# Earnings inequality and informal employment in Russia[1]

Anna Lukiyanova

Centre for Labor Market Studies, Higher School of Economics, Moscow.

E-mail: alukyanova@hse.ru.

In this paper, we explore the relationship between informality and earnings inequality using the RLMS-HSE data for 2000-2010. We find that during the whole period earnings inequality was substantially higher in the informal sector. Informality increases the earnings polarization widening both tails of the wage distribution. Nonetheless, inequality has declined in both formal and informal sectors. In the formal sector, changes in the distribution of monthly earnings between 2000 and 2010 were mainly generated by changes in the distribution of hourly earnings. In the informal sector, reduction of variation in monthly earnings went through two channels: declining differences in hourly rates and considerable compositional shifts within the informal sector. The results point at importance of distributional analysis of earnings gaps and explicit accounting for the sector choice.

Keywords: earnings inequality, informal economy, decomposition, recentered influence functions

JEL-classification: C21, D63, J31, J42.

## 1. Introduction

In recent years, there has been a growing interest in earnings inequality in transition and emerging economies following the evidence of rapidly increasing earnings dispersion during the early reform period. Previous studies of the rise in earnings inequality have identified several causes of this trend: the rise of returns to education, growth of the private sector, and failure of labor market institutions (e.g. Rutkowski, 1996; Brainerd, 1998; Milanović, 1999; Mitra and Yemtsov, 2006). Less attention has been paid to the expansion of informal economy. Meanwhile, the share of the informal sector has grown since the start of the economic transition in many post-communist countries. Informality can be one of the 'unnoticed' important sources of earnings inequality. Only few papers explicitly relate inequality and informality in the context of transition. Rosser et al. (2000) find a positive relationship between income inequality and the

[1] This article is part of the project 'Informality in the Russian Labor Market' funded by the HSE Research Programme. I wish to thank Rostislav Kapeliushnikov, Vladimir Gimpelson, participants of the 1st International Russia Longitudinal Monitoring Survey of HSE User Conference and two anonymous referees for helpful comments and suggestions.

size of the informal sector in transition countries. Krstic and Sanfey (2007, 2011) report significant impact of informality on earnings inequality in Serbia and in Bosnia and Herzegovina.

The understanding of the nature of informality has developed over time causing the changes in the expected effects of the informal sector on inequality. According to the traditional paradigm labor markets are segmented and workers are rationed out of the formal sector because of excessive regulation (Harris and Todaro, 1970). In this view informality is involuntary and incomes of informal workers are only marginally higher than of those in unemployment. Wages in the formal sector are driven up to the above market-clearing level because of minimum wages, trade unions, or efficiency wages. The pay gap between the sectors pushes up the overall inequality while within the informal sector inequality can be quite modest. Some authors later criticized the traditional view of the informal sector arguing that markets are integrated and majority of informal jobs reflect the voluntary choice. Informal jobs can fit workers' preferences and their skill endowments and offer better earnings prospects (De Soto, 1989; Maloney, 2004). In the case of competitive markets it means that in the long run the formal-informal wage gaps should be equal to the compensating wage differentials. This model suggests the smallest impact of informality on overall inequality. Recent developments combine these two views into a dual theory of the informal sector where a voluntary entry upper-tier segment coexists with an involuntary lower-tier segment (Fields, 2005). In this model, the contribution of informality comes largely from higher inequality within the very heterogeneous informal sector while between-sector inequality is relatively small. The effect of institutions is also uncertain for the informal sector. Generally, stronger labor market institutions are associated with lower earnings inequality (e.g. DiNardo et al, 1996; Fortin and Lemieux, 1997). However, this inequality-reducing effect is likely to exist only in the formal sector. In the informal sector, non-compliance is higher and many workers are out of the reach of legislation (no work contracts, flexible working hours, wages below the legal minimum, etc.). On the other hand, some regulations were confirmed to have positive spillover effects, resulting in wage compression in both sectors. For example, in Brazil minimum wage increases between 1984 and 2002 were found to have decreased wage inequality in the formal and informal sectors (Lemos, 2009).

The informal sector is not a new phenomenon in transition countries. It was present in centrally planned economies in the form of hidden transactions of goods or services that were privately produced or sold by individuals and in the form of hidden activities within large enterprises. Individual informal activities were quite common in a situation of pervasive goods shortages and widespread theft from the state enterprises. Enterprise-level informal activities were often tolerated by state officials in order to achieve the goals of the central plan or because of corruption. At the start of the transition, many of informal activities such as production and exchange of private goods and services were legalized. However, informal activities still take place within medium- and large-size enterprises and many of the new private

micro-firms and entrepreneurs choose to operate in the informal sector. Along with traditional reasons for informalization (high tax and regulation burden, weak rule of law, inefficient law enforcement, small scale of production, etc.) there have been offered explanations that are specifically relevant to transition economies. They are related to general lack of confidence in state authorities and institutional uncertainty that emerged between the demolishing of old institutions and the construction of new ones. Additionally, many individuals were pushed into informal subsistence activities by rising unemployment and lack of work opportunities (Earle and Sakova, 2000).

Russia has experienced striking increase in informality over the years of transition to market economy. Buehn and Schneider (2012) provide evidence that the underground economy in Russia is large compared to other post-communist economies, exceeding 40 percent of official GDP in 1999-2007. Lehmann and Zaiceva (2013) employ a wide range of definitions of informality and find that dependent informal employment covers 7 to 20 percent of all employees and the share of informal self-employment lies between 45 and 73 percent of all self-employed. Being an important phenomenon, informality may have large effect on earnings inequality. Our primary objective in this paper is to explore this relationship[2].

The paper contributes to the existing literature in a number of ways. First, using the data from the Russia Longitudinal Monitoring Survey (RLMS-HSE) for 2000-2010 we provide an accurate measurement of the earnings gaps and earnings inequality for the informal sector as a whole and for different groups of informal workers in Russia. Second, we decompose earnings into hours of work and hourly earnings and study the relative importance of these two explanations. Third, we provide an econometric analysis of the heterogeneity observed in the informal sector by considering different types of informal employment and accounting for selection into informality. The understanding of heterogeneity that characterizes the informal sector has significant implications for any strategy and policy interventions aiming to reduce overall inequality and improve economic welfare. Fourth, unlike most studies which have focused on the formal-informal wage gap at the mean, we compare the shapes of the distributions estimate the earnings gaps along the distribution. The decomposition technique proposed by Firpo, Fortin, and Lemieux (2007) is applied to identify the contribution of endowments and returns to characteristics to differences in inequality between formal and informal workers.

The paper is organized as follows. Section 2 describes the sample, discusses the definition of key variables used in analysis and provides a decomposition of earnings inequality into inequality of hours and that of hourly earnings. Sector choice and selection issues are considered in Section 3. Section 4 presents the methodology of unconditional quantile regressions and Oaxaca-Blinder decompositions based on this type of regressions. This methodology is applied to study wage-setting mechanisms in the formal and

[2] In this paper, informality is defined to include workers employed without a work contract in the corporate sector, workers employed in private unincorporated units, and workers involved in remunerated irregular activities.

informal sector to quantify the relevance of various covariates in affecting the observed differences earnings inequality. The final section presents concluding comments.

## 2. Data description

The data used in this paper come from the 2000-2010 waves of the Russia Longitudinal Monitoring Survey (RLMS-HSE)[3,4]. The RLMS-HSE is a well-known panel survey of Russian households based on the national probability sample. Together with standard demographic variables at the individual and household level, the RLMS-HSE contains detailed information about labor market experience of individuals. This dataset has been previously used by a number of researchers to analyze informal employment relationships in the Russian labor market (Slonimczyk, 2012 and 2013; Lehmann et al., 2012 and 2013; Lehmann and Zaiceva, 2013; Zudina, 2013). Data from the RLMS-HSE Phase II sample are available for the period starting from 1994. However, the questions used to identify informal sector workers were introduced in 2000. So we do not employ the data for earlier years.

The sample used in this paper includes full-time and part-time workers who report to have a main job or are involved in other income-generating – mostly irregular – activities. The earnings variable is based on actual labor incomes received during the last 30 days from either the main job or irregular activities. Earnings are taken net of taxes and social security contributions. We eliminate observations with missing data in key variables, including age, nationality, education, earnings, and hours worked. Furthermore, we exclude individuals who report zero earnings and earnings which are at least two times higher than the 99.5$^{th}$ percentile of the distribution for respective year. These restrictions leave 53,965 observations in the baseline sample[5].

There is no consensually accepted definition of the informal sector in the literature. The definition in this paper is similar in many respects to the definition employed by Slonimczyk (2012)[6]. It covers:

1. workers employed without a work contract in the corporate sector,

[3] Russia Longitudinal Monitoring survey (RLMS-HSE) has been conducted since 1992 by the National Research University Higher School of Economics and ZAO "Demoscope" together with Carolina Population Center, University of North Carolina at Chapel Hill and the Institute of Sociology RAS. See detailed project description at: http://www.cpc.unc.edu/projects/rlms-hse.

[4] The round of 2001 is excluded for consistency reasons because the question on work contracts was not asked in this round.

[5] The original size of the subsample of employed individuals totaled to 69,089. The informal employment status could not be identified for 82 respondents. Additionally, 14,524 respondents were excluded from the sample because they had missing or zero responses to questions about wages and/or working hours (20.0 percent of formal sector workers and 25.1 percent of informal sector workers). Extra 518 observations were dropped because of other restrictions. The proportions of excluded observations are somewhat higher for the informal sector. However, the results presented in the paper are still representative of entire working population Russia.

[6] Slonymczyk (2012) offers detailed discussion of the definition of informal employment accounting for the limitations of the RLMS-HSE dataset. In particular, using the validation questions from a special informality supplement conducted in 2009, he shows that informal workers defined in this way have low level of compliance with labor and tax regulations and share many characteristics of informality documented in other countries. Lehmann and Zaiceva (2013) compare various definitions of informal employment with implication to the RLMS-HSE data. According to the definition which is very similar to the one used in this paper, the incidence of informal employment is close to average values across different measures. They also find that determinants of informal employment are robust across all definitions except for firm-size based measure.

2. workers employed outside the corporate sector in private unincorporated units that do not constituted separate legal entities,
3. people involved in irregular remunerated activities.

The first subgroup includes those who work in informal jobs in the registered firms and organizations. We classify the respondent into this group if she gives positive answer to the question: '*At this job do you work at an enterprise or organization? I mean any organization or enterprise where more than one person works, no matter if it is private or state-owned. For example, any establishment, factory, firm, collective farm, state farm, farming industry, store, army, government service, or other organization*' and negative answer to the question: '*Are you employed in this job officially, in other words, by labor book, labor agreement, or contract?*' Work contracts give right to receive state-mandated benefits such as paid annual leave, sick leave payments, severance pay and pension. Those who give negative answer to the former question fall into the second subgroup. Effectively, it consists of own-account professionals, unincorporated self-employed and their employees. Finally, the third subgroup covers those who claim not to have main job but report incomes from remunerated irregular activities. Involvement in such activities is identified by the question: '*In the last 30 days did you engage in some additional kind of work for which you were paid or will be paid? Maybe you sewed someone a dress, gave someone a ride in a car, assisted someone with apartment or car repairs, purchased and delivered food, looked after a sick person, sold purchased food or goods in a market or on the street, or did something else that you were paid for?*' For this subgroup the data do not allow making distinction between salaried employees and the self-employed. For this reason, we do not consider salaried employees and the self-employed separately in this paper.

**Table 1. Percentages of formal and informal employment**

| | 2000 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 |
|---|---|---|---|---|---|---|---|---|---|---|
| Formal | 80.3 | 80.8 | 80.2 | 78.7 | 80.0 | 79.9 | 80.1 | 82.8 | 79.2 | 81.4 |
| Informal employment | 19.7 | 19.2 | 19.8 | 21.3 | 20.0 | 20.1 | 18.9 | 17.2 | 20.8 | 18.6 |
| Employed without contract in the corporate sector | 3.1 | 3.8 | 4.7 | 5.1 | 5.5 | 6.0 | 4.9 | 5.3 | 6.2 | 5.6 |
| Self-employed and their employees | 6.8 | 6.1 | 5.8 | 7.4 | 6.6 | 6.4 | 7.9 | 6.7 | 7.7 | 8.2 |
| Employed in irregular activities | 9.8 | 9.3 | 9.3 | 8.8 | 8.0 | 7.8 | 6.2 | 5.1 | 6.9 | 4.8 |
| Number of observations | 3,521 | 4,426 | 4,600 | 4,879 | 4,675 | 5,774 | 5,693 | 5,844 | 5,770 | 8,788 |

Table 1 presents the employment shares of each subgroup in the sample. Changes in economic conditions hardly had any effect on the size of the informal sector as a whole. Despite strong economic growth in 2000-2007 the share of informal workers stayed roughly stable amounting to about 20 percent of total employment. The response to the global economic crisis of 2008-2009 was rapid but shallow and

transient. Informal employment shrank somewhat in 2008 and bounded back to pre-crisis levels in 2009. Fundamental changes occurred *within* the informal sector. The share of irregular workers has been steadily declining while both types of 'regular' informality have been expanding over the period. This trend may indicate the diminishing role of informality as a survival strategy since irregular activities were most likely used by the households to cope with the shocks of the early transition period. These changes in the structure of informal employment are prominent because in the beginning of the period irregular activities were the most common form of informal work in Russia.

On the other hand, the cumulative share of 'regular' informal employment increased from 9.8 percent in 2000 to 13.8 percent in 2010. This rise in informal employment has come about despite economic growth and the efforts to improve business climate. Slonimczyk (2012) has found that the introduction of the flat rate of personal income tax in 2001 had adverse effect on the fraction of informal employees decelerating the rate of growth of informal employment. However, this tax reform failed to break an upward trend in informality. Ivanova et al. (2005) and Gorodnichenko et al. (2009) report positive effects of the flat tax reform on compliance. More general conclusion from these studies is that informality in Russia is sensitive to changes in the tax system. This conclusion equally applies to social security contributions that are not balanced with the benefits. High and regressive social security contributions discourage firms from hiring low-skilled labor on the formal basis. Additionally, the incentive for workers to join the formal economy is lowered by the ease in gaining access to social benefits. In Russia, health insurance is universally provided to unemployed and non-employed. The minimum contribution period to qualify for labor pension is as low as 5 years. Muravyev and Oshchepkov (2013) claim that minimum wages hikes in the 2000s have been associated with rising costs of hiring low-skilled workers and increased informality. Thus, excessive regulation is likely to blame for the growth of the `regular' informal sector in Russia.

Table 2 summarizes demographic characteristics and key indicators of human capital for formal and informal workers. The numbers in the table generally support the view of the informal sector as a disadvantaged segment of the labor market. It employs higher proportions of rural inhabitants and ethnic minorities. Informal workers tend to be younger and less educated. The informal sector lags behind in human accumulation reflected in lower increase in the share of university graduates over the period. Finally, the proportion of men is higher among informal workers. The latter seems to be the only observable characteristic that may drive up earnings in the informal sector.

The table also reports mean monthly earnings and working hours for formal and informal workers. Monthly earnings are used instead of hourly wages given the irregular working time arrangements in the informal sector and often constraint working hours in both sectors. For this reason, we believe that monthly wages reflect earning opportunities in the informal sector better than hourly wages (we come

back to the differences in working hours and hourly wages in the next section). As expected, there is a substantial earnings gap of about 20 percent in favor of formal workers. At the same time, at least part of this gap can be attributed to lower working hours.

**Table 2. Descriptive statistics**

| | 2000 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 |
|---|---|---|---|---|---|---|---|---|---|---|
| Formal employment | | | | | | | | | | |
| Average age, years | 39.8 | 39.7 | 39.8 | 39.8 | 39.9 | 40.2 | 40.1 | 40.0 | 40.6 | 40.4 |
| Females, % | 53.9 | 53.5 | 54.7 | 54.6 | 53.7 | 55.2 | 55.2 | 54.7 | 55.3 | 54.9 |
| Ethnicity - % of Russians | 86.9 | 86.4 | 86.5 | 86.0 | 87.0 | 87.0 | 87.0 | 85.2 | 87.2 | 88.3 |
| Rural, % | 19.1 | 18.0 | 19.0 | 18.3 | 19.6 | 20.6 | 20.8 | 20.6 | 20.7 | 22.3 |
| University degree, % | 24.6 | 26.0 | 25.9 | 26.6 | 27.1 | 27.6 | 28.0 | 28.9 | 30.2 | 32.0 |
| Average net monthly nominal earnings, rubles | 1851 | 3677 | 4487 | 5417 | 6767 | 8191 | 9948 | 13136 | 13142 | 14650 |
| Average net monthly nominal earnings, current international USD*) | 244.9 | 365.9 | 407.7 | 439.7 | 505.3 | 602.2 | 693.0 | 871.6 | 826.4 | 878.3 |
| Average monthly working hours | 167.4 | 167.5 | 167.6 | 167.8 | 168.6 | 173.0 | 174.4 | 172.6 | 171.4 | 171.8 |
| Informal employment | | | | | | | | | | |
| Average age, years | 36.0 | 35.7 | 35.8 | 36.0 | 36.2 | 36.3 | 36.9 | 37.3 | 37.9 | 37.3 |
| Females, % | 45.3 | 45.5 | 45.0 | 44.6 | 44.6 | 44.4 | 44.3 | 45.7 | 46.5 | 43.1 |
| Ethnicity - % of Russians | 83.9 | 79.5 | 80.3 | 81.3 | 78.9 | 80.2 | 82.6 | 79.0 | 82.5 | 82.0 |
| Rural, % | 23.0 | 26.3 | 31.4 | 25.5 | 29.0 | 27.0 | 29.5 | 27.0 | 28.1 | 28.3 |
| University degree, % | 11.7 | 12.2 | 10.8 | 11.7 | 10.6 | 10.6 | 11.3 | 12.4 | 12.6 | 14.3 |
| Average net monthly nominal earnings, rubles | 1654 | 3226 | 3684 | 4726 | 5396 | 6683 | 8319 | 10821 | 10348 | 12579 |
| Average net monthly nominal earnings, current international USD*) | 218.8 | 321.0 | 334.8 | 383.6 | 403.0 | 491.4 | 579.5 | 718.0 | 650.7 | 754.1 |
| Average monthly working hours | 140.7 | 137.5 | 137.1 | 141.6 | 145.8 | 152.9 | 158.9 | 155.1 | 153.2 | 160.1 |

Note: *) PPP conversion rates are taken from http://data.worldbank.org/indicator/PA.NUS.PRVT.PP

For closer look at earnings and working hours, we break them down by types of informal employment. Figure 1 plots the average earnings and working hours in each group divided by the respective averages for formal workers. The informal sector shows heterogeneous structure. We observe that irregular workers are very different from both formal workers and two other groups of informal workers. They are often employed part-time and monthly earnings are only half of those of formal workers. On the contrary, two other groups of informal workers have many common features with formal workers. The group of self-employed and their employees enjoyed up to 20 percent premium in the beginning of the period but this advantage fade out in the mid-2000s. The earnings gap for employed without contract in the corporate sector fluctuated around zero during the whole period. However, both groups of informal workers tend to have longer working hours.

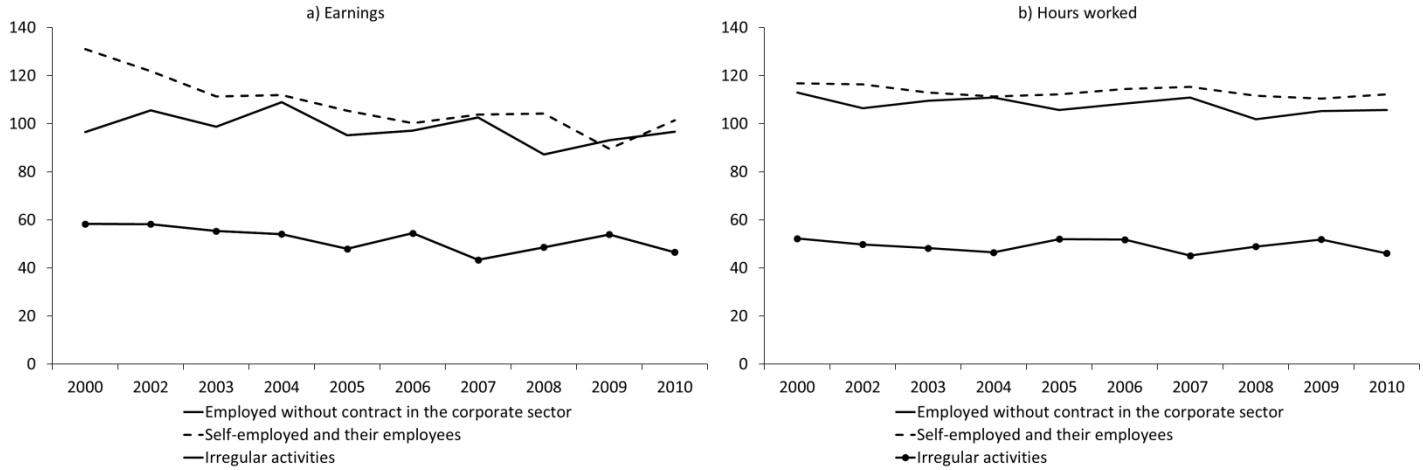


**Figure 1. Average earnings and working hours by types of informality, in percent (100 = formal employment)**

### 3. Descriptive analysis of inequality

To address the question of how much change there has been in the formal and informal earnings distributions over the period, we plot kernel density functions and percentiles graphs for 2000 and 2010 (Figure 2) and calculate inequality measures for all years of the period (Tables 3 and 4). Standard errors in Table 3 were calculated by bootstrap method.

The general picture that emerges from inequality measures is that during the whole period variation of earnings was considerably higher in the informal sector. Comparison of kernel-densities suggests that in both sectors earnings are more dispersed among low-income earners than in the upper part of the distribution. Upper tails of the two distributions almost coincide. However, the informal sector distribution is broader with more mass at the lower part reflecting larger proportion of low paid workers in this sector. The major pay differences between the sectors are at the bottom of the distribution that is among the low-wage earners. One possible reason for this result is the existence of the minimum wage that is enforced in the formal sector, but not in the informal sector. Other reasons include the prevalence of part-time work and higher proportion of poorly educated workers in the informal sector.

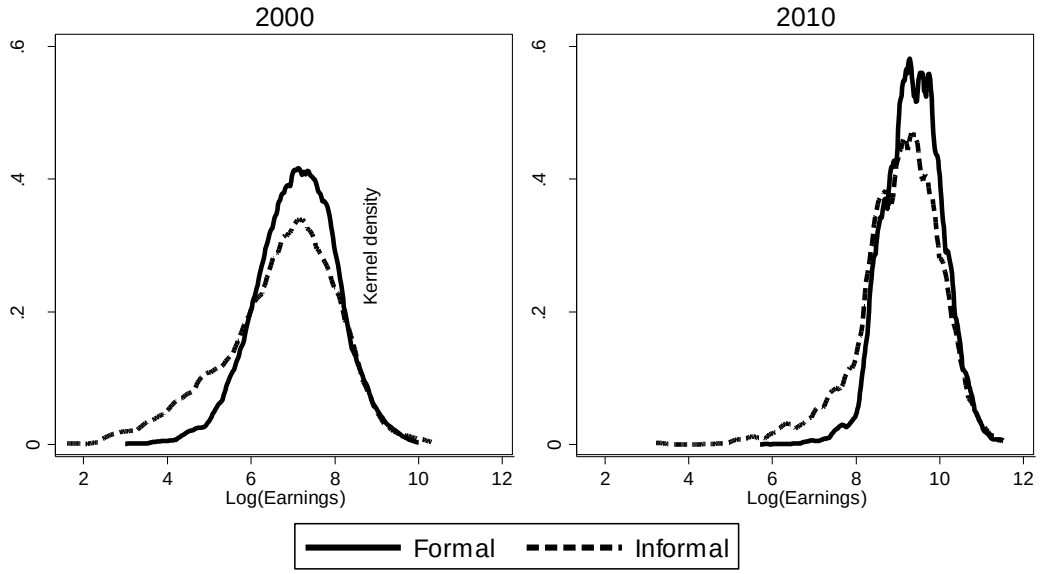


**Figure 2. Kernel density estimates of log earnings**

Earnings inequality has declined substantially between 2000 and 2010 in both sectors. Changes in the inequality measures indicate that the pace of decline has been remarkably similar between the sectors. The Gini coefficient suggests a contraction in earnings inequality in a range between 20 and 25 percent (Table 3). The variance of log-earnings almost halved in both sectors (Table 5). Interestingly, the response of inequality to the global economic crisis of 2008-2009 was very moderate. The Gini coefficient increased only marginally and insignificantly in both sectors. The sensitivity of wages to economic shocks becomes weaker as market economy matures. These developments are consistent with previous studies that report increased self-insurance of Russian households against permanent income shocks and better smoothing of transitory shocks (Gorodnichenko et al., 2010).

**Table 3. Gini coefficient**

| | 2000 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 |
|---|---|---|---|---|---|---|---|---|---|---|
| Formal employment | 0.481 | 0.434 | 0.435 | 0.417 | 0.409 | 0.396 | 0.375 | 0.381 | 0.370 | 0.365 |
| | (0.006) | (0.007) | (0.006) | (0.005) | (0.006) | (0.005) | (0.004) | (0.004) | (0.004) | (0.003) |
| Informal employment | 0.568 | 0.539 | 0.556 | 0.508 | 0.492 | 0.497 | 0.465 | 0.478 | 0.461 | 0.439 |
| | (0.017) | (0.015) | (0.017) | (0.009) | (0.012) | (0.011) | (0.008) | (0.011) | (0.009) | (0.008) |
| Employed without contract in the corporate sector | 0.446 | 0.451 | 0.450 | 0.434 | 0.431 | 0.403 | 0.390 | 0.391 | 0.409 | 0.395 |
| | (0.029) | (0.030) | (0.020) | (0.018) | (0.021) | (0.013) | (0.013) | (0.012) | (0.015) | (0.015) |
| Self-employed and their employees | 0.477 | 0.449 | 0.489 | 0.434 | 0.443 | 0.441 | 0.413 | 0.443 | 0.402 | 0.382 |
| | (0.030) | (0.028) | (0.028) | (0.012) | (0.020) | (0.020) | (0.011) | (0.017) | (0.013) | (0.009) |
| Irregular informal workers | 0.637 | 0.606 | 0.635 | 0.580 | 0.524 | 0.590 | 0.516 | 0.565 | 0.547 | 0.530 |
| | (0.023) | (0.020) | (0.030) | (0.017) | (0.017) | (0.018) | (0.014) | (0.022) | (0.021) | (0.015) |

Note: Bootstrapped clustered standard errors in parentheses (500 replications).

Closer inspection of inequality measures reveals that higher inequality within the informal sector is mainly driven by those employed in irregular activities. Two other informal subgroups demonstrate levels of earnings dispersion that are surprisingly similar to those in the formal sector. Gini coefficients are on average larger by 5-8% but these differences are not statistically significant. Higher inequality is observed only at the tails of the distribution reflected in larger values of log-variance for these groups.

**Table 4. Raw earnings gaps and inequality differentials between the formal and informal sector**

| | 2000 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Employed without contract in the formal sector** | | | | | | | | | | |
| Mean | 0.002 | 0.038 | -0.054 | 0.046 | -0.096* | -0.046 | 0.000 | -0.162*** | -0.131*** | -0.078** |
| q10 | -0.090 | -0.046 | -0.230 | -0.128 | -0.168*** | -0.119* | -0.026 | -0.281*** | -0.297*** | -0.210*** |
| Median | 0.075 | 0.019 | 0.017 | 0.055 | -0.051 | -0.036 | -0.015 | -0.078** | -0.045 | -0.090* |
| q90 | 0.069 | 0.109 | 0.084 | 0.158 | 0.071 | 0.037 | 0.003 | -0.106 | -0.007 | -0.001 |
| Variance of logs | 0.009 | -0.022 | 0.127 | 0.113 | 0.117 | 0.049 | 0.059 | 0.117** | 0.124* | 0.100** |
| **Self-employed and their employees** | | | | | | | | | | |
| Mean | 0.327*** | 0.170*** | 0.063 | 0.094** | -0.001 | -0.041 | -0.026 | -0.046 | -0.150*** | -0.023 |
| q10 | 0.484*** | 0.060 | 0.141 | 0.073 | -0.134* | -0.081* | -0.108** | -0.200** | -0.093*** | -0.076** |
| Median | 0.283*** | 0.161*** | 0.031 | 0.059 | -0.030 | -0.051 | -0.026 | -0.028 | -0.162*** | -0.167*** |
| q90 | 0.263*** | 0.245*** | 0.130** | 0.298*** | 0.151*** | 0.098 | 0.049 | 0.139** | -0.095 | 0.099*** |
| Variance of logs | -0.155* | 0.033 | 0.058 | 0.027 | 0.079 | 0.060 | 0.138*** | 0.165*** | 0.072 | 0.073** |
| **Irregular informal activities** | | | | | | | | | | |
| Mean | -1.005*** | -1.001*** | -0.990*** | -1.050*** | -1.040*** | -1.036*** | -1.155*** | -1.120*** | -0.977*** | -1.124*** |
| q10 | -1.826*** | -1.875*** | -1.664*** | -1.872*** | -1.766*** | -1.804*** | -1.795*** | -1.799*** | -1.609*** | -1.921*** |
| Median | -0.884*** | -0.952*** | -0.919*** | -0.989*** | -0.929*** | -0.929*** | -0.988*** | -1.167*** | -0.834*** | -0.931*** |
| q90 | -0.291*** | -0.306*** | -0.421*** | -0.364*** | -0.558*** | -0.364*** | -0.642*** | -0.513*** | -0.483*** | -0.532*** |
| Variance of logs | 1.108*** | 1.150*** | 1.013*** | 1.167*** | 0.899*** | 1.066*** | 0.857*** | 0.963*** | 0.831*** | 0.980*** |

Note: Values in the table are differences in means and distributional statistics between the informal and formal sector. ***, **, and * denote statistical significance at 1 percent, 5 percent, and 10 percent level, respectively.

Table 4 presents the raw differentials in various distributional statistics for various types of informal employment. Positive values in the table refer to the cases when means, relevant percentiles or variances are higher in the informal sector than in the formal sector. There are few significant values for those employed without contract in the formal sector. All of them are found for the later years of the period and may represent short-term consequences of the global financial crisis. The earnings distribution of those employed without contract in the formal sector is very similar to that in the rest of the formal sector. Jobs in this part of the informal sector yield comparable incomes, albeit these incomes may be more responsive to negative economic shocks. It looks like that formal and informal workers work side by side at the registered enterprises blurring the formal-informal dichotomy. Those who work without contract probably lack job protection and have restricted access to some social benefits. However, provision of major social benefits like health care, pensions and even unemployment allowances is almost universal in Russia. Workers may attribute less value to formal jobs, especially if these jobs imply less flexibility and more

taxation. Moreover, additional social protection benefits offered by formal job contracts may be seen of low quality and therefore not worth the contributions. On the other hand, tax underreporting is widespread and formal workers often receive part of their remuneration in the envelopes. The formal sector by our definition may be also involved into informal labor practices.

The relative position of the self-employed and their employees changed over the period. They had higher earnings and lower within-group inequality than formal workers in the beginning of the period, but have gradually lost these advantages and finally the premium turned into the penalty. These developments are most likely the result from rapid economic growth which Russia experienced in 2000-2007. Growing revenues in the private formal sector and improving budget conditions fueled earnings increases in the formal sector. Many of Russian formal sector workers are employed in state-funded establishments and have salaries below those in the private sector (Gimpelson and Lukiyanova, 2009). Public sector workers shared the benefits of economic growth as additional budget revenues were poured into their salary bills. The reform of income tax of 2001 and the social security reform of 2005 that lower the rates of mandatory contributions from 35.6 to 26 percent provided additional stimulus to formalization. The costs of being formal dropped significantly in 2000-2010. Formal sector employers increased earnings of their employees and got chance to attract more productive informal workers into the formal sector jobs.

The group of self-employed and their employees can be suspected to be segmented into the upper and lower tiers. There are only subtle signs of such segmentation in raw differentials. In the early 2000s there were no signs of duality within this segment, rather gains from working in this segment were shared between the self-employed and their salaried workers. In 2000, the raw differentials were positive and significant both in the bottom and in the top distribution. The segment experienced a horizontal shift in raw differentials over the period suggesting that relative losses were shared by majority of workers. The increase of inequality differentials indicates that relative losses were lower for the top percentiles for whom the differential remained positive in 2010.

Raw differentials are significantly negative in all years for irregular informal workers. There no signs that they narrowed between 2000 and 2010. In spite of high within-group inequality the gap is persistently negative even for top deciles of the distribution. It is very likely that the greater part of this segment follows common beliefs about the informal sector as comprised of subsistence-level production activities, motivated by the need for survival and characterized by low levels of income, productivity, skills, technology and capital, and weak linkages with the rest of the economy.

Figure 3 gives more detailed picture of raw differentials depicting them for all quantiles of the earnings distribution. It confirms conclusions that higher inequality in informal sector is largely attributed to low wages of irregular workers and to high within-group inequality among such workers. The gap is the largest for low pay workers and bridges only at the top of distribution. Informality has unequalizing effect

contributing to the polarization of the labor market widening the top and the bottom of distribution. However, this effect was been diminishing over the period as the group of self-employed and their employees was losing its premium and converging with the formal sector. However, in general, contrary to the widely shared believes, for many workers informal employment has been financially more attractive or at least yielded same incomes as work in the formal sector, especially for middle- and high-skilled workers.

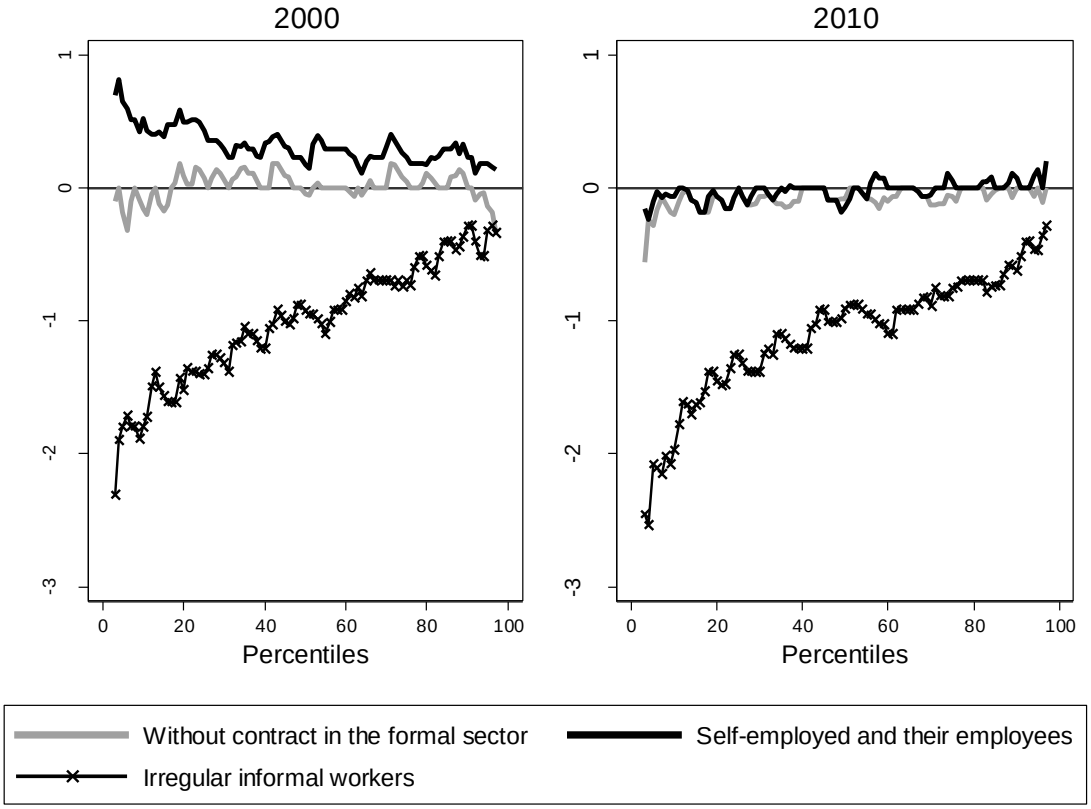


**Figure 3. Raw differentials between informal and formal workers by percentiles of the earnings distribution**

The earnings distribution results from the joint distributions of hourly wages and workings hours. We decompose earnings into hours of work and hourly wages and investigate the importance of differences in the distributions of hours in explaining the formal-informal gap in earnings. The literature on inequality provides limited and often indirect evidence on the effects of hours of work on changes in inequality over time. Doiron and Barrett (1996) proposed the methodology that decomposes total earnings inequality into the contributions of hours, wages, and their Interaction. The shortcoming of this methodology is that implicitly assumes the independence between informality and wages/hours. Since being formal is associated with a worker's choice, the results of this decomposition do not necessarily have a causal interpretation. Rather they provide a descriptive comparison of the earnings distribution in the formal and informal sectors. Later on in this paper we discuss the endogeneity of the sector choice in more detail.

The decomposition is straightforward for the variance of log-earnings[7]:

$$Var\left(\ln mw\right) = Var\left(\ln h\right) + Var\left(\ln hw\right) + 2Cov(\ln h, \ln hw)\,, \qquad (1)$$

where *h* denotes hours, *hw* denotes hourly wages, and *mw* is total earnings. The covariance term describes the interaction between hours and hourly wages. This interaction component reflects the relationship between hours and hourly rates at the individual level. The interaction effect can be inequality-reducing if workers with higher hourly wages tend to work shorter hours, or an inequality-enhancing if the relationship goes in the opposite direction.

The decomposition is compared for the formal sector and various forms of informal employment in Table 5. The results are presented in absolute values and in percentages. Over the period, there has been substantial inequality compression both in monthly earnings and in hourly rates for the formal sector and for all types of informal employment. In the formal sector, the changes in the earnings distribution are generated almost exclusively by the declining variation in hourly earnings, while the distribution of working hours hardly changed over the period. Changes in the covariance term are also of minor importance for the formal sector.

**Table 5. Decomposition of inequality within subgroups of informal workers: variance of logs**

| Year | Formal employment | | | | Employed without contract in the corporate sector | | | | Self-employed and their employees | | | | Irregular informal workers | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MW | HW | H | Cov (HW, H) | MW | HW | H | Cov (HW, H) | MW | HW | H | Cov (HW, H) | MW | HW | H | Cov (HW, H) |
| 2000 | 0.915 | 0.944 | 0.183 | -0.212 | 0.931 | 1.078 | 0.221 | -0.368 | 0.793 | 1.009 | 0.202 | -0.418 | 2.022 | 1.462 | 2.080 | -1.520 |
| 2005 | 0.645 | 0.679 | 0.117 | -0.152 | 0.762 | 0.698 | 0.171 | -0.107 | 0.724 | 0.729 | 0.193 | -0.198 | 1.546 | 1.114 | 1.861 | -1.429 |
| 2010 | 0.473 | 0.509 | 0.127 | -0.163 | 0.579 | 0.564 | 0.267 | -0.252 | 0.533 | 0.576 | 0.207 | -0.249 | 1.451 | 1.115 | 1.871 | -1.535 |
| Change over 2000-2010 | -0.441 | -0.435 | -0.056 | 0.049 | -0.352 | -0.513 | 0.045 | 0.116 | -0.260 | -0.433 | 0.005 | 0.169 | -0.571 | -0.347 | -0.209 | -0.015 |
| Percentage contribution to the reduction of inequality: | | | | | | | | | | | | | | | | |
| 2000 | 100 | 103 | 20 | -23 | 100 | 116 | 24 | -40 | 100 | 127 | 26 | -53 | 100 | 72 | 103 | -75 |
| 2005 | 100 | 105 | 18 | -24 | 100 | 92 | 22 | -14 | 100 | 101 | 27 | -27 | 100 | 72 | 120 | -92 |
| 2010 | 100 | 108 | 27 | -34 | 100 | 97 | 46 | -43 | 100 | 108 | 39 | -47 | 100 | 77 | 129 | -106 |
| Change over 2000-2010 | 100 | 98 | 13 | -11 | 100 | 146 | -13 | -33 | 100 | 167 | -2 | -65 | 100 | 61 | 37 | 3 |

The results for the informal sector show that it presents a high degree of heterogeneity in the sources of inequality. The structure of inequality in the 'regular' forms of informal employment resembles that in the formal sector (with few important differences), while irregular informal activities clearly follow a distinctive pattern. On average, the inequality of both monthly and hourly earnings among those employed without a contract in the corporate sector and among the self-employed and their employees is 10-20 percent higher than in the formal sector. On the contrary, for hours the inequality is on average higher by

[7] Decomposition for Gini is slightly more complicated but results are qualitatively the same. They are not provided in the paper, but available upon request and can be found in the working paper version.

40-60 percent. This higher dispersion of hours in these two informal subgroups is largely offset by the larger negative interaction component. The compression of wintin-group inequalities for regular forms of informal employment between 2000 and 2010 can be attributed to a decline in the variance of hourly earnings. Changes in two other components have a positive sign for both subgroups, implying that their impact was inequality-enhancing. Changes in hours inequality are quite small and definitely have not been an important driving force under the evolution of earnings inequality in this part of the informal sector.

Large and persistent interaction term differentiates regular informal employment from the formal sector. It has a negative sign in the decomposition (see upper panel of Table 5). Therefore, this component has an equalizing effect, implying that workers with lower hourly rates generally work longer hours. However, the equalizing effect is diminishing over time. This may reflect a shift from the demand side. More workers from the informal sector who initially involuntarily worked short hours found jobs with longer hours. For whatever reasons the interaction term is coming closer to the value observed in the formal sector. Together with declining variation of hourly rates, this finding suggests further 'regularization' of 'regular' types of informal employment.

The subgroup of employed in irregular activities evidently stands apart from the formal workers and two other informal subgroups. In this subgroup both higher inequality of hours and higher inequality of hourly rates contribute to higher total variance of monthly earnings. The negative interaction component is large in absolute terms, but it is not sufficiently large to counterbalance higher inequality in hours and hourly earnings. It is noteworthy that reduction in the hours inequality, though smaller compared to the decline in hourly earnings inequality, played important role in improvement of overall inequality between 2000 and 2010. This finding means that analysis of differences in inequality should account for specific pattern of hours determination in this subgroup.

What we call 'irregular' employment is likely to offer more flexible working arrangements than other types of formal and informal employment. Other types of employment offer limited number of jobs with part-time working arrangements and flexible hours while more workers gain higher utility from being able to choose their working schedule. They may opt for irregular work that provides them target income while working distantly or part-time. Many of those who work short hours are mothers with small children and pensioners. For some workers irregular work is a deliberately chosen strategy of job search. They treat is a temporary job that leaves time for active job search. The RLMS data gives some empirical evidence of this hypothesis. Those in irregular employment 5-6 times more frequently than formal workers and 3 times more frequently than other informal workers report applying somewhere or asking someone for a job in last 30 days. Of course, some workers end up in irregular activities if they only able to find a seasonal job or a job with daily payments. These findings on the differences in hours distributions and their interactions with hourly earnings need further study.

Table 5 is informative about the changes in within-group inequality but there remains to explain the role of between-group differentials and changes in the shares of different subgroups of workers. To answer this question, we decompose the 2000-2010 change in the variance of the log of monthly earnings into within- and between-group components and compositional effects. The variance of logs in year *t* ($Var(Y_t)$) can be broken down into population subgroups using the following formula[8]:

$$Var(Y_t) = \sum_{j=1}^{K} f_{jt} Var(Y_{jt}) + \sum_{j=1}^{K} f_{jt} \left(\overline{Y}_{jt} - \overline{Y}_t\right)^2 \quad (2)$$

where $Y_t = \ln(mw_t)$, $f_{jt}$ is the fraction of workers in subgroup *j* in year *t* (j=1, 2, .., K), $Var(Y_{jt})$ is the variance of earnings in group *j* in year *t*, $\overline{Y}_{jt}$ is the average earnings in group *j* in year *t*, $\overline{Y}_t$ – is the average earnings of all workers in year *t*.

The change in variance over time can then be decomposed into shifts in the sectoral composition (changes in fractions of population subgroups), and changes in within- and between-group inequality:

$$\begin{aligned} \Delta Var(Y_t) = &\sum_{j=1}^{K} f_{j0} \left(Var(Y_{j1}) - Var(Y_{j0})\right) + \sum_{j=1}^{K} f_{j0} \left((\overline{Y}_{j1} - \overline{Y}_1)^2 - (\overline{Y}_{j0} - \overline{Y}_0)^2\right) \\ &+ \sum_{j=1}^{K} \left(f_{j1} - f_{j0}\right) Var(Y_{j1}) + \sum_{j=1}^{K} \left(f_{j1} - f_{j0}\right) (\overline{Y}_{j1} - \overline{Y}_1)^2 \end{aligned} \quad (3)$$

The first two terms in equation (3) reflects changes in within- and between-group variance, respectively. The third and the fourth term account for the compositional effects. The compositional effects reduce the variance if employment shifts either towards the groups with low within-group variance or towards the groups with average wage closer to the mean.

The results of decomposition are presented in Table 6. Changes in inequality in 2000-2010 are mainly explained by narrowing inequality within the groups (83 percent of total reduction). Changes in the composition of employment stand for another 20 percent of total reduction. However, the gaps in average wages across the groups remained roughly unchanged and had insignificant impact on total inequality.

Formal employment has the largest contribution to the reduction of the total variance mostly due to the decline in within-group earnings dispersion. It is not surprising as this group composes 80 percent of total employment. In fact the contribution of formal employment (66 percent of total reduction) is relatively small compared to the share of this group in total employment. Really striking impact was made by the subgroup of irregular workers. Their contribution to the reduction of total inequality is equal to 30%

[8] This methodology provides rough estimates of changes in variance for the RLMS-HSE data as it assumes that 2000 and 2010 samples are independent cross-sections. Given that the RLMS-HSE follows individuals over time that may not be the case even though the sample was substantially replenished and expanded between 2000 and 2010. As a result, about 19 percent of the pooled sample are the same in both years. The results hold if we limit the sample to the individuals who participated only in one of the years. Moreover, the survey of repeated cross-sections design with questions about informality does not exist in Russia. At this stage we ignore the issues of selection as well. So the presented analysis of compositional shifts and changes in within- and between-group variances is purely descriptive.

which much larger than their employment share. For this subgroup, the shrinking of within-group dispersion was coupled with the decrease in the relative share of this subgroup. The latter change is at least twice more important than changes in inequality within the group. Irregular informal workers moved towards the groups with lower earnings dispersion and with average earnings which are closer to the overall mean.

**Table 6. Decomposition of changes in inequality (2000-2010): variance of logs**

| | All workers | Formal employment | Informal employment | | |
|---|---|---|---|---|---|
| | | | Employed without contract in the corporate sector | Self-employed and their employees | Irregular informal workers |
| Total change (2000-2010) | 0.527 | 0.350 | -0.004 | 0.022 | 0.159 |
| 1. Within-group | 0.438 | 0.354 | 0.011 | 0.017 | 0.056 |
| 2. Between-group | -0.016 | 0.002 | 0.000 | 0.012 | -0.029 |
| 3. Compositional effects | 0.105 | -0.006 | -0.015 | -0.007 | 0.132 |
| In percent: | | | | | |
| Total change | 100 | 66 | -1 | 4 | 30 |
| 1. Within-group | 83 | 67 | 2 | 3 | 11 |
| 2. Between-group | -3 | 0 | 0 | 2 | -6 |
| 3. Compositional effects | 20 | -1 | -3 | -1 | 25 |
| Employment share – 2000 | | 80.3 | 3.1 | 6.8 | 9.8 |
| Employment share – 2010 | | 81.4 | 5.6 | 8.2 | 4.8 |

**3. Selection bias corrections**

There are many possible costs and benefits of being informal. Individuals can consider the informal sector more attractive than the formal sector even if earnings are lower in the former sector. Voluntary employment in the informal sector can be due to desirable non-wage features of the informal sector where individuals maximize their utility rather than their earnings. Moreover, workers may have individual comparative advantages in the informal sector and would earn more in the formal sector (Maloney, 2004). On the other hand, individuals could want to work in the formal sector, but they did not find a job, this happens when markets are segmented and there is a queue to enter in the formal sector (Harris and Todaro, 1970). The informal sector itself can highly heterogeneous and contain a competitive part into which individuals enter voluntarily along with a lower tier where workers are rationed out of the formal market (Fields, 2005). Another reason for selection bias may be related to sorting if, for instance, low ability individuals are also those who are more likely to work in the informal sector. Empirical studies conclude that important to accounts for sample selection caused by individuals' self-selection into informality (e.g., Günther and Launov, 2012).

For Russia, Lehmann and Zaiceva (2013) show that risk-averse individuals are less likely to select themselves into informal employment. Several papers relate selection into informality with job mobility.

Lehmann, Razzolini and Zaiceva (2012) and Lehmann et al. (2013) find that those who separate from jobs, whether voluntarily or involuntarily, have a higher probability of finding a subsequent job in the informal sector or where part of earnings is off-book. They stress that this probability is particularly high for workers who were displaced or separated from informal jobs and for those with low human capital. Slonimczyk (2013) employs a dynamic multinomial logit framework to model complex mobility decisions and finds little evidence of entry barriers to the formal sector. He concludes that informal employees are not more likely than the unemployed to get a formal position. My reading of this literature is that self-selection into informality clearly dominates over selection by employers and segmentation issues.

There are two types of selection problems in the data. First, there is a choice to work or to stay out of the labor market. Second, there is selection between working in the formal market and various forms of informal employment. We combine the two decision into a single multinomial model and estimate it applying the selection bias correction method proposed by Lee (1983) and later modified by Dubin and McFadden (1984) and Bourguignon, Fournier and Gurgand (2007). This approach is a generalization of Heckman two-step procedure that is widely used to correct for sample selection. It was recently applied to study labor market segmentation in Indonesia (Comola and de Mello, 2009), Sri Lanka (Gunatilaka and Vodopivec, 2010) and West Africa (Barlet, 2013).

In our set-up individuals face five possible choices: not participating in the labor market, working in the formal sector (this sector is assumed homogeneous), and working in the informal sector which consists of three segments as defined above (employed without contract in the corporate sector; self-employed and their employees; employed in irregular activities). These choices are the outcomes in the multinomial selection model:

$$y_s = x_s\beta_s + e_s \text{ and } y_s^* = z_s\gamma_s + u_s, \qquad (4)$$
$$\text{where } s = 1, \dots, 5 \;;\; E(e_s|x,z) = 0 \text{ and } E(u_s|x,z) = \sigma_s^2.$$

Each individual selects the option that yields the highest predicted utility $y_s$. The outcome variable $y_s^*$ is observed if and only if $y_s^* > max_{j \neq s} y_j^*$, so that alternative $s$ is selected. The vector $z_s$ represents the set of explanatory variables for the sector choice and the vector $x_s$ contains all determinants of earnings.

Each segment of the labor market has its own unique wage function. The problem is to estimate the coefficients $\beta_s$ taking into account that the error term in the earnings equation is not independent of all error terms in the selection equation. These correlations, if existent, lead to the inconsistency of the OLS estimates. The idea is to calculate the bias correction terms that are equivalent to the inverse Mills ratios in the Heckman correction procedure and then introduce them into the earnings equations. Bourguignon, Fournier and Gurgand (2007) compare the methods available in the literature to deal with selectivity bias in the situations with multiple choices. The methods differ in functional forms of the correction terms. In this paper, we follow the approach of Dubin and McFadden (1984) in its flexible version proposed in

Bourguignon, Fournier and Gurgand (2007)[9]. The most serious assumption within the multinomial logit frameworks is the assumption of the independent of irrelevant alternatives (IIA). This assumption may be violated especially given that participation and sector selection decisions may not be simultaneous. However, Bourguignon, Fournier, and Gurgand (2007) use Monte Carlo simulations to show that the results of the second step are robust even when the IIA assumption is violated.

The sector choice is assumed to depend on demographic characteristics, human capital, family composition, and location. The set of explanatory variables includes age, age squared, gender, marital status, education (4 categories), household size, region (7 categories), an urban dummy, dummies for the presence of children aged 0-6 and 7-16 in the household, and a dummy for Russian nationality. Children variables and marital status are additionally interacted with gender.

For the second step, we estimate an earnings equations for each type of employment (except non-participation). The dependent variable is the logarithm of monthly earnings. The set of explanatory variables in the earnings equations contains the same variables as in the selection equation except for household characteristics, marital status and all interactions. The identification implies that these excluded variables determine sector choice but not earnings. Additionally, the earnings equation includes the log of working time and selection correction terms calculated from the first-step equation. Table A1 in Appendix shows the results of the sector choice estimations. The reference category is the formal sector. For all years, the LR-tests of whether dependent groups can be combined rejects the hypothesis that any pair of groups can be collapsed into a single group at the 5 percent significance level.

Age has a significant impact on sector choice. Young people are more likely to be inactive because many of them are in education. After entering the labor market, young people have a greater probability to work without contract in the formal sector or to be engaged in irregular informal activities than to work in the formal sector. This finding suggests that these two types of informal employment are a stepping stone before entry into a more desirable sector (the formal sector or any form related to self-employment).

Having any kind of post-school education raises the probability of participation at the labor market as it increases the opportunity costs of non-employment. Tertiary education (college and university) also generally enhances the chances of working in the formal sector. For the segment of self-employed and their employees this regularity holds strongly only for the university graduates, but not for college graduates. Workers with primary and low secondary education are more likely to end up in inactivity or irregular employment.

Rural inhabitants are more likely to stay out of the labor market. Subsistence agriculture continues to be an alternative to paid employment in rural areas. However, the effect of location is ambiguous for the sector choice. Living in the city increases the probability to be employed without contract in formal

[9] The model was estimated in Stata with *selmlog* command developed by M.Gurgand and M.Fournier.

sector but reduces the probability of being in other forms of informal employment. Living in rural areas reduces the probability of being engaged in irregular activities but has no clear effect on other types of informality.

Single women are as likely to be employed as males. Getting married and having children aged under six significantly increases the likelihood of inactivity. When children become elder women tend to return to the labor market. Men are more likely to work informally, but for the whole period that is the case only for irregular informal activities. This impact is especially robust for single males. For other forms of informal employment the effect is significant for the second half of the period but not for the early 2000s. Nationality seems to be an important determinant of informality status. Belonging to national minorities positively affects the risks to stay out of labor market or to be employed in irregular informal activities. Household size mostly affects the probability of employment and to a certain extent the likelihood of being employed in the informal sector. Individuals from larger households are more likely to be inactive and to be employed in irregular informal activities. This result can be explained by the fact that Irregular activities are often part of family subsistence economy.

The estimation results for the earnings equation are reported in Table A2 in Appendix. Equation for formal workers is highly significant. Equations for informal workers have fewer significant coefficients partly because of much smaller sample sizes of these groups, especially in the early years of the period. However, many of selection terms are significant – at least, at 10 percent level. We will describe the general regularities that emerge from regressions.

Gender differences exist in the formal sector and all types of informal employment and they are roughly of the same magnitude. The returns to education are higher in the formal sector – for informal workers the returns to education are almost always insignificant. Surprisingly, a notable exception is presented by those in irregular informal activities. Coefficients of university education dummy are positive in most years and significant in a few years. This finding may reflect the diversity of activities in this subsector: it covers a wide range of production – from simple personal services to business services provided, for instance, by highly qualified women on maternity leave.

Informal work seems to be much less rewarding in urban areas. Living in the city is associated with wage premium for formal workers and self-employed and their employees. Agglomeration effects do not affect wages of those who work without contract in the informal sector while the city premium is unexpectedly negative for irregular activities. Wages in irregular informal activities are often higher for ethnic minorities probably because a large share of micro-firms are run by non-Russians and recruiting is made via family and diaspora channels. However, the coefficient of Russian dummy is significantly negative in the formal sector for a number of years. We think that this “penalty” reflects more stringent selection in non-Russians into the formal sector. Those non-Russians who get jobs in the formal sector in fact possess

better unobserved characteristics than Russian workers with similar observed characteristics. Further research is needed for robust conclusions on the effect of ethnicity on informal work.

## 4. Earnings and variance decomposition

In this section we adopt recent advances in methodology proposed by Firpo-Fortin-Lemieux (hereafter, FFL) to decompose the differences in inequality between the formal and informal sector (Firpo et al., 2009; Fortin at al., 2011). The FFL method extends the well-known Oaxaca-Blinder approach for decomposing the differences along the entire distribution and applying it to statistics other than the mean (Blinder, 1973; Oaxaca, 1973). The method allows dividing the differences in inequality into two main sources attributed: (1) to differences in observable characteristics (composition effect) and (2) to differences in the returns to these characteristics (wage structure effect). The composition and wage structure effects can be further decomposed to the contributions of individual covariates as in standard Oaxaca-Blinder decomposition.

### *4.1 Methodology*

The FFL decomposition procedure is based on computing the recentered influence function (RIF) for the quantile $q(\tau)$ or other distributional statistics of interest and using this variable instead of the outcome variable (earnings) in regression equation. FFL define the RIF as $RIF(y;\nu) = \nu(F_Y) + IF(y;\nu)$ where $\nu$ is some parameter of the distribution and $IF(y;\nu)$ is the influence function which comes from the statistics theory (Hampel, 1974). The RIF for the variance is given by: $RIF(Y,\sigma^2) = \sigma^2 + [(Y-\mu)^2 - \sigma^2] = (Y-\mu)^2$. The RIF for quantile $q(\tau)$ has a more complicated form $RIF(y;q_\tau) = q_\tau + \frac{\tau - 1\{y \le q_\tau\}}{f_Y(q_\tau)}$, where $f_Y$ is the marginal density function of *Y* and $I\{.\}$ is an indicator function of whether the particular earnings observation y is at or under $q(\tau)$. In empirical applications the RIF for each observation is obtained by estimating the sample quantile $\hat{q}_\tau$ and the density $\hat{f}_Y(\hat{q}_\tau)$ using kernel functions.

FFL demonstrate that the RIF-regression for any distributional statistics $\nu$ ($E[RIF(Y;\nu)|X]$) can be modeled by a linear regression model:

$$E[RIF(Y;\nu)|X] = X\gamma_\nu \tag{5}$$

The coefficients of Equation (5) are estimated by OLS and represent the mean marginal effects of explanatory variables on considered parameters of the earnings distribution.

Using the estimates from Equation (5) it is possible to compute an Oaxaca-Blinder decomposition for differences between formal and informal workers as follows:

$$\hat{v}_F - \hat{v}_N = (\bar{X}_F - \bar{X}_N)\hat{\gamma}_{N,\nu} + \bar{X}_F(\hat{\gamma}_{F,\nu} - \hat{\gamma}_{N,\nu}) \tag{6}$$

The first term represents the composition effect which stands for differences explained by observable characteristics. The second term is the wage structure effect, or unexplained part of the differential[10].

*4.2 Empirical results*

In this section we explore the impact of covariates on inequality of monthly earnings with cross-sectional data. The analysis employs RIF-regressions for the 10th, 50th and 90th percentiles and the variance. The list of covariates in regressions includes gender, working time (log), a dummy for Russian nationality, age, age squared, education, urban dummy and regional dummies plus selection correction to correct for the endogeneity of the sector choice. Since the differences in earnings between formal workers and those employed without contract in the formal sector are marginally significant (see Table 4 and Fig.3) we will exclude this group form further analysis an focus on two other informal groups with apparent disparities. Moreover, the group of those employed without contract in the formal sector has extremely low sample size in 2000 (N=109).

Table A3 and A4 in Appendix report the results of the RIF-regressions for 2000 and 2010[11]. The coefficients in the tables were obtained by running a liner regression of the RIF (either for the appropriate percentile or the variance) on explanatory variables. The RIF estimates give more information about the earnings structure and inequality patterns in each type of employment.

Working time is the most important explanatory variable for the dispersion of earnings. Differences in working hours are significant in most of percentiles and variance equation and have equalizing effect on earnings in all sectors. This finding confirms conclusions from our descriptive analysis – workers higher hourly wages tend to work shorter hours.

Males and females often have differentiated response to various policy changes which can be important in the context of informality since proportion of males is higher among informal workers. Duncan and Peter (2010) find that the tax reform led to a statistically significant increase in male hours of work but had no effect on that of females. The RIF estimates confirm that gender is an important determinant of earnings but its effect varies along the distribution, across types of employment and over time. Gender has strong impact on inequality only in the formal sector as suggested by significant coefficients of the male dummy in the variance equations. The effect of gender on inequality seems to decline over time for this sector the formal sector because of increased gender gap for the bottom deciles.

[10] We also performed decomposition with reweighting in the spirit of DiNardo et al. (1996). However, the estimation results are not reliable because of small sizes of informality groups and therefore they are not reported in the paper. However, it means that the results presented in the next section may be subject to the specification error if linearity assumption in Equation (5) does not hold (Barsky et al., 2002).

[11] The RIF-regression estimations use the Stata command rifreg. The code is available from N.Fortin home page: http://faculty.arts.ubc.ca/nfortin/datahead.html.

In the formal sector sector, the gender gap in favor of males is large and significant in the middle and upper part of the distribution while at the bottom the gap was is either small (in 2010) or insignificant. In the informal sector, the gap is either insignificant for all deciles (as for self-employment in 2000) or rather uniform across the deciles of the distribution.

Age can be viewed as a crude proxy for experience. Experience has much higher effect on the formal sector distribution and generally is of low importance in the informal sector. In the formal sector experience is granted with high and stable rewards in the upper part of the distribution. At the same time experience/age was penalized among low paid formal workers in the beginning of the period. This negative effect disappeared by 2010. The evaporation of this penalty can be attributed to cohort shifts at the Russian labor market. Old cohorts with obsolete human capital gained in a command economy and with few years of experience in a market economy were leaving the market during the 2000s. These cohorts were replaced by younger cohorts who were also educated during the Soviet era but acquired more market-relevant experience since the start of economic transition. Education yields considerable and sustained economic returns only in the formal sector. There is no clear pattern of returns from education in any type of informal employment – education seems not be valued in the informal sector. However, high returns to education do not increase inequality in the formal sector as they are reasonably even across the deciles of the distribution.

Agglomeration effects are working in the formal sector where they increase earnings at all deciles of the distribution and, thus, have no significant effect on the variance. There is no city premiums in the group of self-employed and their employees. In 2000 living in urban areas was penalized among irregular workers, especially at the top of the distribution.

Controlling for selection appears to have important effect on earnings and inequality. Selection effects were especially strong in the early 2000s and essentially vanish by the end of the period. Selection into the formal sector stretches the earnings distribution in this sector: it increasing polarization in this sector. On the contrary, selection has equalizing effects on earnings from irregular employment. The importance of selection was declining over time for all forms of employment.

Comparison of the RIF estimates between the formal sector and informal self-employment segment reveals that these two segments follow very different wage-setting mechanisms in spite of having similar earnings distributions. Key human capital variables are relevant for wage determination only in the formal sector and have no effect on wages in the informal sector. Employment in the formal sector is associated with a greater use of physical capital that requires human capital acquisition on the part of the employed workers, while the informally employed often work with little or no physical capital and their human capital is underutilized.

In order to examine and further disentangle the informal sector wage gap, we apply the Oaxaca-Blinder method to decompose the differences between the $10^{th}$, $50^{th}$, $90^{th}$ and the variances in the formal sector and considered types of informal employment Formal sector workers are used as the reference group. Regional, education and selection variables are grouped into broad categories. Age is grouped with age squared into a single category.

The composition effect is not statistically significant for differences the self-employed and formal workers in 2000. Observable characteristics did not contribute much to the differentials in the early 2000s. Significant advantage of the self-employed over formal sector workers was entirely due to differences in returns to characteristics. Remarkably, the differences in returns have negative signs while the total wage structure effect is positive. As we have seen before there are positive returns to human capital and job characteristics in the formal sector and virtually no returns in the segment of self-employment. However, the positive constant term more than offsets the negative wage structure effects of observable characteristics. Therefore, unobservable characteristics were extremely important for wage-setting in the informal sector in the early 2000s. Including more covariates into the earnings equation is potentially interesting, More specifically, looking at informal employment patterns by industry and occupation merit further research.

The group of self-employed and their employees experienced dramatic changes both in the size of the gap which became negative for the bottom and the middle of distribution and in its determinants. The contribution of endowments increased in importance by 2010. Their total effect is negative for all deciles and significantly so for the median and the upper part of the distribution. The group of regional dummies has the largest negative value, suggesting that this type of employment has higher incidence in the low-wages regions. Differences in human capital endowments (education and experience) do not contribute much to the wage gap between formal workers and the self-employed, except for the very top of distribution. However, this effect is sufficiently large to reduce the differences in inequality. At the same time there are some employment-related characteristics that increase the relative earnings of the self-employed and their employees. In particular, in 2010 the self-employed worked longer hours than formal workers in all parts of distribution narrowing the negative gap in monthly earnings and lowering the differences in inequality. Higher proportion of men also boosts the relative incomes of the self-employed in all deciles, but it has no effect on the differences in the variance of earnings between the sectors.

Finally, the wage structure effect diminished over time in the percentile decompositions for the self-employed. This was the major reason for reversal of the earnings gap at the bottom and in the middle of the distribution. This reduction was caused by the decline of returns to both observable and unobservable characteristics.

At the same time we see an increase in the contribution wage structure effects in explaining the differences in equality. This developments are related to faster decline in relative earnings of low paid workers in the group of self-employed and their employees. Repeated minimum wage hikes between 2000 and 2010 enhanced position of low paid workers in the formal sector. Employers in the informal sector failed to offer comparable wage increases to their low paid employees. The “lighthouse effect” is not perfect in Russia as formal and informal wages do not grow in tandem even at the bottom of the distribution. Additionally, minimum wages could cause non-trivial employment effects. First, minimum wage increases destroy jobs in the formal sector and force workers to look for jobs in the informal sector. Low ability individuals are in greater risk to leave job in the formal sector and enter the informal sector while high ability individuals move in the opposite direction. These flows lead to changes in the composition of unobservable characteristics in both sectors. Second, wage increases in the formal and informal sector motivate taking jobs by those who were initially out labor force and encourage transitions from irregular employment to full-time employment in the formal or informal sector. All these story lines are consistent with shrinking share of irregular employment and rising share of other types of informal employment in the RLMS-HSE dataset. The behavior of selection terms in the decomposition in Table 7 provides additional evidence for changes in unobservable characteristics. These selection variables have no direct interpretations but changing signs and declining significance point at the transformation in the selection mechanisms.

At the same time minimum wages are unlikely to be the only reason for the changes in the earnings gap because they cannot explain the narrowing of the gap in upper half of the distribution. More general reasons are related to economic growth and changes in regulation and deserve further research.

Table 8 reports the decomposition results for difference in earnings between formal workers and irregular informal workers. The earnings gap was relatively stable during the whole period and if widened then only at the top of the distribution. Therefore, for formal-irregular divide major differences are not over time but across the deciles. However, composition effects are negative and significant in both years and for all percentiles. Such persistence is associated with the negative impact of working time. Irregular work shorter hours and therefore have lower earnings. Differences in working hours are the principal reason for higher inequality among irregular workers. The differences in hours are particularly striking for low deciles suggesting the higher incidence of part-time and casual work in this end of the distribution.

It is appealing that the effects of covariates are not evenly distributed through the income ranges. Apart from hours gender appears as positive and significant in several decompositions mainly for the middle and upper half of the distribution: the prevalence of males among irregular workers reduces the

**Table 7. FFL decomposition: Formal workers VS Self-employed and their employees**

| | q10 | | | | Median | | | | q90 | | | | Variance of logs | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2000 | | 2010 | | 2000 | | 2010 | | 2000 | | 2010 | | 2000 | | 2010 | |
| | Composition effect | Wage structure effect | Composition effect | Wage structure effect | Composition effect | Wage structure effect | Composition effect | Wage structure effect | Composition effect | Wage structure effect | Composition effect | Wage structure effect | Composition effect | Wage structure effect | Composition effect | Wage structure effect |
| Gender | -0.009 | -0.081 | 0.023** | 0.061 | -0.012 | -0.329*** | 0.037*** | -0.016 | 0.017 | -0.124 | 0.044*** | 0.069 | 0.011 | -0.082 | 0.009 | 0.004 |
| | *(0.022)* | *(0.130)* | *(0.009)* | *(0.040)* | *(0.013)* | *(0.075)* | *(0.010)* | *(0.032)* | *(0.023)* | *(0.131)* | *(0.013)* | *(0.050)* | *(0.026)* | *(0.147)* | *(0.009)* | *(0.045)* |
| Age | 0.061 | 2.680 | 0.005 | -0.053 | -0.006 | -2.887** | -0.002 | -0.831* | -0.064 | -3.757 | 0.011 | -0.413 | -0.167 | -5.505** | -0.008 | 0.215 |
| | *(0.093)* | *(2.344)* | *(0.010)* | *(0.575)* | *(0.056)* | *(1.351)* | *(0.008)* | *(0.465)* | *(0.098)* | *(2.384)* | *(0.013)* | *(0.726)* | *(0.113)* | *(2.683)* | *(0.012)* | *(0.656)* |
| Russian | 0.002 | -0.048 | 0.004 | -0.044 | 0.001 | -0.146 | -0.003 | 0.071 | 0.000 | -0.086 | -0.002 | 0.094 | -0.003 | 0.215 | -0.009 | 0.090 |
| | *(0.008)* | *(0.293)* | *(0.007)* | *(0.083)* | *(0.004)* | *(0.169)* | *(0.006)* | *(0.068)* | *(0.003)* | *(0.295)* | *(0.009)* | *(0.105)* | *(0.010)* | *(0.330)* | *(0.008)* | *(0.094)* |
| Working time | 0.023 | -1.801* | 0.023*** | -0.448 | 0.025* | -1.019* | 0.011** | -0.527* | -0.027 | -2.591*** | 0.014* | 0.049 | -0.029 | 0.098 | -0.022*** | 0.078 |
| | *(0.024)* | *(0.959)* | *(0.007)* | *(0.342)* | *(0.015)* | *(0.553)* | *(0.005)* | *(0.282)* | *(0.026)* | *(0.951)* | *(0.007)* | *(0.435)* | *(0.029)* | *(1.054)* | *(0.008)* | *(0.387)* |
| Education | -0.002 | 0.053 | -0.009 | -0.017 | 0.024 | -0.022 | -0.027 | -0.020 | -0.006 | -0.037 | -0.052* | -0.001 | 0.031 | -0.027 | -0.063** | 0.036* |
| | *(0.037)* | *(0.041)* | *(0.025)* | *(0.018)* | *(0.021)* | *(0.024)* | *(0.019)* | *(0.015)* | *(0.035)* | *(0.041)* | *(0.031)* | *(0.023)* | *(0.041)* | *(0.045)* | *(0.029)* | *(0.021)* |
| Urban | 0.003 | -0.298 | -0.006 | -0.062 | 0.013 | -0.149 | 0.002 | -0.116 | -0.019 | -0.401 | -0.013 | 0.005 | 0.029 | 0.254 | 0.003 | 0.016 |
| | *(0.045)* | *(0.364)* | *(0.015)* | *(0.094)* | *(0.026)* | *(0.210)* | *(0.012)* | *(0.076)* | *(0.047)* | *(0.370)* | *(0.019)* | *(0.119)* | *(0.054)* | *(0.417)* | *(0.018)* | *(0.108)* |
| Region | -0.016 | -0.132** | 0.001 | -0.012 | 0.024 | -0.038 | -0.058*** | 0.004 | 0.018 | 0.083 | -0.055* | 0.039 | 0.021 | 0.129** | -0.042 | 0.030 |
| | *(0.022)* | *(0.053)* | *(0.026)* | *(0.024)* | *(0.017)* | *(0.030)* | *(0.021)* | *(0.019)* | *(0.023)* | *(0.053)* | *(0.033)* | *(0.030)* | *(0.028)* | *(0.059)* | *(0.030)* | *(0.027)* |
| Selection | -0.017 | -4.877* | -0.049 | 2.056** | -0.008 | -3.012** | -0.012 | 1.234 | 0.091 | 2.591 | -0.023 | 1.772 | 0.002 | 1.698 | 0.069 | -1.079 |
| | *(0.119)* | *(2.543)* | *(0.051)* | *(0.952)* | *(0.077)* | *(1.465)* | *(0.040)* | *(0.759)* | *(0.128)* | *(2.593)* | *(0.063)* | *(1.196)* | *(0.144)* | *(2.921)* | *(0.059)* | *(1.091)* |
| Constant | | 4.943 | | -1.549 | | 7.824*** | | 0.086 | | 4.574 | | -1.440 | | 3.170 | | 0.746 |
| | | *(4.070)* | | *(1.064)* | | *(2.345)* | | *(0.846)* | | *(4.185)* | | *(1.334)* | | *(4.738)* | | *(1.220)* |
| Total | 0.045 | 0.439*** | -0.008 | -0.068** | 0.061 | 0.222*** | -0.052*** | -0.115*** | 0.011 | 0.252** | -0.075*** | 0.174*** | -0.105 | -0.050 | -0.063*** | 0.136*** |
| | *(0.073)* | *(0.106)* | *(0.022)* | *(0.034)* | *(0.045)* | *(0.061)* | *(0.019)* | *(0.027)* | *(0.076)* | *(0.108)* | *(0.028)* | *(0.043)* | *(0.086)* | *(0.122)* | *(0.024)* | *(0.039)* |
| Raw difference | 0.484*** | | -0.076** | | 0.283*** | | -0.167*** | | 0.263*** | | 0.099*** | | -0.155* | | 0.073** | |
| | *(0.084)* | | *(0.030)* | | *(0.051)* | | *(0.026)* | | *(0.083)* | | *(0.038)* | | *(0.092)* | | *(0.033)* | |
| Observations | 3,065 | | 7,872 | | 3,065 | | 7,872 | | 3,065 | | 7,872 | | 3,065 | | 7,872 | |

Note: Standard errors are in parentheses. ***, **, and * denote statistical significance at 1 percent, 5 percent, and 10 percent level, respectively.

**Table 8. FFL decomposition: Formal workers VS Irregular workers**

| | q10 | | | | Median | | | | q90 | | | | Variance | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2000 | | 2010 | | 2000 | | 2010 | | 2000 | | 2010 | | 2000 | | 2010 | |
| | Composition effect | Wage structure effect | Composition effect | Wage structure effect | Composition effect | Wage structure effect | Composition effect | Wage structure effect | Composition effect | Wage structure effect | Composition effect | Wage structure effect | Composition effect | Wage structure effect | Composition effect | Wage structure effect |
| Gender | 0.169* | 0.768** | 0.072 | 0.192 | 0.039*** | -0.079 | 0.023** | -0.117*** | 0.020 | -0.126* | 0.127*** | 0.283** | -0.009 | -0.187 | 0.056 | 0.146 |
| | *(0.088)* | *(0.349)* | *(0.064)* | *(0.210)* | *(0.015)* | *(0.053)* | *(0.011)* | *(0.038)* | *(0.013)* | *(0.067)* | *(0.041)* | *(0.116)* | *(0.038)* | *(0.181)* | *(0.044)* | *(0.144)* |
| Age | -0.484 | 8.115 | -0.063 | 1.074 | -0.133** | 1.338 | -0.003 | -0.631 | 0.048 | -2.074** | -0.424*** | 6.165*** | 0.255 | -5.447** | -0.157 | 2.331 |
| | *(0.412)* | *(5.250)* | *(0.209)* | *(3.748)* | *(0.057)* | *(0.816)* | *(0.036)* | *(0.663)* | *(0.064)* | *(1.043)* | *(0.127)* | *(2.060)* | *(0.212)* | *(2.730)* | *(0.145)* | *(2.562)* |
| Russian | 0.109 | -1.070 | -0.022 | 0.345 | 0.005 | -0.053 | 0.002 | 0.006 | -0.004 | -0.011 | -0.023 | 0.432 | -0.070* | 0.709* | -0.016 | 0.231 |
| | *(0.076)* | *(0.829)* | *(0.032)* | *(0.479)* | *(0.009)* | *(0.125)* | *(0.005)* | *(0.087)* | *(0.011)* | *(0.160)* | *(0.019)* | *(0.264)* | *(0.042)* | *(0.431)* | *(0.022)* | *(0.327)* |
| Working time | -1.992*** | 4.840*** | -1.847*** | 4.689*** | -0.458*** | -0.191 | -0.338*** | -0.023 | -0.263*** | -0.668* | -0.531*** | 1.094*** | 0.278* | 0.086 | 0.561*** | -0.588 |
| | *(0.304)* | *(1.151)* | *(0.201)* | *(0.655)* | *(0.045)* | *(0.296)* | *(0.035)* | *(0.177)* | *(0.047)* | *(0.394)* | *(0.102)* | *(0.398)* | *(0.146)* | *(0.627)* | *(0.127)* | *(0.456)* |
| Education | -0.445 | 0.111 | -0.366 | 0.115 | 0.004 | -0.010 | -0.043 | -0.016 | 0.003 | -0.051** | -0.552*** | 0.205*** | 0.166 | -0.083 | -0.212 | 0.119 |
| | *(0.307)* | *(0.110)* | *(0.294)* | *(0.107)* | *(0.039)* | *(0.018)* | *(0.050)* | *(0.019)* | *(0.048)* | *(0.023)* | *(0.164)* | *(0.060)* | *(0.158)* | *(0.057)* | *(0.201)* | *(0.073)* |
| Urban | 0.038 | -0.472 | 0.096 | -0.526 | 0.048* | -0.442*** | 0.050** | -0.328*** | 0.054* | -0.476*** | 0.027 | -0.190 | 0.046 | -0.151 | -0.104 | 0.495 |
| | *(0.187)* | *(0.765)* | *(0.128)* | *(0.566)* | *(0.025)* | *(0.121)* | *(0.023)* | *(0.101)* | *(0.031)* | *(0.155)* | *(0.069)* | *(0.312)* | *(0.097)* | *(0.398)* | *(0.088)* | *(0.387)* |
| Region | 0.174 | 0.039 | 0.010 | 0.170 | 0.014 | 0.068*** | -0.006 | 0.074*** | 0.027 | 0.059** | -0.066 | 0.065 | -0.003 | 0.049 | -0.071 | 0.013 |
| | *(0.149)* | *(0.158)* | *(0.099)* | *(0.156)* | *(0.019)* | *(0.023)* | *(0.018)* | *(0.027)* | *(0.024)* | *(0.030)* | *(0.057)* | *(0.086)* | *(0.073)* | *(0.082)* | *(0.067)* | *(0.107)* |
| Selection | -0.167 | 3.186 | 0.097 | 3.035 | -0.066 | 0.160 | -0.136 | 1.457** | -0.143 | 1.021 | 0.377 | 0.173 | -0.106 | 0.375 | 0.295 | -4.843* |
| | *(0.649)* | *(3.774)* | *(0.514)* | *(4.062)* | *(0.085)* | *(0.682)* | *(0.088)* | *(0.734)* | *(0.103)* | *(0.886)* | *(0.287)* | *(2.241)* | *(0.334)* | *(1.981)* | *(0.355)* | *(2.779)* |
| Constant | | -14.744** | | -8.995** | | -1.127 | | -0.902 | | 2.293* | | -7.695*** | | 5.200* | | 2.724 |
| | | *(5.806)* | | *(4.541)* | | *(0.976)* | | *(0.816)* | | *(1.259)* | | *(2.502)* | | *(3.033)* | | *(3.106)* |
| Total | -2.599*** | 0.773 | -2.021*** | 0.100 | -0.548*** | -0.336*** | -0.451*** | -0.480*** | -0.259*** | -0.032 | -1.065*** | 0.533*** | 0.556*** | 0.552** | 0.353** | 0.627*** |
| | *(0.406)* | *(0.470)* | *(0.266)* | *(0.298)* | *(0.058)* | *(0.064)* | *(0.047)* | *(0.051)* | *(0.064)* | *(0.080)* | *(0.144)* | *(0.163)* | *(0.199)* | *(0.243)* | *(0.179)* | *(0.204)* |
| Raw difference | -1.826*** | | -1.921*** | | -0.884*** | | -0.931*** | | -0.291*** | | -0.532*** | | 1.108*** | | 0.980*** | |
| | *(0.321)* | | *(0.188)* | | *(0.054)* | | *(0.036)* | | *(0.058)* | | *(0.101)* | | *(0.153)* | | *(0.125)* | |
| Observations | 3,171 | | 7,576 | | 3,171 | | 7,576 | | 3,171 | | 7,576 | | 3,171 | | 7,576 | |

Note: Standard errors are in parentheses. ***, **, and * denote statistical significance at 1 percent, 5 percent, and 10 percent level, respectively.

differentials in these parts of the distribution. For low deciles gender acts predominantly through higher premiums to male. In 2010 gender has no effect on earnings differentials in low quantiles. Human capital deficits among irregular workers also contribute to widening of the earnings gaps, especially for high-paid workers. Irregular work is relatively more common in the urban areas and this composition effect reduces the gap. However, the positive composition effect is dominated by the negative returns effect. As a result, the overall city effect is highly unequalizing except for low paid workers whose earnings do not depend on settlement types. Regional dummies also affect only the upper of distribution. The effect of regional dummies equalizing – irregular work brings higher returns in the regions where pay-offs for formal sector are quite low.

**5. Conclusion**

The informal sector plays an important role in labor market adjustment in many low- and middle-income emerging economies. However little is known about the formal-informal earnings gaps beyond the mean and generally about the impact of informality on the earnings inequality. This paper employs various decomposition techniques to study the formal-informal differences in earnings inequality using the data from the Russian Longitudinal Monitoring Survey for 2000-2010. New results complement the existing literature on informality in Russia in several ways.

First, we find that during the whole period earnings inequality was substantially higher in the informal sector. The decomposition of earnings inequality into inequality of working hours and that of hourly earnings shows that in the formal sector inequality of monthly earnings is almost completely determined by the distribution of hourly wages. Higher inequality in the informal sector is explained both by larger within-sector differences in working hours and higher dispersion of highly earnings. The distributions of formal and informal hourly earnings are much closer than distributions of hours.

Second, earning inequality was declining in both sectors between 2000 and 2010. The analysis in this paper shows that the underlying force was the contraction of the share of workers employed in irregular activities and further ‘regularization’ of the informal sector. This result is important since in the beginning of the period irregular activities were the most common form of informal work in Russia.

Third, standard measures of returns to the informal employment at the mean fail to capture the important heterogeneity found in the data. This finding illustrates the explicit dealing with the sector choice when studying the formal-informal earnings gaps in Russia. In particular, we find that for many workers self-employment and working as an employee for the self-employed is as desirable as working in the formal sector. For most of the period this segment offered comparable or higher wages with fewer human capital requirements. Informal workers working without contract in the formal sector receive comparable earnings throughout the earnings distribution. Employers keep pay similar wage but save on

social security and taxes, offer lower fringe benefits and no job protection. Irregular informal workers are the only segment within the informal sector that follows common beliefs about the informal sector as comprised of subsistence-level production activities, motivated by the need for survival and characterized by low levels of income, productivity, skills, technology and capital, and weak linkages with the rest of the economy.

Forth, our study shows the importance of distributional analysis. Our results confirm our results conform to the stylized fact that workers in the formal sector have better observed characteristics at all wage levels, but these characteristics are rewarded mostly in the formal sector and generally irrelevant for wage setting in the informal sector. That is not necessarily true for unobserved characteristics – our results point towards better unobserved characteristics of the self-employed, particularly in the early 2000s. Differences in wage determination across quantiles are specifically striking for irregular informal workers.

Our analysis highlights several phenomena that merit further research. We focus on monetary remuneration and do not account for other cash or non-pecuniary advantages attached to a particular sector. Our analysis also ignores the differences in distribution of formal and informal workers across industries and occupations. The ways to account for industry specifics and non-monetary components of compensation in the formal and informal sector are fruitful area for future research.

**References**


Barlet, M. (2013). Urban Labor Market Segmentation in West Africa, in P.De Vreyer, F.Roubaud (eds.) ***Urban Labor Markets in Sub-Saharan Africa***, World Bank, Washington DC.

Barsky, R., Bound, J., Lupton, J. (2002). Accounting for the Black-White Wealth Gap: A Nonparametric Approach, *Journal of the American Statistical Association*, 97: 663–673.

Buehn, A., Schneider, F. (2012). Shadow Economies around the World: Novel Insights, Accepted Knowledge, and New Estimates, *International Tax and Public Finance*, 19 (1): 139–171.

Blinder, A. (1973). Wage Discrimination: Reduced Form and Structural Estimates, *Journal of Human Resources*, 8(4): 436–455.

Bourguignon, F., Fournier, M., Gurgand, M. (2007). Selection Bias Corrections Based on the Multinomial Logit Model: Monte Carlo Comparisons, *Journal of Economic Surveys*, 21(1): 174-205.

Brainerd, E. (1998). Winners and Losers in Russia's Economic Transition, *American Economic Review*, 88(5): 1094–1116.

Comola, M., De Mello, L. (2009). The Determinants of Employment and Earnings in Indonesia: A Multinomial Selection Approach, OECD Economics Department Working Paper No. 690.

Dubin, J., McFadden, D. (1984). An Econometric Analysis of Residential Electric Appliance Holdings and Consumption, *Econometrica*, 52: 345-362.

De Soto, H. (1989). ***The Other Path: The Invisible Revolution in the Third World,*** New York: Harper and Row.

DiNardo, J., Fortin, N., Lemieux, T. (1996). Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach, *Econometrica*, 64(5): 1001–1046.

Doiron, D., Barrett, G. (1996). Inequality in Male and Female Earnings: The Role of Hours and Wages, *Review of Economics and Statistics*, 78(3): 410–420.

Earle, J., Sakova, Z. (2000). Business Start-Ups or Disguised Unemployment? Evidence on the Character of Self-Employment from Transition Economies, *Labour Economics*, 7(5): 575–601.

Fields, G. (2005). A Guide to Multisector Labor Markets Models, Paper prepared for the World Bank labor Market Conference.

Firpo, S., Fortin, N., Lemieux, T. (2009). Unconditional Quantile Regressions, *Econometrica*, 77(3): 953–973.

Fortin, N., Lemieux, T. (1997). Institutional Changes and Rising Wage Inequality: Is there a Linkage*? Journal of Economic Perspectives*, 11(2): 75–96.

Fortin, N., Lemieux, T., Firpo, S. (2011). Decomposition Methods, in O.Ashenfelter, D.Card (eds.) ***Handbook of Labor Economics***, Vol. 4A, Amsterdam: North-Holland.

Gimpelson, V., Lukiyanova, A. (2009). Are Public Sector Workers Underpaid in Russia? Estimating the Public-Private Wage Gap, IZA Discussion Paper No. 3941.

Gorodnichenko, Y., Martinez-Vazquez, J., Sabirianova Peter, K. (2009). Myth and Reality of Flat Tax Reform: Micro Estimates of Tax Evasion Response and Welfare Effects in Russia, *Journal of Political Economy*, 117(3): 504–554.

Gorodnichenko, Y., Sabirianova Peter, K. Stolyarov, D. (2010). Inequality and Volatility Moderation in Russia: Evidence from Micro-level Panel Data on Consumption and Income, *Review of Economic Dynamics*, 13: 209–237.

Gunatilaka, R., Vodopivec, M. (2010). Labor Market Institutions and Market Segmentation in Sri Lanka. Chapter 2 in Gunatilaka, R., Vodopivec, M., Mayer, M. ***The Challenge of Youth Employment in Sri Lanka***, World Bank, Washington DC.

Günther, I.,Launov, A. (2012). Informal Employment in Developing Countries: Opportunity or Last Resort? *Journal of Development Economics*, 97(1): 88-98.

Hampel, F. (1974). The Influence Curve and its Role in Robust Estimation, *Journal of the American Statistical Association*, 69(436): 383–393.

Harris, J., Todaro M. (1970). Migration, Unemployment and Development: A Two Sector analysis, *American Economic Review*, 60(1): 126–142.

Ivanova, A., Keen, M., Klemm, A. (2005). The Russian 'Flat Tax' Reform, *Economic Policy*, 43: 397–444.

Krstic, G., Sanfey, P. (2007). Mobility, Poverty and Well-Being among the Informally Employed in Bosnia and Herzegovina, *Economic Systems*, 31(3): 311–335.

Krstic, G., Sanfey, P. (2011). Earnings Inequality and the Informal Economy: Evidence from Serbia, *Economics of Transition*, 19(1): 179–199.

Lehmann, H., Razzolini, T., Zaiceva, A. (2012). Job Separations, Job Loss and Informality in the Russian Labor Market, in H.Lehmann, K.Tatsiramos (eds.) Informal Employment in Emerging and Transition Economies (Research in Labor Economics, Vol.34): 257–290.

Lehmann, H., Muravyev, A., Razzolini, T., Zaiceva, A. (2013). The Wage and Non-Wage Costs of Displacement in Boom Times: Evidence from Russia, *Journal of Comparative Economics*, 41: 1184–1201.

Lehmann, H., Zaiceva, A. (2013). Re-Defining Informal Employment and Measuring its Determinants: Evidence from Russia, IZA Discussion Paper No. 7844.

Lemos, S. (2009). Minimum Wage Effects in a Developing Country, *Labour Economics*, 16: 224–237.

Lee, L. (1983). Generalized econometric models with selectivity, *Econometrica*, 51: 507-512.

Maloney, W. (2004). Informality revisited, *World Development*, 32(7): 1159–1178.

Milanović, B. (1999). Explaining the Increase in Inequality during Transition, *Economics of Transition*, 7(2): 299–341.

Mitra, P., Yemtsov, R. (2006). Increasing Inequality in Transition Economies: Is There More to Come? World Bank, Policy Research Working Paper No. 4007.

Muravyev, A., Oshchepkov, A. (2013). Minimum Wages and Labor Market Outcomes: Evidence from the Emerging Economy of Russia, HSE Working Papers, Series: Economics, WP BRP 29/EC/2013.

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets, *International Economic Review*, 14(3): 693–709.

Slonimczyk, F. (2012). The Effect of Taxation on Informal Employment: Evidence from the Russian Flat Tax Reform, in H.Lehmann, K.Tatsiramos (eds.) Informal Employment in Emerging and Transition Economies (Research in Labor Economics, Vol.34): 55–99.

Slonimczyk, F. (2013). Informality as a Stepping Stone, Higher School of Economics Working Paper WP3/2013/05, Moscow.

Rosser, J., Rosser, M., Ahmed, E. (2000). Income Inequality and the Informal Economy in Transition Economies, *Journal of Comparative Economics*, 28(1): 156–171.

Rutkowski, J. (1996). Changes in the Wage Structure during Economic Transition in Central and Eastern Europe, World Bank Technical Paper No. 340 – Social Challenges Series.

Zudina, A. (2013). Informal Employment and Subjective Social Status: The Case of Russia, Higher School of Economics Working paper WP3/2013/01, Moscow (in Russian).

## Appendix

### Table A1. Estimation of sector choice equations

| | 2000 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 |
|---|---|---|---|---|---|---|---|---|---|---|
| *Not working* | | | | | | | | | | |
| Children aged 0-6 (1=yes) | 0.286** | 0.370*** | 0.387*** | 0.374*** | 0.181 | 0.272** | 0.311*** | 0.212** | 0.222** | 0.483*** |
| Children aged 0-6 * Male | -0.861*** | -1.378*** | -1.802*** | -1.280*** | -1.073*** | -1.459*** | -1.792*** | -1.299*** | -1.447*** | -1.576*** |
| Children aged 7-16 (1=yes) | -0.139 | -0.084 | -0.091 | -0.009 | 0.025 | -0.215** | -0.112 | -0.035 | -0.287*** | -0.198** |
| Children aged 7-16 * Male | 0.329** | -0.005 | 0.092 | -0.264 | -0.027 | 0.092 | 0.073 | -0.137 | 0.065 | 0.135 |
| HH size | 0.074** | 0.094*** | 0.053*** | 0.078*** | 0.060*** | 0.097*** | 0.106*** | 0.098*** | 0.117*** | 0.078*** |
| Male | -0.013 | 0.110 | 0.068 | 0.092 | -0.188** | -0.065 | -0.007 | -0.013 | -0.047 | -0.010 |
| Married | 0.020 | -0.004 | 0.084 | 0.019 | -0.035 | 0.075 | 0.034 | 0.036 | 0.144* | 0.132** |
| Married * Male | -0.861*** | -0.858*** | -0.757*** | -0.986*** | -0.606*** | -0.788*** | -0.894*** | -0.760*** | -0.606*** | -0.793*** |
| Age | -0.414*** | -0.420*** | -0.447*** | -0.414*** | -0.416*** | -0.417*** | -0.417*** | -0.395*** | -0.410*** | -0.416*** |
| Age squared | 0.005*** | 0.005*** | 0.006*** | 0.005*** | 0.005*** | 0.005*** | 0.005*** | 0.005*** | 0.005*** | 0.005*** |
| Russian | -0.331*** | -0.412*** | -0.536*** | -0.465*** | -0.646*** | -0.417*** | -0.424*** | -0.208*** | -0.329*** | -0.523*** |
| Education (upper secondary) | | | | | | | | | | |
| Low secondary and less | 0.405*** | 0.499*** | 0.423*** | 0.601*** | 0.484*** | 0.439*** | 0.507*** | 0.548*** | 0.593*** | 0.452*** |
| College | -0.651*** | -0.653*** | -0.563*** | -0.630*** | -0.716*** | -0.623*** | -0.685*** | -0.576*** | -0.785*** | -0.825*** |
| University | -1.070*** | -0.999*** | -1.055*** | -1.055*** | -1.146*** | -1.072*** | -1.225*** | -1.182*** | -1.185*** | -1.324*** |
| Urban | -0.508*** | -0.530*** | -0.441*** | -0.499*** | -0.614*** | -0.588*** | -0.507*** | -0.491*** | 0.500*** | -0.409*** |
| Intercept | 7.537*** | 7.512*** | 8.203*** | 7.385*** | 7.914*** | 7.438*** | 7.257*** | 6.575*** | 7.158*** | 7.464*** |
| *Employed without contract in the corporate sector* | | | | | | | | | | |
| Children aged 0-6 (1=yes) | 0.551 | 0.553* | -0.035 | 0.130 | -0.220 | 0.449** | 0.318 | -0.012 | -0.041 | 0.283 |
| Children aged 0-6 * Male | -1.136** | -0.843** | 0.371 | 0.208 | -0.094 | -0.178 | -0.077 | 0.091 | 0.331 | -0.196 |
| Children aged 7-16 (1=yes) | -0.326 | -0.116 | -0.183 | -0.465 | -0.190 | -0.34 | -0.181 | 0.039 | -0.008 | -0.248 |
| Children aged 7-16 * Male | 0.482 | -0.042 | 0.276 | -0.048 | 0.120 | 0.043 | -0.111 | 0.104 | 0.209 | 0.526** |
| HH size | 0.046 | 0.013 | -0.131** | -0.018 | 0.014 | 0.063 | -0.069 | -0.028 | 0.039 | 0.045 |
| Male | 0.121 | 0.240 | 0.318 | 0.254 | 0.244 | 0.445*** | 0.634*** | 0.258 | 0.210 | 0.354*** |
| Married | -0.566* | -0.365 | -0.279 | -0.431** | -0.390* | -0.369** | -0.272 | -0.715*** | -0.320 | -0.598*** |
| Married * Male | 0.113 | -0.060 | -0.633** | -0.120 | -0.189 | -0.362 | -0.484* | -0.056 | -0.211 | 0.082 |
| Age | 0.006 | -0.082* | -0.131*** | -0.013 | -0.112*** | -0.066** | -0.091** | -0.103*** | -0.076*** | -0.081*** |
| Age squared | -0.001 | 0.001 | 0.001*** | 0.000 | 0.001** | 0.001 | 0.001** | 0.001*** | 0.001** | 0.001*** |
| Russian | 0.534 | -0.207 | 0.139 | -0.053 | -0.268 | -0.096 | -0.242 | -0.018 | -0.420*** | -0.453*** |
| Education (upper secondary) | | | | | | | | | | |
| Low secondary and less | 0.391 | -0.276 | 0.153 | 0.206 | 0.185 | 0.157 | -0.093 | 0.165 | 0.158 | 0.128 |
| College | -0.819*** | -0.612*** | -0.588*** | -0.673*** | -0.500*** | -0.558** | -0.520*** | -0.498*** | -0.671*** | -0.558*** |
| University | -1.137*** | -1.405*** | -1.032*** | -0.954*** | -1.369*** | -1.089*** | -0.981*** | -1.280*** | -1.349*** | -1.068*** |
| Urban | 1.040*** | 0.214 | 0.263 | 0.488*** | 0.228 | 0.522*** | 0.511*** | 0.250* | 0.352** | 0.721*** |
| Intercept | -3.685*** | -0.537 | 0.688 | -1.816** | 0.161 | -1.260** | -0.535 | -0.233 | -0.467 | -0.878* |
| *Self-employed and their employees* | | | | | | | | | | |
| Children aged 0-6 (1=yes) | -0.377 | -0.438 | -0.233 | -0.550** | -0.406 | -0.103 | -0.328 | -0.029 | 0.178 | 0.111 |
| Children aged 0-6 * Male | 0.204 | 0.549 | 0.290 | 0.496 | 0.337 | 0.122 | 0.281 | -0.009 | 0.174 | -0.096 |

| | 2000 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 |
|---|---|---|---|---|---|---|---|---|---|---|
| Children aged 7-16 (1=yes) | 0.047 | -0.568*** | -0.047 | -0.470** | -0.336 | 0.054 | -0.675*** | 0.051 | -0.322* | -0.290** |
| Children aged 7-16 * Male | 0.033 | 0.509* | 0.342 | 0.637** | 0.421 | 0.236 | 0.529** | 0.337 | 0.618** | 0.576*** |
| HH size | -0.059 | 0.079* | -0.137*** | -0.064 | 0.037 | -0.012 | 0.076** | -0.028 | 0.030 | 0.014 |
| Male | 0.345 | -0.087 | -0.247 | -0.038 | 0.067 | 0.144 | 0.248 | 0.452*** | -0.027 | 0.274** |
| Married | -0.178 | -0.174 | -0.240 | -0.090 | -0.073 | -0.324* | -0.035 | -0.253 | -0.380** | -0.121 |
| Married * Male | -0.169 | -0.101 | 0.098 | -0.109 | -0.104 | 0.007 | -0.278 | -0.517** | -0.309 | -0.349** |
| Age | 0.161** | 0.105** | 0.105** | 0.094** | 0.074* | 0.027 | 0.049 | 0.031 | 0.031 | 0.056** |
| Age squared | -0.003*** | -0.002** | -0.002** | -0.002*** | -0.001** | -0.001 | -0.001** | -0.001 | -0.001 | -0.001** |
| Russian | -0.143 | 0.093 | -0.364** | -0.231 | -0.221 | -0.390** | -0.348*** | -0.227 | -0.187 | -0.400*** |
| Education (upper secondary) | | | | | | | | | | |
| Low secondary and less | 0.132 | 0.186 | -0.016 | 0.428** | 0.140 | 0.360** | -0.046 | -0.050 | 0.123 | 0.093 |
| College | -0.213 | 0.006 | -0.319** | 0.076 | -0.178 | -0.130 | -0.136 | -0.206 | -0.496*** | -0.463*** |
| University | -0.626*** | -0.642*** | -0.901*** | -0.644*** | -0.877*** | -0.982*** | -1.217*** | -0.845*** | -1.146*** | -0.888*** |
| Urban | 0.931*** | 0.203 | -0.131 | -0.076 | 0.062 | 0.130 | -0.200* | 0.009 | -0.228* | -0.246*** |
| Intercept | -4.827*** | -4.269*** | -2.834** | -3.067*** | -3.322*** | -2.344*** | -2.453*** | -2.677*** | -2.192*** | -2.710*** |
| *Irregular informal workers* | | | | | | | | | | |
| Children aged 0-6 (1=yes) | 0.217 | 0.233 | 0.216 | -0.185 | 0.167 | -0.097 | 0.139 | -0.295 | -0.041 | -0.039 |
| Children aged 0-6 * Male | -0.229 | -0.484 | -0.356 | 0.138 | -0.441 | -0.193 | -0.511 | 0.273 | -0.560 | -0.416 |
| Children aged 7-16 (1=yes) | 0.401* | 0.119 | -0.080 | 0.068 | -0.089 | -0.227 | -0.198 | 0.217 | -0.221 | -0.656*** |
| Children aged 7-16 * Male | -0.014 | 0.422 | 0.459* | 0.326 | 0.467 | 0.361 | 0.557* | -0.184 | 0.801** | 0.908*** |
| HH size | 0.070* | 0.063* | 0.067** | 0.083** | 0.070* | 0.140*** | 0.090** | 0.100** | 0.009 | 0.053 |
| Male | 0.575*** | 0.752*** | 0.811*** | 0.501*** | 0.435*** | 0.570*** | 0.651*** | 0.533*** | 0.462*** | 0.484*** |
| Married | -0.842*** | -0.469*** | -0.606*** | -0.751*** | -0.549*** | -0.493*** | -0.562** | -0.628*** | -0.309* | -0.229 |
| Married * Male | -0.281 | -0.664*** | -0.656** | -0.229 | -0.353 | -0.566** | -0.817** | -0.386 | -0.245 | -0.530** |
| Age | -0.182*** | -0.211*** | -0.177*** | -0.152*** | -0.194*** | -0.143*** | -0.160*** | -0.147*** | -0.171*** | -0.152*** |
| Age squared | 0.002*** | 0.003*** | 0.002*** | 0.002*** | 0.002*** | 0.002*** | 0.002*** | 0.002*** | 0.002*** | 0.002*** |
| Russian | -0.471*** | -0.559*** | -0.373*** | -0.206 | -0.591*** | -0.372*** | 0.029 | -0.319** | -0.020 | -0.332** |
| Education (upper secondary) | | | | | | | | | | |
| Low secondary and less | 0.247 | 0.343** | 0.139 | 0.641*** | 0.226 | 0.228* | 0.442*** | 0.156 | 0.140 | 0.232* |
| College | -0.474*** | -0.613*** | -0.849*** | -0.628*** | -0.641*** | -0.668*** | -0.656*** | -0.827*** | -0.745*** | -0.989*** |
| University | -0.986*** | -0.716*** | -0.972*** | -0.981*** | -1.010*** | -1.191*** | -0.945*** | -1.033*** | -0.935*** | -1.476*** |
| Urban | -0.533*** | -0.406*** | -0.675*** | -0.480*** | -0.836*** | -0.724*** | -0.537*** | -0.636*** | -0.554*** | -0.569*** |
| Intercept | 1.846*** | 1.857** | 1.412*** | 0.692 | 1.957*** | 0.385 | 0.299 | -0.057 | 1.027** | 0.673 |
| Regions | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES |
| N | 7576 | 9031 | 9140 | 9202 | 8855 | 10585 | 10354 | 10477 | 10389 | 15813 |
| Pseudo R2 | 0.2780 | 0.2833 | 0.2910 | 0.2731 | 0.2738 | 0.2708 | 0.2779 | 0.2686 | 0.2599 | 0.2658 |

Note: Formal sector workers are the base category. Standard errors (available upon request) are derived from 50 bootstrap replications, with ***, **, and * denoting statistical significance at 1 percent, 5 percent, and 10 percent level, respectively.

**Table A2. Earnings equations with selection terms**

| | 2000 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 |
|---|---|---|---|---|---|---|---|---|---|---|
| *Formal sector* | | | | | | | | | | |
| Male | 0.390*** | 0.428*** | 0.416*** | 0.375*** | 0.373*** | 0.323*** | 0.374*** | 0.407*** | 0.335*** | 0.326*** |
| Age | 0.006 | 0.048** | 0.018 | -0.003 | -0.005 | 0.001 | 0.018 | 0.036** | 0.029* | 0.040*** |
| Age squared | 0.000 | -0.001** | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000*** | 0.000* | -0.001*** |
| Russian | -0.037 | -0.040 | -0.060 | -0.066** | -0.141*** | -0.096*** | -0.036 | -0.078*** | 0.018 | -0.029 |
| Working time (ln) | 0.366*** | 0.342*** | 0.353*** | 0.316*** | 0.251*** | 0.348*** | 0.285*** | 0.322*** | 0.327*** | 0.298*** |
| Education (upper secondary) | | | | | | | | | | |
| Low secondary and less | -0.136* | -0.075 | -0.098* | -0.099** | -0.045 | -0.113*** | -0.092** | -0.115*** | -0.085* | -0.092** |
| College | 0.116** | 0.113** | 0.137*** | 0.040 | -0.048 | 0.074* | 0.045 | 0.069* | 0.065 | 0.086*** |
| University | 0.323*** | 0.450*** | 0.380*** | 0.329*** | 0.418*** | 0.398*** | 0.346*** | 0.360*** | 0.426*** | 0.403*** |
| Urban | 0.361*** | 0.445*** | 0.425*** | 0.438*** | 0.249*** | 0.284*** | 0.337*** | 0.245*** | 0.180*** | 0.173** |
| Selection terms | | | | | | | | | | |
| Non-employment | 0.353 | -0.168 | 0.331** | 0.259* | 0.256 | 0.308*** | 0.147 | 0.297*** | 0.227* | 0.122* |
| Formal workers | -0.330** | 0.010 | -0.073 | -0.162 | -0.229 | -0.122 | -0.045 | -0.031 | 0.004 | 0.032 |
| Without contract in the formal sector | -3.861** | 0.273 | -1.069 | -0.148 | -1.375 | -0.798 | -0.156 | -1.983*** | -1.471 | -2.171*** |
| Self-employed and their employees | 0.436 | -1.392* | 0.609 | -0.408 | -1.349 | -0.706 | -0.312 | 1.874*** | 0.530 | 0.528 |
| Irregular informal workers | 1.235 | 0.949 | 0.574 | -0.139 | 0.922 | 0.626 | 0.948* | 1.445** | 1.209** | 1.213* |
| Intercept | 4.812*** | 4.663*** | 5.704*** | 6.509*** | 7.227*** | 6.844*** | 7.034*** | 6.950*** | 6.865*** | 6.890*** |
| *Employed without contract in the formal sector* | | | | | | | | | | |
| Male | 0.517 | 0.540*** | 0.504** | 0.330** | 0.654*** | 0.294*** | 0.557*** | 0.217** | 0.451*** | 0.347*** |
| Age | 0.270 | 0.028 | 0.076 | -0.002 | 0.109 | 0.054 | 0.060 | 0.020 | 0.053 | 0.084** |
| Age squared | -0.004 | -0.000 | -0.001 | 0.000 | -0.002* | -0.001 | -0.001 | 0.000 | -0.001 | -0.001** |
| Russian | 0.633 | -0.012 | -0.390* | -0.160 | -0.094 | -0.057 | 0.066 | -0.008 | 0.042 | -0.132 |
| Working time (ln) | 0.199 | 0.474*** | 0.360** | 0.632*** | 0.627*** | 0.401*** | 0.386*** | 0.509*** | 0.421*** | 0.431*** |
| Education (upper secondary) | | | | | | | | | | |
| Low secondary and less | -0.287 | -0.024 | -0.298 | 0.109 | -0.127 | -0.106 | -0.010 | -0.003 | -0.016 | -0.034 |
| College | 0.125 | 0.416* | -0.042 | 0.079 | 0.236 | 0.181 | 0.163 | 0.293* | 0.137 | 0.148 |
| University | 0.076 | 1.004** | 0.424 | 0.090 | 0.074 | 0.415* | 0.329 | 0.123 | 0.064 | 0.231* |
| Urban | 1.051 | 0.204 | 0.159 | 0.380* | 0.589* | 0.211 | 0.117 | 0.156 | 0.198 | 0.193 |
| Selection terms | | | | | | | | | | |
| Non-employment | 1.219 | 0.172 | 0.132 | -0.013 | -0.850 | -0.446 | 0.358 | 1.268** | -0.188 | 0.177 |
| Formal workers | 1.401 | 1.058 | 0.554 | -1.969 | -0.010 | -0.528 | 0.822 | 0.052 | -0.597 | 0.902 |
| Without contract in the formal sector | 0.688 | -0.550 | -0.327 | 0.050 | 1.066 | -0.010 | -0.052 | -0.162 | -0.052 | -0.053 |
| Self-employed and their employees | 4.475 | -0.998 | 0.316 | 3.195 | 1.766 | -2.312 | 0.617 | 2.350 | 0.789 | 0.166 |
| Irregular informal workers | -2.419 | 0.525 | 2.490 | -0.950 | 1.146 | 0.424 | 3.516 | 0.197 | 1.021 | 1ю564 |
| Intercept | 0.158 | 6.643** | 6.434*** | 4.187** | 0.475 | 5.062*** | 6.934*** | 7.263*** | 5.521*** | 6.6480*** |
| *Self-employed and their employees* | | | | | | | | | | |
| Male | 0.079 | 0.259 | 0.526** | 0.599*** | 0.199 | 0.325** | 0.313*** | 0.332*** | 0.330*** | 0.470*** |
| Age | -0.084 | 0.074 | 0.035 | 0.125*** | -0.054 | -0.030 | 0.009 | 0.004 | 0.038 | 0.009 |
| Age squared | 0.001 | -0.001* | 0.000 | -0.002*** | 0.001 | 0.000 | 0.000 | 0.000 | -0.001 | -0.000 |
| Russian | -0.029 | -0.053 | 0.344 | 0.018 | -0.222 | -0.050 | -0.166 | -0.194 | -0.088 | 0.009 |

| | 2000 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 |
|---|---|---|---|---|---|---|---|---|---|---|
| Working time (ln) | 0.054 | 0.426*** | 0.487*** | 0.196** | 0.383*** | 0.383*** | 0.439*** | 0.373*** | 0.475*** | 0.249*** |
| Education (upper secondary) | | | | | | | | | | |
| Low secondary and less | 0.003 | -0.128 | -0.135 | -0.048 | -0.005 | -0.089 | -0.110 | 0.070 | -0.136 | -0.171** |
| College | -0.307** | -0.107 | 0.300 | 0.329** | -0.046 | -0.087 | -0.059 | -0.184 | -0.055 | -0.005 |
| University | -0.121 | 0.269 | 0.546 | 0.610*** | 0.759 | 0.229 | 0.348 | 0.171 | 0.159 | 0.287** |
| Urban | 0.182 | 0.468** | 0.536*** | 0.467*** | 0.139 | 0.158 | 0.354*** | 0.123 | 0.277** | 0.127 |
| Selection terms | | | | | | | | | | |
| Non-employment | 2.900 | -0.094 | 0.526 | 0.383 | 1.057 | 0.865 | 0.456 | 0.955 | 0.758* | 0.386 |
| Formal workers | -0.377 | -0.098 | 3.386* | 1.047 | 1.626 | 0.100 | -0.334 | 0.296 | 0.036 | 0.872 |
| Without contract in the formal sector | -2.554 | 7.681** | 7.690** | 0.513 | 3.183 | 1.611 | 2.146 | 0.479 | 2.569 | -0.773 |
| Self-employed and their employees | 0.514 | -0.329 | -1.652** | 0.284 | -1.572** | -0.530 | -0.210 | -0.483 | -0.112 | -0.571 |
| Irregular informal workers | 1.411 | -1.881 | 1.631 | 1.497 | 1.085 | 2.294 | -0.160 | 3.910 | 0.182 | 5.247*** |
| Intercept | 8.503* | 5.350*** | 10.143*** | 4.671** | 12.158*** | 8.838*** | 6.882*** | 8.909*** | 6.749*** | 9.698*** |
| *Irregular informal workers* | | | | | | | | | | |
| Male | 0.596*** | 0.304** | 0.267* | 0.235 | 0.449*** | 0.524*** | 0.355*** | 0.253* | 0.252 | 0.399*** |
| Age | 0.085 | -0.049 | -0.045 | 0.006 | 0.060 | 0.028 | 0.021 | 0.081 | 0.000 | 0.098* |
| Age squared | -0.001 | 0.000 | 0.000 | 0.000 | -0.001 | 0.000 | 0.000 | -0.001 | 0.000 | -0.001** |
| Russian | -0.415** | -0.088 | -0.547*** | -0.202 | -0.102 | -0.292** | -0.167 | -0.370*** | 0.062 | 0.019 |
| Working time (ln) | 0.600*** | 0.623*** | 0.607*** | 0.571*** | 0.573*** | 0.597*** | 0.516*** | 0.527*** | 0.537*** | 0.518*** |
| Education (secondary) | | | | | | | | | | |
| Incomplete secondary and less | 0.021 | -0.461** | -0.051 | 0.130 | -0.125 | -0.093* | 0.056 | -0.097 | -0.179 | -0.120* |
| College | 0.157 | 0.266 | -0.122 | 0.025 | -0.099 | -0.024 | 0.003 | -0.102 | 0.224 | 0.033 |
| University | -0.011 | 0.582** | 0.326 | 0.456* | -0.029 | 0.265 | 0.158 | 0.511** | 0.583** | 0.407 |
| Urban | -0.306 | -0.324** | 0.041 | -0.221 | 0.077 | -0.299* | -0.174 | -0.368 | -0.153 | -0.429* |
| Selection terms | | | | | | | | | | |
| Non-employment | 0.014 | 0.015 | 0.000 | 0.573 | -0.250 | -0.131 | 0.082 | -0.368 | 0.156 | -0.250 |
| Formal workers | -0.561 | -0.215 | -1.886 | -0.387 | -0.574 | 0.147 | -0.876 | -0.474 | -0.368 | 1.375 |
| Without contract in the formal sector | -5.609 | -1.154 | 0.206 | -3.830 | -1.046 | -2.487 | -2.666 | -1.597 | -1.302 | -5.912** |
| Self-employed and their employees | 2.585 | -3.555* | -1.720 | 0.637 | 1.333 | -3.809** | -0.437 | 3.170 | -0.303 | -1.282 |
| Irregular informal workers | 0.961*** | -0.143 | 0.115 | 0.489 | 0.636 | 1.120*** | 0.808* | 0.434 | -0.249 | 1.552*** |
| Intercept | 0.947 | 5.676*** | 4.956*** | 4.279*** | 2.393 | 2.467 | 3.417** | 4.059** | 6.649*** | 1.690 |

Note: All equations additionally include regional dummies. Selection equations in Table A1. Standard errors (available upon request) are derived from 50 bootstrap replications, with ***, **, and * denoting statistical significance at 1 percent, 5 percent, and 10 percent level, respectively.

**Table A3. RIF-regression estimation, 2000**

| | Formal sector | | | | Informal sector | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Self-employed and their employees | | | | Irregular workers | | | |
| | q10 | Median | q90 | Variance | q10 | Median | q90 | Variance | q10 | Median | q90 | Variance |
| Male | 0.063 | 0.568*** | 0.481*** | 0.313*** | -0.112 | -0.147 | 0.212 | 0.135 | 1.731** | 0.397*** | 0.209* | -0.093 |
| | *(0.102)* | *(0.060)* | *(0.085)* | *(0.074)* | *(0.287)* | *(0.150)* | *(0.255)* | *(0.311)* | *(0.729)* | *(0.102)* | *(0.101)* | *(0.386)* |
| Age | -0.136*** | 0.009 | 0.080** | 0.138*** | 0.031 | -0.140** | -0.133 | -0.198 | 0.335 | 0.093*** | -0.033 | -0.176 |
| | *(0.042)* | *(0.024)* | *(0.036)* | *(0.031)* | *(0.145)* | *(0.070)* | *(0.132)* | *(0.151)* | *(0.275)* | *(0.034)* | *(0.046)* | *(0.143)* |
| $Age^2$ | 0.002*** | -0.000 | -0.001** | -0.002*** | -0.001 | 0.002* | 0.002 | 0.003 | -0.004 | -0.001*** | 0.000 | 0.002 |
| | *(0.001)* | *(0.000)* | *(0.000)* | *(0.000)* | *(0.002)* | *(0.001)* | *(0.002)* | *(0.002)* | *(0.003)* | *(0.000)* | *(0.001)* | *(0.002)* |
| Russian | -0.264* | -0.005 | 0.073 | 0.152 | -0.320 | -0.172 | -0.026 | 0.399 | -1.496* | -0.066 | 0.061 | 0.968** |
| | *(0.145)* | *(0.076)* | *(0.097)* | *(0.093)* | *(0.307)* | *(0.178)* | *(0.270)* | *(0.368)* | *(0.873)* | *(0.117)* | *(0.171)* | *(0.487)* |
| Working time (ln) | 0.518*** | 0.377*** | 0.327*** | -0.223*** | 0.161 | 0.176** | -0.186 | -0.204 | 1.475*** | 0.340*** | 0.195*** | -0.206* |
| | *(0.126)* | *(0.052)* | *(0.064)* | *(0.062)* | *(0.191)* | *(0.088)* | *(0.207)* | *(0.199)* | *(0.271)* | *(0.024)* | *(0.033)* | *(0.108)* |
| Education (secondary) | | | | | | | | | | | | |
| Incomplete secondary and less | -0.348** | -0.128 | -0.217** | 0.108 | -0.783** | -0.004 | 0.045 | 0.317 | -1.031 | -0.067 | 0.168 | 0.642 |
| | *(0.172)* | *(0.081)* | *(0.090)* | *(0.096)* | *(0.358)* | *(0.147)* | *(0.277)* | *(0.323)* | *(0.842)* | *(0.101)* | *(0.131)* | *(0.418)* |
| College | 0.202* | 0.120* | 0.013 | -0.130 | -0.141 | -0.277** | -0.148 | -0.171 | 1.423* | -0.075 | -0.003 | -0.747 |
| | *(0.123)* | *(0.070)* | *(0.092)* | *(0.085)* | *(0.229)* | *(0.140)* | *(0.254)* | *(0.284)* | *(0.813)* | *(0.130)* | *(0.171)* | *(0.528)* |
| University | 0.345*** | 0.296*** | 0.301** | -0.041 | -0.018 | -0.187 | 0.109 | -0.269 | 0.791 | -0.057 | 0.181 | 0.154 |
| | *(0.134)* | *(0.085)* | *(0.120)* | *(0.103)* | *(0.314)* | *(0.183)* | *(0.421)* | *(0.389)* | *(1.286)* | *(0.165)* | *(0.223)* | *(0.702)* |
| Urban | 0.425** | 0.333*** | 0.346*** | -0.044 | 0.028 | 0.134 | -0.189 | 0.295 | -0.204 | -0.257** | -0.289** | -0.245 |
| | *(0.187)* | *(0.094)* | *(0.117)* | *(0.115)* | *(0.471)* | *(0.273)* | *(0.492)* | *(0.544)* | *(1.279)* | *(0.117)* | *(0.129)* | *(0.519)* |
| Selection terms | | | | | | | | | | | | |
| Non-employment | 0.633 | 0.199 | 0.905*** | 0.509 | 2.860 | 3.894*** | 0.614 | 0.023 | 0.128 | 0.019** | -0.001 | 0.004 |
| | *(0.487)* | *(0.275)* | *(0.332)* | *(0.406)* | *(2.719)* | *(1.474)* | *(2.490)* | *(3.015)* | *(0.104)* | *(0.009)* | *(0.014)* | *(0.062)* |
| Formal workers | -2.727*** | -1.157*** | 1.157* | 2.714*** | 2.744 | 0.757 | -0.887 | -2.045 | 0.204 | 0.359 | -0.723 | -0.513 |
| | *(0.870)* | *(0.435)* | *(0.614)* | *(0.551)* | *(2.840)* | *(1.463)* | *(3.027)* | *(3.055)* | *(4.731)* | *(0.557)* | *(0.856)* | *(2.496)* |
| Without contract in the formal sector | -12.035*** | -4.390*** | 0.627 | 7.729*** | -3.109 | -1.895 | -1.682 | 4.592 | -40.114* | -1.907 | 0.873 | 20.796* |
| | *(2.870)* | *(1.571)* | *(2.047)* | *(1.860)* | *(4.560)* | *(3.188)* | *(5.576)* | *(6.769)* | *(21.899)* | *(2.725)* | *(3.303)* | *(11.393)* |
| Self-employed and their employees | 2.150 | 0.577 | -0.245 | -0.940 | 3.125 | 2.168 | -3.177 | -1.536 | 19.062* | 0.718 | -3.057 | -15.060*** |
| | *(1.668)* | *(0.912)* | *(1.234)* | *(1.143)* | *(3.701)* | *(1.862)* | *(3.218)* | *(3.696)* | *(9.787)* | *(1.353)* | *(1.946)* | *(5.654)* |
| Irregular informal workers | 0.373 | 0.267* | 0.251 | 0.092 | 0.166 | 0.088 | 0.723 | -0.671 | 3.026* | 0.842*** | -0.088 | -1.182 |
| | *(0.273)* | *(0.142)* | *(0.174)* | *(0.169)* | *(0.486)* | *(0.325)* | *(0.516)* | *(0.660)* | *(1.742)* | *(0.232)* | *(0.250)* | *(0.920)* |
| Intercept | 3.003*** | 3.193*** | 5.078*** | 1.267* | 8.057* | 10.934*** | 9.814** | 4.645 | -10.875** | 2.347*** | 7.353*** | 6.361** |
| | *(1.100)* | *(0.580)* | *(0.700)* | *(0.766)* | *(4.542)* | *(2.320)* | *(3.886)* | *(4.693)* | *(5.136)* | *(0.712)* | *(0.883)* | *(2.851)* |
| Regions | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES |
| R-squared | 0.104 | 0.187 | 0.093 | 0.061 | 0.097 | 0.211 | 0.090 | 0.073 | 0.240 | 0.461 | 0.209 | 0.091 |
| Observations | 2827 | | | | 238 | | | | 344 | | | |

Note: Selection terms are derived from the estimation reported in Table A1. Standard errors are in parentheses. ***, **, and * denote statistical significance at 1 percent, 5 percent, and 10 percent level, respectively.

**Table A4. RIF-regression estimation, 2010**

| | Formal sector | | | | Informal sector | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Self-employed and their employees | | | | Irregular workers | | | |
| | q10 | Median | q90 | Variance | q10 | Median | q90 | Variance | q10 | Median | q90 | Variance |
| Male | 0.109** | 0.433** | 0.320** | 0.091** | 0.388** | 0.516** | 0.321* | 0.051 | 0.244** | 0.398** | 0.472** | 0.101 |
| | *(0.023)* | *(0.025)* | *(0.037)* | *(0.024)* | *(0.131)* | *(0.116)* | *(0.137)* | *(0.117)* | *(0.085)* | *(0.060)* | *(0.099)* | *(0.097)* |
| Age | 0.001 | 0.045** | 0.084** | 0.033** | 0.067 | 0.078 | 0.214** | 0.054 | -0.002 | -0.005 | 0.061+ | 0.038 |
| | *(0.009)* | *(0.010)* | *(0.015)* | *(0.009)* | *(0.060)* | *(0.051)* | *(0.059)* | *(0.048)* | *(0.030)* | *(0.023)* | *(0.037)* | *(0.034)* |
| $Age^2$ | -0.000 | -0.001** | -0.001** | -0.000** | -0.001 | -0.001+ | -0.003** | -0.001 | 0.000 | 0.000 | -0.001 | -0.000 |
| | *(0.000)* | *(0.000)* | *(0.000)* | *(0.000)* | *(0.001)* | *(0.001)* | *(0.001)* | *(0.001)* | *(0.000)* | *(0.000)* | *(0.000)* | *(0.000)* |
| Russian | -0.001 | -0.042 | -0.080+ | 0.012 | -0.490** | -0.115 | 0.186 | 0.244+ | -0.051 | 0.039 | 0.027 | 0.114 |
| | *(0.031)* | *(0.034)* | *(0.048)* | *(0.031)* | *(0.128)* | *(0.141)* | *(0.177)* | *(0.143)* | *(0.084)* | *(0.069)* | *(0.116)* | *(0.102)* |
| Working time (ln) | 0.363** | 0.239** | 0.154** | -0.275** | 0.583** | 0.275** | 0.290** | -0.425** | 0.275** | 0.136** | 0.163* | -0.259** |
| | *(0.036)* | *(0.026)* | *(0.036)* | *(0.025)* | *(0.133)* | *(0.088)* | *(0.081)* | *(0.081)* | *(0.074)* | *(0.047)* | *(0.077)* | *(0.072)* |
| Education (secondary) | | | | | | | | | | | | |
| Incomplete secondary and less | -0.033 | -0.113** | -0.129** | 0.005 | 0.053 | -0.014 | -0.172 | -0.163 | -0.051 | -0.086 | -0.174+ | -0.057 |
| | *(0.042)* | *(0.038)* | *(0.041)* | *(0.034)* | *(0.120)* | *(0.130)* | *(0.160)* | *(0.129)* | *(0.100)* | *(0.071)* | *(0.092)* | *(0.100)* |
| College | 0.117** | 0.049 | 0.043 | -0.053 | -0.129 | 0.082 | 0.201 | -0.020 | -0.071 | -0.002 | -0.034 | 0.207+ |
| | *(0.036)* | *(0.038)* | *(0.048)* | *(0.034)* | *(0.187)* | *(0.170)* | *(0.173)* | *(0.156)* | *(0.091)* | *(0.081)* | *(0.131)* | *(0.117)* |
| University | 0.298** | 0.352** | 0.381** | -0.006 | -0.407 | 0.079 | 0.641* | 0.265 | 0.060 | 0.142 | 0.277 | 0.308* |
| | *(0.043)* | *(0.048)* | *(0.065)* | *(0.043)* | *(0.259)* | *(0.221)* | *(0.254)* | *(0.209)* | *(0.111)* | *(0.102)* | *(0.179)* | *(0.150)* |
| Urban | 0.132** | 0.145** | 0.096+ | -0.045 | 0.144 | 0.270 | 0.469+ | -0.025 | 0.046 | -0.016 | 0.102 | -0.022 |
| | *(0.061)* | *(0.061)* | *(0.085)* | *(0.056)* | *(0.250)* | *(0.302)* | *(0.401)* | *(0.283)* | *(0.176)* | *(0.144)* | *(0.245)* | *(0.228)* |
| Selection terms | | | | | | | | | | | | |
| Non-employment | 0.089* | 0.097** | 0.256** | 0.006 | 0.234** | 0.089 | 0.110+ | -0.085 | -0.787* | -0.191 | -0.160 | 0.197 |
| | *(0.045)* | *(0.025)* | *(0.034)* | *(0.048)* | *(0.090)* | *(0.070)* | *(0.064)* | *(0.111)* | *(0.368)* | *(0.213)* | *(0.277)* | *(0.423)* |
| Formal workers | -0.300+ | -0.052 | 0.813** | 0.702** | -0.496 | 0.303 | 3.246** | 1.045 | -1.108* | 0.050 | 1.077+ | 1.461* |
| | *(0.168)* | *(0.184)* | *(0.262)* | *(0.170)* | *(0.928)* | *(0.804)* | *(1.002)* | *(0.791)* | *(0.519)* | *(0.406)* | *(0.637)* | *(0.624)* |
| Without contract in the formal sector | -2.044** | -3.238** | -1.457+ | 1.095* | 1.484 | 1.593 | 0.121 | 0.044 | -1.586 | -1.524 | -1.745 | -1.525 |
| | *(0.567)* | *(0.601)* | *(0.822)* | *(0.551)* | *(2.937)* | *(2.740)* | *(2.765)* | *(2.595)* | *(1.358)* | *(1.214)* | *(2.067)* | *(1.779)* |
| Self-employed and their employees | 0.138 | 1.617** | 1.152* | -0.260 | 3.387 | 0.001 | -1.724 | -1.102 | -1.233 | -2.185+ | -1.638 | 0.733 |
| | *(0.553)* | *(0.541)* | *(0.580)* | *(0.483)* | *(3.366)* | *(2.681)* | *(2.275)* | *(2.360)* | *(1.386)* | *(1.131)* | *(1.694)* | *(1.643)* |
| Irregular informal workers | 0.003 | 0.253** | 0.223* | 0.165* | 0.057 | 0.223 | 0.422 | 0.229 | 0.538+ | 0.677** | 0.897** | -0.126 |
| | *(0.083)* | *(0.080)* | *(0.106)* | *(0.077)* | *(0.500)* | *(0.389)* | *(0.412)* | *(0.375)* | *(0.324)* | *(0.218)* | *(0.344)* | *(0.347)* |
| Intercept | 6.190** | 6.531** | 7.998** | 1.452** | 4.644** | 6.494** | 5.331** | 1.625 | 4.660** | 6.686** | 6.613** | 2.151+ |
| | *(0.276)* | *(0.248)* | *(0.362)* | *(0.247)* | *(1.354)* | *(1.044)* | *(1.294)* | *(1.080)* | *(1.044)* | *(0.695)* | *(1.033)* | *(1.176)* |
| Regions | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES |
| R-squared | 0.111 | 0.220 | 0.107 | 0.035 | 0.176 | 0.179 | 0.133 | 0.093 | 0.137 | 0.225 | 0.136 | 0.068 |
| Observations | 7156 | | | | 716 | | | | 420 | | | |

Note: Selection terms are derived from the estimation reported in Table A1. Standard errors are in parentheses. ***, **, and * denote statistical significance at 1 percent, 5 percent, and 10 percent level, respectively.