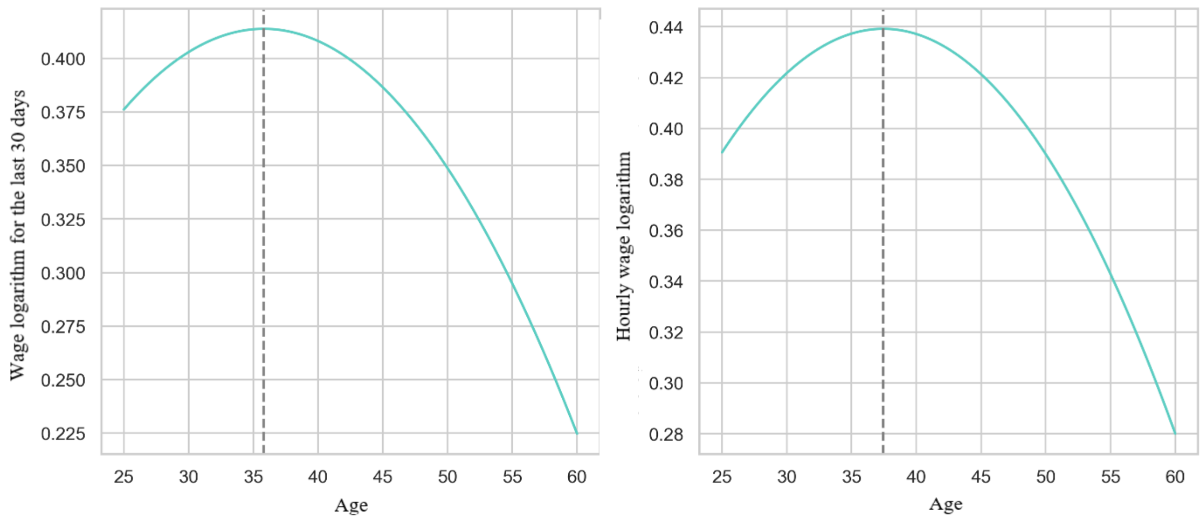

Fig. 1. **Wage dependence on age**

Source: Authors' calculations based on RLMS-HSE and Rosstat data.

language.[2] The Statsmodels library was used to build econometric models, NumPy — or matrix calculations, and SciPy — to find minima of functions.[3]

RESULTS

The results of the first step regression evaluation are presented in *Table 1*. All signs of coefficient are as expected and the estimates are statistically significant at any reasonable level of significance.

Other things being equal, the wages of individuals living in the city are 31–33% higher than those in rural areas. Monthly wages for residents of Moscow and the Moscow region are 55–57% higher than those of rural residents, and hourly wages are 61–62%.

Workers with higher education, all other things being equal, earn 35% per month and 41% per hour more than those with incomplete secondary education. In addition, the wages of workers with specialized secondary education are higher than those with secondary education. It is worth noting that estimates of binary variables for education levels are statistically higher for regressions in which the dependent variable is hourly wage. That is, the higher the level of education, the less people work. This may be due, for example, to the specificity of the work or to the fact that the income effect of an increase in the hourly wage is greater than the substitution effect — individuals prefer additional leisure hours to higher labor incomes.

Age and age squared coefficient estimates were significant, indicating a quadratic dependence. Based on the signs of coefficients, the wage depends nonlinearly on age — has the inverted U shape (*Fig. 1*). Depending on the specification, all things being equal, wages go up to 35 to 37 years, and then they fall. This result is consistent with the empirical evidence presented in [25] that this age-specific wage profile is a characteristic and stable Russian feature. While wages in foreign countries grow monotonously and at a slowing rate throughout life, in Russia the maximum wages reach up to 40 years. V. Gimpelson attributes this empirical fact to a lack of investment in human capital after completion of formal education and the resulting decline in cognitive abilities, deteriorating health, and declining in personal characteristics such as "openness to new experience".

The results of the estimation by the generalized method of moments are presented in *Table 2*. As can be seen from the results obtained, the addition of individual effects leads to a lower estimate

[2] Python 3.8.8 version.

[3] Statsmodels 0.12.2 version, NumPy 1.20.1 version, SciPy 1.6.2 version.

Table 1

Estimates of regression coefficients of real wages for demographic variables

Variable	Dependent variable: last 30 days' wages		Dependent variable: hourly wages	
	Individuals	Heads of households	Individuals	Heads of households
	(1)	(2)	(3)	(4)
Constant	8.181*** (0.075)	8.215*** (0.088)	3.010*** (0.076)	3.027*** (0.088)
Urban area	0.310*** (0.007)	0.330*** (0.008)	0.305*** (0.007)	0.330*** (0.008)
St Petersburg	0.355*** (0.018)	0.349*** (0.021)	0.366*** (0.019)	0.357*** (0.022)
Moscow and Moscow region	0.574*** (0.010)	0.555*** (0.011)	0.623*** (0.010)	0.610*** (0.011)
Higher education	0.346*** (0.0128)	0.354*** (0.014)	0.414*** (0.013)	0.411*** (0.014)
Secondary special education	0.077*** (0.013)	0.081*** (0.014)	0.119*** (0.013)	0.120*** (0.014)
Complete secondary education	0.036*** (0.013)	0.039*** (0.014)	0.056*** (0.013)	0.057*** (0.014)
Age	0.023*** (0.003)	0.024*** (0.004)	0.023*** (0.003)	0.024*** (0.004)
Square age	−0.0003*** (0.000)	−0.0003*** (0.000)	−0.0003*** (0.000)	−0.0003*** (0.000)
R^2	0.318	0.326	0.326	0.330
R^2_{adj}	0.318	0.325	0.326	0.330

Source: compiled by the authors based on regression analysis.

Notes: standard errors are shown in parentheses. Estimates for binary variables for each year are omitted. *** – significance at the 1% level.

of autoregression ρ in most of the proposed specifications, which, by and large, corresponds to the conclusions of work [17], although the difference is not so great. The autoregression coefficient does not change much depending on the dependent variable (hourly or wage for the last 30 days) and on the sample (all individuals or heads of households). In work with similar specifications made on the USA data, the autoregression coefficients are slightly higher: in [13, 14, 17] they were about 0.97–0.99, while in the *Table 2* estimates range from 0.89 to 0.93. It can be concluded that the stability of labor incomes in Russia is somewhat lower than in the USA.

Table 2

Estimates of the parameters of the stochastic component of wages

	Dependent variable: last 30 days' wages				Dependent variable: hourly wages			
	Individuals		Heads of households		Individuals		Heads of households	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ρ	0.931	0.886	0.931	0.886	0.931	0.886	0.903	0.916
	(0.922, 0.935)	(0.87, 0.893)	(0.923, 0.935)	(0.869, 0.893)	(0.923, 0.935)	(0.872, 0.892)	(0.894, 0.909)	(0.898, 0.926)
σ_ν^2	0.017	0.014	0.024	0.050	0.076	0.100	0.090	0.00
	(0.017, 0.018)	(0.014, 0.015)	(0.024, 0.025)	(0.05, 0.052)	(0.076, 0.079)	(0.1, 0.103)	(0.089, 0.093)	(0.000, 0.000)
σ_η^2	0.083	0.079	0.072	0.071	0.069	0.079	0.097	0.047
	(0.081, 0.086)	(0.077, 0.081)	(0.071, 0.075)	(0.069, 0.074)	(0.067, 0.071)	(0.077, 0.081)	(0.095, 0.1)	(0.046, 0.049)
σ_ϵ^2	0.017	0.079	0.100	0.086	0.052	0.036	0.021	0.016
	(0.017, 0.018)	(0.079, 0.081)	(0.1, 0.103)	(0.086, 0.088)	(0.052, 0.053)	(0.036, 0.037)	(0.021, 0.022)	(0.011, 0.011)
σ_α^2		0.079		0.086		0.057		0.100
		(0.078, 0.081)		(0.086, 0.088)		(0.057, 0.059)		(0.1, 0.103)
$\sigma_\alpha^2 = 0$	-	+	-	+	-	+	-	+

Source: compiled by the authors based on the estimation of the model by the generalized method of moments.

Note: The lower and upper bounds of the confidence intervals calculated with the bootstrap are given in parentheses.

As for shock variance σ_η^2, that the estimates of this variable are not too different depending on the specification. An exception is the case where the dependent variable is the hourly wage, and the sample includes heads of household. Compared to similar specifications estimated on the USA data, for example [17], the variance of shocks in Russia was higher. It can be assumed that households in Russia face greater uncertainty in labor income than households in the USA.

GENERAL EQUILIBRIUM MODEL

This section presents a neoclassical general equilibrium model based on a model from [26] extended by the introduction of social security contributions and lump-sum taxes (transfers) into the model. The model has three sectors: firms, households and the government. Calibration of the model is described, which uses estimates from previous sections. There is no aggregate uncertainty in the model. Households maximize utility by making decisions about their supply of labor and consumption in an elastic-supply model, and only in a consumption — in the model variant with an inelastic labor supply defined exogenously. At the same time, households are exposed to idiosyncratic shocks to productivity and employment that cannot be insure. Firms maximize profits and produce goods using labor and capital as factors of production. The government collects revenue from taxes and spends it on government consumption and unemployment benefits. The government balances the budget with lump-sum taxes or transfers to households.

HOUSEHOLD

The economy has a continuum of households with a measure of one. They differ in employment status, labor roductivity ϵ and asset stock k. Productivity follows a first-order Markov process with the transition matrix $\pi(\epsilon'|\epsilon)$, which specifies the probability of transition from the current state ϵ to the new state ϵ'. In a model with an inelastic supply of labor, the household maximizes its utility and decides only on current consumption, taking into account the current stock of capital k and the productivity ϵ, of labor, in a model with elastic labor — also relatively the supply of labor n.

The instantaneous utility function of households is a function with constant relative risk aversion (CRRA):

$$u(c_t,n_t)=\frac{c_t^{1-\sigma}-1}{1-\sigma}-\gamma_0\frac{n_t^{1+\gamma_1}}{1+\gamma_1},$$ where c_t — household consumption during t; n_t — hours worked; σ — elasticity marginal utility by consumption (relative measure of risk aversion); γ_1 — parameter reflecting the elasticity of hours worked to the wage rate (inverse value to labor supply elasticity by Frisch); γ_0 — normalization coefficient.

The household budget constraint is:

$$k'=-T_{lump\ sum}+(1+r)k+(1-\tau_{income})w\,\epsilon\,n-$$
$$-(1+\tau_{consumption})c+I_{state=unemployed}b,$$

where k' — household asset stock in the next period; k — stock of assets in the current period; $T_{lump\ sum}$ — lump-sum taxes; r — real interest rate; τ_{income} — labor income tax; w — hourly rate of individual wages; ϵ — Individual productivity as measured by changes in hourly wage rates and described by the Markov chain of the first order; $\tau_{consumption}$ — consumption tax; $I_{state=unemployed}$ — indicator function equal to one if the agent is unemployed; b — unemployment benefit.

FIRM

Firms maximize profits based on labor and capital demand:

$$\max_{N_t,K_t}\left\{F(K_t,N_t)-(1+\tau_{insurance})w_tN_t-R_{K,t}K_t\right\},$$

where $F(K_t,N_t)$ — production function; K_t — total capital stock in the economy; N_t — cost effective labor; w_t — wage; $\tau_{insurance}$ — social security contribution rate paid by the firm per employee. User capital costs are defined as

$R_{K,t}=\frac{r_t}{1-\tau_K}+\delta$, where τ_K — rental tax on capital (counterpart of profit tax in the model); r_t — real interest rate; δ — rate of capital depreciation.

Production is described by the Cobb-Douglas function:

$$F(K,N)=AK^{\alpha}N^{1-\alpha}.$$

In equilibrium, firms' profits are zero and factor prices are equal to their marginal product:

$$r_t=(1-\tau_K)\left(\alpha A\left(\frac{N_t}{K_t}\right)^{1-\alpha}-\delta\right),$$
$$w_t=\frac{(1-\alpha)A}{1+\tau_{insurance}}\left(\frac{K_t}{N_t}\right)^{\alpha}.$$

GOVERNMENT

The government has a balanced budget in each period. The government revenue consists of lump-sum taxes $T_t^{lump\ tax}$, income taxes T_t^K, defined as a rental tax on capital for firms τ_K, final consumption taxes $T_t^{consumption}$, employee insurance payments $T_t^{insurance}$. Expenses consist of government consumption of G_t, which is a fixed share γ_g of the output, and unemployment benefits B_t:

$$G_t+B_t=T_t^{lump\ tax}+T_t^K+T_t^{income}+T_t^{consumption}+T_t^{insurance}.$$

The government balances the budget with lump-sum taxes (transfers).

STATIONARY EQUILIBRIUM

Stationary equilibrium in the proposed model for a given tax policy is the value function $V(\epsilon,k)$, set of decision rules for consumption $c(\epsilon,k)$, labor supply $n(\epsilon,k)$ and capital stock in the next period $k'(\epsilon,k)$, time-invariant prices for labor w and capital r, as well as set of variables K,N,B,T,C such as:

1. Total capital stock and volume of efficient labor, consumption, tax revenues and unemployment benefits are calculated as:

$$K=\sum_{\epsilon\in\mathcal{E}}\int_0^{\infty}kf(\epsilon,k)dk,$$

$$N=\sum_{\epsilon\in\mathcal{E}}\int_0^{\infty}\epsilon n(\epsilon,k)f(\epsilon,k)dk,$$

$$C=\sum_{\epsilon\in\mathcal{E}}\int_0^{\infty}c(\epsilon,k)f(\epsilon,k)dk,$$

$$T=T_{lump\ tax}+\tau_K\left(F(K,N)-(1+\tau_{insurance})wN-\delta K\right)+$$
$$+\tau_{income}wN+\tau_{consumption}C+\tau_{insurance}wN,$$

$$B=\int_0^{\infty}bf(\epsilon_0,k)dk.$$

2. $c(\epsilon,k),k'(\epsilon,k),n(\epsilon,k)$ — optimal decision functions of the household, which are the solution to its problem.
3. Labor and capital prices equal marginal productivities.
4. The goods market clears.
5. The government budget is balanced.
6. Distribution of household variables is stationary.

PARAMETER CALIBRATION

To construct the transition matrix $\pi(\epsilon'|\epsilon)$, the estimations of the stochastic process of wages, obtained in the previous sections of the article, were used. Thus, for a model with an inelastic supply of labor, estimates ρ and σ_η^2 were taken from a column (4) (*Table 2*), and for a model with elastic labor supply — from a column (8). The estimates ρ and σ_η^2 served as input parameters for the discretization algorithm proposed in [12], which produces a productivity distribution and a transition matrix for employed workers at the output.

Since the estimates in the previous sections were given conditional upon the employment of the head of household, the matrix also takes unemployment separately. The productivity of the unemployed was equal to zero. Two facts were taken into account in the simulation of transition from employment to unemployment. First, the long-term unemployment rate was assumed to be 6%. Second, the average time taken to find a job was 6 months.[4] In addition, it has been suggested that households at all levels of productivity are equally likely to be unemployed. If an unemployed person finds a job, he falls into one of the classes with a probability proportional to the stationary distribution of employed workers.

The final distribution of productivity was normalized so that in the stationary state the sum of the productivity of the work was equal to one. As a result, for models with inelastic and elastic labor supply, productivity in different states was defined as follows:

$$\mathcal{E}_m=\{0,0.2823,0.5044,0.9011,1.6098,2.8759\},$$

$$\mathcal{E}_h=\{0,0.2894,0.5121,0.906,1.6029,2.8359\}.$$

The unemployment benefit in the model was set at 33.75% of the wage since in Russia the unemployed receive a benefit of 75% of their previous earnings for the first three months and 60% for the next three months, not more than six months.

The selected elasticity coefficients of marginal utility of consumption σ and output elasticity of capital α in the Cobb-Douglas function are standard in the literature that uses general equilibrium models. The depreciation rate δ was based on the average depreciation rate of fixed capital in Russia.[5]

The discount rate β is chosen so that the real interest rate in the stationary state in the base model with inelastic labor supply is 5%, which corresponds to the real interest rate in Russia in

[4] Federal State Statistics Service. Employment and unemployment in the Russian Federation in October 2020. URL: https://rosstat.gov.ru/storage/mediabank/wFtwOnek/zanyatost-i-bezrabotitsa.pdf (accessed on 11.11.2021).

[5] Penn World Table version 10.0. URL: https://www.rug.nl/ggdc/productivity/pwt/?lang=en (accessed on 17.11.2021).

Table 3

Calibration of general equilibrium model parameters

Parameter value	Description
$\beta = 0{,}94$	Discount factor
$\sigma = 2$	Elasticity of the marginal utility of consumption
$\gamma_0 = 1{,}57, \gamma_1 = 4$	Parameters of disutility of labour
$A = 0{,}63$, $\alpha = 0{,}35$	Total factor productivity and capital elasticity coefficient in Cobb-Douglas functions with constant returns
$\delta = 0{,}03$	Capital depreciation rate
$\gamma_g = 0{,}2$	Share of government expenditure in output

Source: compiled by the authors.

recent years.[6] Coefficient A was selected so that the interest rate is targeted and $F(K, L) = 1$.

Along with a specification with an inelastic labor supply, we are considering a scenario with an elastic labor supply in which the Frisch elasticity of labor supply is calibrated at a sufficiently moderate level of 0.25, which was used in similar works [27]. An estimate of the Frisch elasticity of labor supply proposal for married women of 0.16 was obtained from [28] based on RLMS-HSE data. In the work [29] the estimations of the labor supply elasticity were negative. Under these conditions, the selection of an elasticity value of 0.25 seems appropriate in order to compare model results with the case where the labor supply elasticity is zero. The chosen value agrees with the notion that in the Russian economy there is a low elasticity of the labor supply on wages.

Normalization coefficient γ_0 was selected based on the first-order condition for on labor supply:

$$-\frac{u_n^{'}(c,n)}{u_c^{'}(c,n)} = \frac{(1-\tau_{income})\epsilon w}{(1+\tau_{consumption})},$$

i.e.

$$\gamma_0 = \frac{(1-\tau_{income})\epsilon w c^{-\sigma}}{(1+\tau_{consumption})n^{\gamma_1}},$$

where w and c were used out of equilibrium of the base model with an inelastic labor supply, ϵ assumed to be equal to one, taking into account the productivity ratios, the labor supply of n was also equal to one.

The share of government expenditure in output was assumed to be 0.2 according to statistics on GDP by expenditure.[7]

The final calibrated parameters are presented in *Table 3*.

DESCRIPTION OF THE EXPERIMENTS

After the model calibration stage, we compare the stationary equilibrium parameters for the four tax policy options:

1. Initial parameters of tax policy: final consumption tax (value added tax, VAT) — 18%, rate of insurance premiums 30%, profit tax rate (rental tax on capital) — 20%.

2. Increase in final consumption tax rate to 20%, which occurred in the Russian Federation in 2019.

3. Increase of the final consumption tax

[6] The World Bank. Real Interest Rate (%) — Russian Federation. URL: https://data.worldbank.org/indicator/FR.INR.RINR?locations=RU (accessed on 08.11.2021).

[7] Federal State Statistics Service. On the production and use of gross domestic product (GDP) for 2020. URL: https://www.gks.ru/bgd/free/B 04_03/IssWWW.exe/Stg/d02/18.htm (accessed on 08.11.2021).

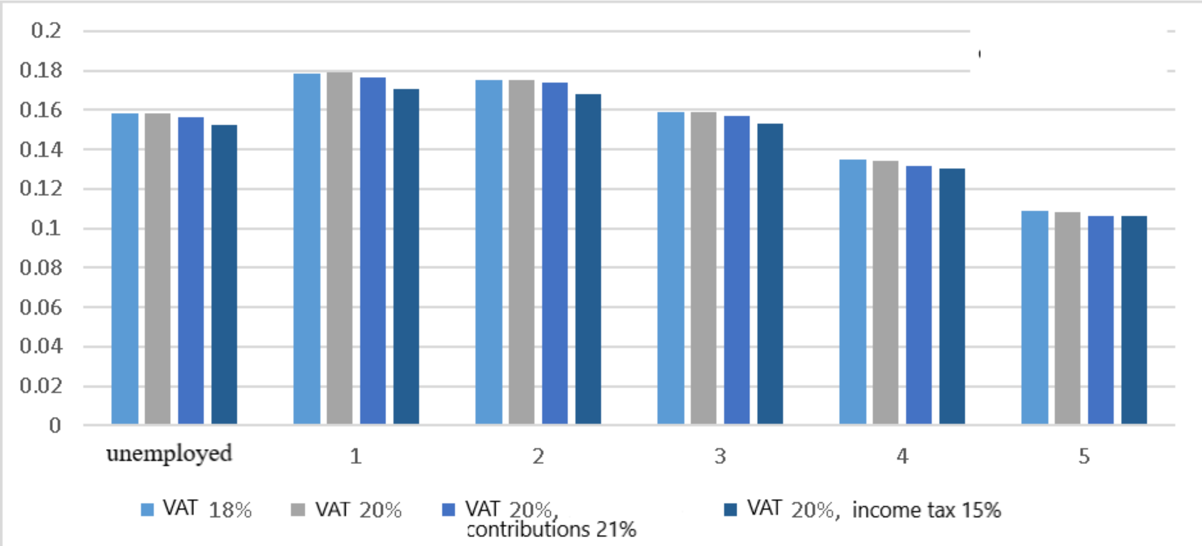

Fig. 2. **Share of consumption in household expenditures depending on tax policy**

Source: author's calculations.

rate to 20% while reducing social security contribution rate to 21% in the context of the option of tax maneuver corresponding to "fiscal devaluation" discussed at the time [30].

4. Increase the final consumption tax rate to 20%, while reducing the profit tax (rent tax) rate to 15%.

RESULTS

The stationary equilibrium of the model was calculated by solving the Bellman equation by value function iteration and finding stationary equilibrium in the asset market. Calculations were made using the Python programming language using NumPy libraries, as well as QuantEcon and Numba to speed up the calculations. Results are presented in *Table 4*.

When the VAT rate increases and other taxes remain unchanged, tax revenues increase, which, according to the construction, increase the lump-sum transfers. Since transfers are lump-sum, all individuals have the same increase in transfers in absolute terms. For rich individuals, however, the increase in VAT tax is greater in absolute terms than for poor people because they spend more resources on consumption. This scenario therefore shows a decline in inequality. In a model with an inelastic labor supply, although aggregate labor is fixed, there is nevertheless a slight decline in output due to the decline in capital in the economy. This is because the rich invest more of their income and consume less (*Fig. 2*), and this tax manoeuvre redistributes aggregate income to the poor, which reduces aggregate capital. In the case of elastic labor supply, the increase in VAT increases the distortion of the limit for the substitution of leisure consumption, which leads to a reduction in the supply of labor. Accordingly, in the model with endogenous labor we see a more significant decline in output and capital in the economy.

If, along with the increase in VAT from 18 to 20%, social security contribution rate is reduced from 30 to 21%, budget revenues will fall against the base scenario, which will consequently cause a decrease in transfers and increase in inequality. This will increase capital as rich individuals are more likely to save and thus increase output (by about 1%). In the specification with endogenous labor this will also increase the hours worked (about 8%), as the reduction of social security contribution rate firms will increase demand for labor , which will have an additional positive impact on capital accumulation and output (output increases in this scenario by about 1.5). The effects are quite moderate as we have

Table 4

Stationary equilibrium in the model depending on tax policy

	Inelastic Labor Supply				Elastic Labor Supply			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Insurance premiums, %	30	30	21	30	30	30	21	30
Income tax, %	20	20	20	15	20	20	20	15
VAT rate, %	18	20	20	20	18	20	20	20
Output index	100	99.82	100.95	101.88	100	99.58	101.55	101.50
Consumption index	100	99.85	100.69	101.29	100	99.62	101.38	100.95
Index capital	100	99.47	102.74	105.45	100	99.29	103.19	105.02
Labour input index	100	100	100	100	100	99.75	100.67	99.65
Interest rate, %	5.09	5.12	4.96	5.14	4.97	4.99	4.85	5.02
Wage index	100	99.82	108.46	101.88	100	99.84	108.37	101.85
Income tax benefit index	100	99.82	108.46	101.88	100	99.58	109.10	101.50
Index of VAT payments	100	110.94	111.88	112.54	100	110.68	112.64	112.16
Profit tax payment index	100	99.98	100.11	75.14	100	99.73	100.75	74.85
Insurance premium payment index	100	99.82	75.92	101.88	100	99.58	76.37	101.50
Compensation transfer index	100	107.65	89.40	102.10	100	107.54	89.98	101.95
Gini coefficient	0.244	0.241	0.253	0.244	0.246	0.242	0.254	0.245

Source: compiled by the authors.

suggested a low value for labor supply elasticity.

In the case of a simultaneous increase in VAT and a reduction in income tax, the increase in tax revenue relative to the base scenario is still positive, which contributes to a negligible reduction of inequality through a small increase in lump-sum transfers. Reducing the user cost of capital due to the reduction of income tax acutely stimulates capital accumulation, which increases by about 5%, leading to an increase in output of about 1.5%. Households become richer and, in endogenous specifications, aggregate labor use declines despite real wage growth, as the income effect exceeds the substitution effect.

CONCLUSION

The paper attempted to provide a broad range of estimates of the stochastic wage process based on RLMS microdata. The estimation included two steps. In the first step the regression of wages of Mincer type was estimated. In the second step, the autocovariance matrix was estimated on the basis of the regression residuals, and then on the basis of this, estimates of the parameters of the stochastic component of wages were made using the generalized method of moments.

Signs of coefficients of Mincer regression obtained in the first step correspond to theoretical representations. Age dependency of wages takes the inverted U shape: wages rise to the age of 35–37 and then fall. With regard to education, the monthly salary of university graduates is 35% higher than that of those who have not completed secondary education and — 41% higher by the hour. Graduates of secondary institutions receive more students who have completed secondary education.

With regard to the estimation of the stochastic component of wages, depending on the specification and the sample, the estimates of the autoregression coefficient vary between 0.89 and 0.93, which is slightly lower than the estimates from similar specifications in the USA data. This suggests that labor incomes in Russia are less stable. In addition, variance of shocks of the stochastic income component is higher for Russian households, which leads to the conclusion that they face greater uncertainty in labor incomes.

The results obtained can be used for the calibration of models with heterogeneous agents. Estimates of autoregression parameters ρ and variance of shocks η are input parameters for discretizing the AR (1) process sampling using methods suggested in the works [10, 11].

The last section of the article presented an example of using the econometric estimates obtained on the example of the model with heterogeneous agents, calibrated with Russian data. On the basis of the theoretical model, alternative tax policies were compared: increase in VAT rate, increase in VAT rate with simultaneous reduction of the rate of insurance payments per employee, increase of VAT rate with simultaneous reduction of income tax.

According to the model, the increase in the VAT rate led to a decrease in total output, consumption, wages, capital stock, and in a model with an elastic labor supply, and a decrease in the use of efficient labor. However, economic inequality in this model was the smallest.

Higher VAT rates and lower rates of insurance premiums have led to an increase in total output, consumption, wages, capital stock, and in a model with elastic labor supply — and the cost of efficient labor. However, inequality in both variants of the model was the highest, which could be explained by the decline in lump-sum transfers to households in the context of declining tax revenues.

An increase in the VAT rate to 20% and the simultaneous reduction of income tax to 15% led to an increase in output, consumption, capital and wages. However, in a model with an elastic labor supply, the volume of efficient labor decreased by 0.35% due to the prevalence of income effect over substitution effect in labor supply, inequality in the model decreased slightly compared to the basic tax policy option.

ACKNOWLEDGEMENTS

Research on the topic “Development of a set of general equilibrium models with heterogeneous economic agents for the Russian economy” was carried out at the expense of the grant of the Russian Science Foundation No. 21–78–10020. Russian Presidential Academy of National Economy and Public Administration (RANEPA), Moscow, Russia.

REFERENCES

1. Kaplan G., Moll B., Violante G.L. Monetary policy according to HANK. *American Economic Review.* 2018;108(3):697–743. DOI: 10.1257/aer.20160042
2. Heathcote J. Fiscal policy with heterogeneous agents and incomplete markets. *The Review of Economic Studies*. 2005;72(1):161–188. DOI: 10.1111/0034–6527.00328
3. Nishiyama S., Smetters K. Does social security privatization produce efficiency gains? *The Quarterly Journal of Economics*. 2007;122(4):1677–1719. DOI: 10.1162/qjec.2007.122.4.1677
4. Quadrini V. Entrepreneurship, saving, and social mobility. *Review of Economic Dynamics*. 2000;3(1):1–40. DOI: 10.1006/redy.1999.0077
5. Bassetto M., Cagetti M., De Nardi M. Credit crunches and credit allocation in a model of entrepreneurship. *Review of Economic Dynamics*. 2015;18(1):53–76. DOI: 10.1016/j.red.2014.08.003
6. Koval P., Polbin A. Evaluation of permanent and transitory shocks role in consumption and income dynamics in the Russian Federation. *Prikladnaya ekonometrika=Applied Econometrics*. 2020;(1):6–29. (In Russ.). DOI: 10.22394/1993–7601–2020–57–6–29
7. Imrohoroğlu A. Cost of business cycles with indivisibilities and liquidity constraints. *Journal of Political Economy*. 1989;97(6):1364–1383. DOI: 10.1086/261658
8. Huggett M. The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control*. 1993;17(5–6):953–969. DOI: 10.1016/0165–1889(93)90024-M
9. Aiyagari S.R. Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics*. 1994;109(3):659–684. DOI: 10.2307/2118417
10. Tauchen G. Finite state markov-chain approximations to univariate and vector autoregressions. *Economics Letters*. 1986;20(2):177–181. DOI: 10.1016/0165–1765(86)90168–0
11. Tauchen G., Hussey R. Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. *Econometrica.* 1991;59(2):371–396. DOI: 10.2307/2938261
12. Rouwenhorst K.G. Asset pricing implications of equilibrium business cycle models. In: Cooley T.F., ed. Frontiers of business cycle research. Princeton, NJ: Princeton University Press; 1995:294–330. DOI: 10.1515/9780691218052–014
13. MaCurdy T.E. The use of time series processes to model the error structure of earnings in a longitudinal data analysis. *Journal of Econometrics*. 1982;18(1):83–114. DOI: 10.1016/0304–4076(82)90096–3
14. Abowd J.M., Card D. On the covariance structure of earnings and hours changes. *Econometrica.* 1989;57(2):411–445. DOI: 10.2307/1912561
15. Meghir C., Pistaferri L. Income variance dynamics and heterogeneity. *Econometrica*. 2004;72(1):1–32. DOI: 10.1111/j.1468–0262.2004.00476.x
16. Storesletten K., Telmer C.I., Yaron A. Cyclical dynamics in idiosyncratic labor market risk. *Journal of Political Economy*. 2004;112(3):695–717. DOI: 10.1086/383105
17. Guvenen F. An empirical investigation of labor income processes. *Review of Economic Dynamics*. 2009;12(1):58–79. DOI: 10.1016/j.red.2008.06.004
18. Hoffmann F. HIP, RIP, and the robustness of empirical earnings processes. *Quantitative Economics*. 2019;10(3):1279–1315. DOI: 10.3982/QE 863
19. Browning M., Ejrnæs M., Alvarez J. Modelling income processes with lots of heterogeneity. *The Review of Economic Studies*. 2010;77(4):1353–1381. DOI: 10.1111/j.1467–937X.2010.00612.x
20. De Nardi M., Fella G., Pardo G.P. The implications of richer earnings dynamics for consumption and wealth. NBER Working Paper. 2016;(21917). URL: https://www.nber.org/system/files/working_papers/w21917/w21917.pdf
21. Altonji J.G., Smith A.A., Vidangos I. Modeling earnings dynamics. *Econometrica*. 2013;81(4):1395–1554. DOI: 10.3982/ECTA8415
22. Altonji J.G., Segal L.M. Small-sample bias in GMM estimation of covariance structures. *Journal of Business & Economic Statistics*. 1996;14(3):353–366. DOI: 10.2307/1392447
23. Gorodnichenko Y., Peter K.S., Stolyarov D. Inequality and volatility moderation in Russia: Evidence from micro-

level panel data on consumption and income. *Review of Economic Dynamics*. 2010;13(1):209–237. DOI: 10.1016/j.red.2009.09.006

24. Heathcote J., Perri F., Violante G.L. Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006. *Review of Economic Dynamics*. 2010;13(1):15–51. DOI: 10.1016/j.red.2009.10.010
25. Gimpelson V. Age and wage: Stylized facts and Russian evidence. *Ekonomicheskii zhurnal Vysshei shkoly ekonomiki = The HSE Economic Journal*. 2019;23(2):185–237. (In Russ.). DOI: 10.17323/1813–8691–2019–23–2–185–237
26. Heer B., Trede M. Efficiency and distribution effects of a revenue-neutral income tax reform. *Journal of Macroeconomics*. 2003;25(1):87–107. DOI: 10.1016/S 0164–0704(03)00008–9
27. Heathcote J. Fiscal policy with heterogeneous agents and incomplete markets. *The Review of Economic Studies*. 2005;72(1):161–188. DOI: 10.1111/0034–6527.00328
28. Zamnius A., Polbin A.B. Estimating intertemporal elasticity of substitution of labor supply for married women in Russia. *Prikladnaya ekonometrika = Applied Econometrics*. 2021;(4):23–48. (In Russ.). DOI: 10.22394/1993–7601–2021–64–23–48
29. Klepikova E. Labor supply elasticity in Russia. *Voprosy ekonomiki*. 2016;(9):111–128. (In Russ.). DOI: 10.32609/0042–8736–2016–9–111–128
30. Sokolov I. Is there a need for fiscal devaluation to spur economic growth? *Ekonomicheskoe razvitie Rossii = Russian Economic Development*. 2017;24(6):13–18. (In Russ.).

ABOUT THE AUTHORS

Elizaveta V. Martyanova — Jun. Researcher, Institute of Applied Economic Research, Russian Presidential Academy of National Economy and Public Administration (RANEPA), Moscow, Russia
https://orcid.org/0000-0001-6389-8610
Corresponding author
martyanova-ev@ranepa.ru

Andrei V. Polbin — Can. Sci. (Econ.), Head of the Laboratory of Mathematical Modeling of Economic Processes, Institute of Applied Economic Research, Russian Presidential Academy of National Economy and Public Administration (RANEPA), Moscow, Russia; Deputy Head of the International Laboratory for Mathematical Modeling of Economic Processes, Gaidar Institute for Economic Policy, Moscow, Russia
https://orcid.org/0000-0003-4683-8194
apolbin@iep.ru

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

The article was submitted on 25.12.2021; revised on 14.01.2022 and accepted for publication on 17.05.2022.
The authors read and approved the final version of the manuscript.