Blunch *IZA Journal of Labor & Development* (2015) 4:13
DOI 10.1186/s40175-015-0035-5

IZA Journal of Labor & Development
a SpringerOpen Journal

ORIGINAL ARTICLE Open Access

Bound to lose, bound to win? The financial crisis and the informal-formal sector earnings gap in Serbia

Niels-Hugo Blunch[1,2]

Correspondence: blunchn@wlu.edu
[1]Department of Economics, Washington and Lee University, 204 W. Washington St., Lexington, VA 24450, USA
[2]IZA, Bonn, Germany

Abstract

While the informal sector has received widespread attention in academic and policy arenas in recent decades, knowledge gaps and controversies remain. By examining the incidence and determinants of the formal-informal sector earnings gap for adult male dependent employees using two identical, nationally representative labor force surveys for Serbia—one just prior to the impact of the recent international financial crisis and one about a year into the crisis—for three alternative measures of informality, this paper adds to our understanding in several dimensions. Among the main results is the finding of a substantively large formal-informal sector earnings gap (favoring the formal sector)—across three alternative informality measures—which appears to have decreased substantially overall following the crisis. Additional results suggest that formal sector workers are concentrated in better paying industries and occupations and have more education and other favorable characteristics than informal sector workers, and at the same time also have higher returns to their (already favorable) characteristics overall, with education and part-time status consistently among the main drivers of the observed gap.

JEL classifications: I24, J31, J42, J46.

Keywords: Formal-informal sector earnings gap; Labor market segmentation; Earnings decomposition; Detailed earnings decomposition; International financial crisis; Serbia

1 Introduction

Since the concept of the informal sector was first introduced by Hart (1971), the informal sector has received substantial attention both in the academic literature and in the policy arena.[1] But even despite the wealth of available evidence, some issues are still poorly understood. First, what exactly is the informal sector? Despite years of research and public policy debate, it does not seem that the concept is used in a uniform, transparent way. Some definitions seem to be based on legality versus illegality (especially in terms of tax evasion), others on whether workers receive benefits such as health and/or pensions benefits or not, while still others seem to be based solely on the size of a given enterprise, characterizing small enterprises as informal and larger ones as formal. Yet, what is "small" and "large" is clearly debatable—and likely also highly contextual, as may be some of the other (potential) informality measures. For example,

it has been argued that firm size is not a very precise measure of informality for transition countries (Lehmann and Pignatti 2007).[2] It therefore seems appropriate, when desiring to explore some aspect or other of the informal sector, to adhere to not just a single definition but to apply a multifaceted framework, incorporating instead several alternative definitions of informality.[3] Second, despite the amount of empirical evidence currently available for both developed and developing countries, the evidence from the former socialist countries of Eastern Europe and Central Asia is still scarce, though starting to emerge (Lehmann 2010; 2015). The issue here is the relative scarcity of data available for these countries, though this is likely to change in the coming years as more and higher quality data becomes available from these countries. Third, only little, if anything, is known about the impact of international financial crises on the informal sector. The issue again here is lack of relevant data since, fortunately, the international crises are few and far between.

Again, while the available evidence on the extent and the nature of the informal sector in the former socialist countries of Eastern Europe and Central Asia is scarce in general, it is particularly scarce for the case of Serbia. Krstić and Sanfey (2011) is the closest in spirit to the present study, examining the extent and evolution of informality and earnings inequality in Serbia between 2002 and 2007. Of most interest to the scope and contribution of the present paper is the finding that informal employees earn substantially less than formal employees, namely about 45 percentage-points (in 2007 but not in 2002)—though when controlling for other factors, this gap decreases to about 29 percentage-points. The main emphasis is on earnings inequality *per se*, however, and so this study does not consider decompositions of the formal-informal sector earnings gap. The scope of Koettl (2010) is also on informality in Serbia but focusing on the relationship between labor taxation and benefit design, on the one hand, and work incentives, on the other, as well as the consequences of informality on the Serbian economy focusing on the fiscal implications in terms of lost revenues. The paper therefore does not directly touch on the formal-informal sector earnings gap. However, with the main conclusion that for many low skill/low earning workers it simply does not pay to work in the formal sector (due to minimum social security contributions, coupled with the design of social assistance and family benefits). Koettl (2010) provides a useful analytical motivation for the analysis undertaken here—and therefore also complements the present study well. While the scope of Kogan (2011), examining a specialized survey on school leavers in Serbia in 2006, is also on the informal sector, the focus is more narrowly on the role of credentials and contacts on job entry in Serbia. Among the main findings are that education dropouts are largely attracted to the informal sector, while the better educated are attracted to the formal sector. Networks are found to be important especially for employment in the informal sector. In turn, these results—together with those in Koettl (2010)—help to understand better the nature of the informal sector in Serbia as one of severe marginalization. This point is also the main finding in Macias (2009), which is mainly a descriptive paper, examining the 2008 Labor Force Survey.

In response to these issues, this paper specifically examines the incidence and determinants of the formal-informal sector earnings gap for male dependent employees in Serbia in the context of the recent international financial crisis. In so doing, an important motivation of this paper is to add to our currently fairly limited understanding

regarding the formal-informal sector earnings gap of the former socialist regimes of Eastern Europe and Central Asia in general (see, however, Lehmann (2015)) and of Serbia in particular by trying to understand better both the extent of the formal-informal sector earnings gap and the factors driving this gap across multiple dimensions of informality—as well as whether any such patterns, if found to be present, have changed following the onset of the crisis.

The case of Serbia is particularly interesting in this connection both because the informal sector appears quite pervasive in Serbia, employing about a third of the private sector, but also since there seems to be several plausible causes for a pervasive informal sector in Serbia, not least due to the tax-system (Koettl 2010). Additionally, and not least important, the Serbian Labor Force Survey (LFS) provides high quality data to study these issues.

Specifically, in terms of empirical methodology, the paper first examines the raw formal-informal sector earnings gaps across the three alternative measures of informality for the LFS in October 2008 (i.e., pre-crisis) and again in October 2009 (i.e., post-crisis), as follows: (1) official (non) registration as pertaining to tax-purposes; (2) lack of formal contract; and (3) lack of pension benefits. The next step of the analysis estimates Mincer earnings regressions and then decomposes the resultant formal-informal sector earnings gap in the Blinder (1973); Oaxaca (1973) tradition, as well as performs detailed decompositions following Oaxaca and Ransom (1999); Yun (2005).

The remainder of this paper is structured as follows. First, the next section reviews recent developments in Serbia, focusing at issues relevant for the formal-informal sector earnings gap. Section 3 presents the data, discusses the construction of the dependent and explanatory variables, as well as the three informality measures, and estimates the raw formal-informal sector earnings gaps for all three measures both pre- and post crisis. This is followed, in section 4, by a discussion of the estimation strategy and related issues. Section 5 presents the main results starting with the results from the Mincer earnings regressions, then the overall decompositions and finally the detailed decompositions. Section 6 concludes, discusses policy implications, and provides directions for further research.

2 Background: labor markets, legislation and informality in Serbia

With the collapse of the Berlin Wall in 1989 followed in Serbia, as in other former socialist countries, formally the transition towards a market economy. For the case of Serbia, though, those first 10 years were no race towards a market economy, for sure—as perhaps most clearly summarized by Babović (2008: 13), "During the last decade of the twentieth century, Serbian society was characterized by a state of blocked transformation that included the obstruction of essential changes in market economy and political democracy by the ruling elite. A profound economic crisis, a deterioration of social institutions, wars with grave economic, social and humanitarian consequences, the impoverishment of a large portion of the population, the expansion of the informal economy and the hampering of the development of civil society, were the main characteristics of Serbian society in this period."

Since then, however, Serbia has witnessed substantial growth—about 6% per capita per year between 2000 and 2006—but at the same time still remains one of the poorest countries in Europe (OECD 2008: 15). Turning to the labor market, employment has

decreased about 2.5% per year over the same period, especially in large and medium sized enterprises and in peasant farming (OECD 2008: 17). The reason for the increased unemployment during a time of widespread growth is due to (as was the case for many transition economies besides Serbia) labor hoarding before the transition so that the increased unemployment became a consequence of the economic restructuring and privatization embedded with the transition (Arandarenko 2007).

One measure taken to try to combat the declining employment is improved legislation, where an important step was taken in 2004 with the creation of a new Business Register Agency. The aim of this Agency was to coordinate several administrative functions that previously required contact with different authorities. In addition to this, starting in 2006, this Agency also keeps records of "entrepreneurs," i.e., self-employed own-account workers with or without employees, and handles their enrollment in social insurance (OECD 2008:22). The overall aim of this is to stimulate the creation of individual companies in Serbia and thereby stimulate overall economic growth in the Serbian economy, thus leading also to increased employment.

While there, thus, have been improvements in Serbian legislation vis-à-vis an improved business environment and increased employment opportunities, many obstacles still remain. Indeed, it has been suggested that due to the specific nature of the Serbian tax and benefit system, the value of social security contributions that are associated with formal employment have to be extremely high to offset the opportunity costs of formal employment, particularly for low-wage earners (Koettl 2010: 9).[4] The reason for this is minimum social security contributions, as well as the design of social assistance and family benefits. Considering these as a package, informal workers at low wages would have to give up a considerable amount of their informal earnings were they to "formalize" instead, and it is unlikely that the value of social security entitlement (and other benefits like formal employment protection legislation) that they get in return for formalization will exceed these implicit costs. Notably, the same holds for the inactive part of the labor market, when considering formal work at low wage levels. Koettl (2010: 9) goes on to conclude that, "In other words, so called-mini jobs and midi-jobs—that is, part-time jobs that pay less than the full-time minimum wage—are hardly economically viable in Serbia. Hence, workers with low educational attainment—like the informally employed and the inactive—might by and large be excluded from formal work in Serbia." Additionally, if working, it would likely be for lower wages, just as would be the case for informal sector workers—which is in line with the theories of segmented labor markets (Lewis 1954; Kuznets 1955). Indeed, according to Fields (2009), *the* distinguishing feature of this view is "the fact that workers earn different wages depending on the sector of the economy in which they are able to find work."

The above discussion, on the other hand, provides at least part of the explanation of the generally found "stylized fact" that informal sector workers tend to be less educated and lower earning (or vice versa) than formal sector workers—specifically for the case of Serbia: namely, the role of legislation, and here specifically the design of social security, social assistance and family benefits.

3 Data and descriptive analysis

The empirical analyses of this paper examine household survey data from two rounds of the Serbian Labor Force Survey (LFS), October 2008 (i.e. pre-crisis) and October

2009 (i.e., post-crisis, by about 1 year). The Serbia LFS is a two-stage, stratified survey that is representative at the national level. In the first stage, enumeration areas were selected systematically with probability proportional to the size of the population aged 15 and above (the target population) using the sampling frame of the 2002 Census. In the second stage households were selected within the enumeration areas with equal probability (simple random selection). The initial weight arising from the initial sampling design was further corrected ex-post to allow for non-response, aiming at creating sampling weights that make the sample nationally representative (these weights are used in all subsequent estimations). The survey contains information on labor market status, earnings, occupation, sector, industry, firm size, benefits and other labor related information, and on background variables such as age, gender, educational attainment, and area of residence, which are also important factors[5] in analyses of earnings determinants. A list of all the variables used in these analyses as well as their definitions is given in Table 1. The definition of variables is discussed in more detail in the remainder of this section.

Starting with the informality measures, the survey distinguishes between the different types of ownership—specifically, among private firms, a distinction is made between registered and non-registered firms, where the registration pertains to taxes and other payments and regulations. The first dimension therefore is based on a dummy variable, which is one if a worker works in a private, non-registered firm and zero otherwise. From Table 2 presenting the means and standard deviations of the informality measures, the incidence here is quite small, however—at only about 2–3%, depending on the time period. Fortunately, it is possible to define two additional, alternative informality measures. The second measure is based on a worker's contract status and is defined as one if a worker has a contract and zero otherwise. Benefit receipt is the third dimension of informality explored here. I again construct a binary measure; this time it is defined as one if a worker does not receive pension benefits and zero otherwise.[6] Altogether, the measure based on contract status is the preferred measure, with the pension benefits based measure a fairly close second, followed by the (il)legality measure.[7] In terms of the dynamics of the three informality measures, Table 2 reveals that the informal share of all dependent workers has decreased overall. This is especially true for the pension-based measure, which is likely due to the impact of the crisis putting more pressures on social protection demands, including pensions.

To help better understand these measures, especially their potential interrelatedness, Table 3 presents the correlations among the three informality measures. From these results the three measures appear highly correlated, though the two preferred measures related to contract status and benefits receipts are particularly highly correlated—supporting the desirability of these two measures.

While firm size has been used as a measure of informality in the literature, as also argued in Lehmann and Pignatti (2007); Lehmann and Zaiceva (2015) this measure may not be a very precise measure of informality, especially for transition countries. To examine this issue further, I tabulate firm size across the three informality measures (Table 4). From these results the overlap is quite small, never exceeding about a fifth—and frequently much less—so that most of the employees in small firms actually are formal employees. In turn, this supports the findings and recommendations in Lehmann and Pignatti (2007) for the case of Ukraine—and I therefore limit the subsequent analyses to the three informality measures discussed and defined earlier.

Table 1 Variable Definitions

Variable name:	*Definition:*
Dependent variable:	
Log real earnings (Oct. 2008, Oct. 2009)	Log real monthly earnings (net salary in the previous month)
Informality measures:	
Unregistered firm	1 if firm is not formally registered (for tax purposes, etc.); 0 otherwise
No labor contract	1 if worker has no labor contract; 0 otherwise
No pension benefits	1 if worker does not receive pension benefits; 0 otherwise
Explanatory variables:	
Birth cohorts:	
15–24 (reference)	1 if in age range; 0 otherwise
25–34	1 if in age range; 0 otherwise
35–44	1 if in age range; 0 otherwise
45–54	1 if in age range; 0 otherwise
55–64	1 if in age range; 0 otherwise
65 and above	1 if in age range; 0 otherwise
Educational attainment:	
Primary or less (reference)	1 if completed primary or less; 0 otherwise
Secondary	1 if completed secondary; 0 otherwise
Tertiary	1 if completed tertiary; 0 otherwise
Part-time status:	
Part-time	1 if part-time; 0 otherwise
Industry:	
Agriculture (reference)	1 if Agriculture; 0 otherwise
Man/Min/Electricity	1 if Manufacturing-Mining-Electricity; 0 otherwise
Construction	1 if Construction; 0 otherwise
Trade/Services	1 if Trade/Services; 0 otherwise
Hotels/Restaurants	1 if Hotels/Restaurants; 0 otherwise
Transports	1 if Transports; 0 otherwise
Finance/Real estate	1 if Finance/Real Estate; 0 otherwise
Public sector	1 if Public Sector; 0 otherwise
Other	1 if Other Sector; 0 otherwise
Occupation:	
Legislators	1 if Legislator; 0 otherwise
Professionals	1 if Professional; 0 otherwise
Technicians	1 if Technician; 0 otherwise
Clerks	1 if Clerk; 0 otherwise
Service	1 if Service; 0 otherwise
Skilled agriculture & fishery	1 if Skilled agriculture & fishery; 0 otherwise
Craft & trade	1 if Craft & trade; 0 otherwise
Plant/machine operators	1 if Plant/machine operator; 0 otherwise
Elementary occupations (reference)	1 if Elementary occupation; 0 otherwise
Firm Size:	
Firm size 1–5 (reference)	1 if firm size is 1–5; 0 otherwise
Firm size 6–19	1 if firm size is 6–19; 0 otherwise
Firm size 20-99	1 if firm size is 20–99; 0 otherwise

Table 1 Variable Definitions *(Continued)*

Firm size 100+	1 if firm size is 100 or more; 0 otherwise
Firm size not sure: 10 or less	1 if not sure about firm size but it is 10 or less; 0 otherwise
Firm size not sure: 11 or more	1 if not sure about firm size but it is 11 or more; 0 otherwise
Geographical location:	
Urban	1 if urban; 0 if rural
Central Serbia	1 if Central Serbia; 0 otherwise
Belgrade (reference)	1 if Belgrade; 0 otherwise
Vojvodina	1 if Vojvodina; 0 otherwise

The dependent variable is (log) total earnings (net salary) in the previous month (i.e., the month before October 2008 and October 2009, respectively). Unfortunately, this information is only collected from employees, so that the self-employed and owner-operators which potentially is an important part—and certainly is a large part of the overall Serbian informal sector, as can also be seen from Table 5—must be excluded from the analysis. As a result, the analysis in this paper examines a specific part of the overall Serbian informal sector, namely the part that contains employed workers who obtain a salary. Additionally, in many transition economies even workers in the formal sector receive part of their earnings in an informal fashion, in terms of so-called "envelope payments" (Lehmann (2010). While the survey question pertains to net salary, it is possible that envelope payments may still not be reported fully in the survey in which case the formal-informal earnings gap will be underestimated. Relatedly, non-cash components of formal sector earnings—like social security benefits, employment protection, severance payments, job security, credit worthiness, access to loans, and so on—will also work to widen the earnings gap even further. Unfortunately, however, there is not information on these components in the LFS.

While one might argue that (hourly) wages are preferable to (monthly) earnings, there are at least two reasons why earnings may actually be preferable to wages, even when hours worked is available in the dataset (which is the case here). First, for policy purposes the worker's take-home earnings seems to be the main object of interest since that is what he or she ultimately will use to sustain the livelihood of his or her household. Second, earnings is an already noisy variable, riddled with measurement error, so that dividing earnings with hours worked will attenuate the overall measurement error

Table 2 Incidence of male informal dependent employment in serbia over the first year of the financial crisis using three alternative informality measures: october 2008 and october 2009

	October 2008	*October 2009*
(1) Not formally registered	0.033	0.025
	[0.178]	[0.155]
(2) No labor contract	0.098	0.061
	[0.297]	[0.240]
(3) No pension benefits	0.111	0.064
	[0.314]	[0.245]
N	2,783	2,577

Notes: Estimations incorporate sampling weights. Values in brackets are robust Huber-White (Huber 1967; White 1980) standard errors
Source: Serbia Labor Force Survey (October 2008 and October 2009)

Table 3 Examining the intra-correlation of the three informality measures: October 2008 and October 2009

	October 2008			*October 2009*		
	(1) Not formally registered	*(2) No labor contract*	*(3) No pension benefits*	*(1) Not formally registered*	*(2) No labor contract*	*(3) No pension benefits*
(1) Not formally registered	1			1		
(2) No labor contract	0.522	1		0.478	1	
(3) No pension benefits	0.486	0.912	1	0.388	0.767	1
N	2,783	2,783	2,783	2,577	2,577	2,577

Source: Serbia Labor Force Survey (October 2008 and October 2009)

of the dependent variable—thus potentially leading to biased results. Despite these potential problems with the wage rate measure, a sensitivity analysis will still be performed to check for the robustness of the choice of dependent variable, in terms of (monthly) earnings versus (hourly) wages.

Among the key explanatory variables is age, which helps control for potential general experience, among other things. To obtain as flexible functional form as possible, a series of birth cohorts are created: 15–24, 25–34, 35–44, 45–54, 55–64, and 65 and above.[8] Educational attainment is measured as the highest level completed, ranging from "Without education" through "PhD." I consider a set of three binary variables corresponding to the completion of primary or less (reference), secondary, and tertiary education. Among the work related variables, I first construct a dummy variable for part-time work. Notably, from Tables 6 and 7, part-time status is particularly prevalent for the informal sector, with the incidence of part-time work among informal sector workers being more than 10 times more than that of formal sector workers. Industry and occupation clearly are potentially important determinants of earnings, as well, and are each included as a series of nine dummy variables (reference group for industry: agriculture; reference group for occupation: elementary occupations). From Tables 6 and 7, the informal sector is dominated by Agriculture and Construction, while Manufacturing/Mining/Electricity plays a major role in the formal sector (accounting for about a third of the dependent employment in the sector) but a much smaller role in the informal sector. In terms of occupations, the informal sector is dominated by elementary occupations (including unskilled agriculture and fishery), with between about a third and half of the dependent workers belonging to this group depending on the informality measures. These results indicate once again how the informal sector in Serbia may well be considered "the marginal sector." Firm size is also a potentially important determinant of earnings and is included as a set of six dummies, with 1–5 employees as the reference category. Lastly, the urban dummy and the region of residence cluster fixed effects capture economic conditions specific to the area (as well as everything else related to the region in question), which are potentially important in explaining labor earnings.

Gender-issues may potentially confound the analysis[9], so to enable focusing more narrowly on the formal-informal sector earnings gap, the sample is initially restricted to the 2,978 males in October 2008 and 2,796 males in October 2009 who are employed and are 15 years of age and above. Additionally, information on some observations is missing for either the dependent variable or for one or more of the explanatory

Table 4 Examining overlap of firm size and the three informality measures: October 2008 and October 2009 (Percent)

	October 2008						*October 2009*					
	(1) Not formally registered		*(2) No labor contract*		*(3) No pension benefits*		*(1) Not formally registered*		*(2) No labor contract*		*(3) No pension benefits*	
	No: Formal	Yes: Informal	No: Formal	Yes: Informal	No: Formal	Yes: Informal	No: Formal	Yes: Informal	No: Formal	Yes: Informal	No: Formal	Yes: Informal
6+ employees	98.5	1.5	93.3	6.7	92.3	7.7	98.7	1.3	95.9	4.1	95.7	4.3
1–5 employees	91.3	**8.7**	80.7	**19.3**	78.2	**21.8**	93.8	**6.3**	87.4	**12.7**	86.7	**13.4**
N	2,783		2,783		2,783		2,577		2,577		2,577	

The rows (firm size by informality measure) do not sum to exactly 100% in a few cases due to rounding. Bold numbers: This is the grouping where the informality measure in question "agrees" with the firm size measure in terms of the informality classification–so the closer this number is to 100, the greater the agreement (and vice versa)
Source: Serbia Labor Force Survey (October 2008 and October 2009)

Table 5 Distribution of employees, self-employed and unpaid family workers in the larger sample of all working (but not necessarily enumerated) individuals in the Informal sector using three alternative informality measures: october 2008 and october 2009

	October 2008			*October 2009*		
	(1) Not formally registered	*(2) No labor contract*	*(3) No pension benefits*	*(1) Not formally registered*	*(2) No labor contract*	*(3) No pension benefits*
Employee	14.7	18.0	25.3	17.0	16.1	24.1
Self-employed	77.5	70.6	61.5	71.7	70.0	61.2
Unpaid family worker	7.9	11.4	13.2	11.3	13.9	14.7
N	672	1,644	1,344	606	1,448	973

Notes: Estimations incorporate sampling weights
Source: Serbia Labor Force Survey (October 2008 and October 2009)

variables, leading to final estimation samples of 2,783 observations for the October 2008 estimation sample and 2,577 observations for the October 2009 estimation sample. Descriptive statistics for the analysis samples across informality status are reported in Tables 2, 5, 6 and 7.

To get an initial handle on the formal-informal earnings gap during the first year-and-a-bit of the international financial crisis, Table 8 presents log earnings for the formal and informal sector for the October 2008 sample and the October 2009 sample, using all three alternative informality definitions discussed previously. To obtain results in percent rather than log-points, the earnings gaps are also presented in their de-logged form.[10]

A few results from the table are rather striking. First, the formal-informal sector earnings gap is pervasive, no matter the definition or time period considered.[11] In turn, this is also in line with the predictions from the literature on segmented labor markets that the lower productivity—or "marginal"—sector also receives the lower wages (Lewis 1954; Kuznets 1955). Second, however, there is a bit of a range in the estimated raw earnings gap—ranging from about 47 percentage-points to about 60 percentage-points in October 2008 and ranging from about 20 to about 27 percentage-points in October 2009, depending on the informality measure considered. In turn, these results confirm earlier findings (Krstić and Sanfey 2011) of a substantial formal-informal earnings gap in Serbia in 2007—though, at about 43 percentage-points, this gap was somewhat smaller than the gaps estimated here for the earlier period and somewhat larger than the gaps estimated here for the later period (likely due to differences in both the data and informality measure used).[12] Third, considering the three informality measures overall, the earnings gap narrowed substantially following the crisis for all three measures. Examining the evidence as a whole, however—incorporating the size of the changes, as well as the relative desirability of the measures as "true" informality measures—it is evident that the overall formal-informal sector earnings gap decreased substantially following the crisis. Again, since "envelope payments" likely have fallen more than reported payments, the decrease in the earnings gap is potentially overestimated.

So what might account for these differences in earnings between the two sectors more generally—and for the narrowing of the gap following the crisis? The previous section discussed how Serbian legislation related to minimum social security

Table 6 Means and standard deviations of monthly earnings and explanatory variables by formality status using three alternative informality measures: october 2008

	(1) Formally registered?		*(2) Has labor contract?*		*(3) Receives pension benefits?*	
	Yes	No	Yes	No	Yes	No
Ln Monthly earnings	10.067	9.204	10.102	9.456	10.108	9.482
	[0.528]	[0.823]	[0.496]	[0.762]	[0.493]	[0.736]
Informality measure:						
Unregistered firm	0.000	1.000	0.002	0.311	0.002	0.273
	[0.000]	[0.000]	[0.046]	[0.464]	[0.046]	[0.446]
No labor contract	0.069	0.94	0.000	1.000	0.002	0.863
	[0.253]	[0.239]	[0.000]	[0.000]	[0.045]	[0.345]
No benefits	0.082	0.94	0.017	0.981	0.000	1.000
	[0.275]	[0.239]	[0.128]	[0.135]	[0.000]	[0.000]
Age cohort:						
15–24	0.077	0.254	0.067	0.224	0.066	0.215
	[0.266]	[0.438]	[0.250]	[0.418]	[0.248]	[0.411]
25–34	0.242	0.233	0.239	0.27	0.239	0.268
	[0.429]	[0.425]	[0.427]	[0.445]	[0.427]	[0.444]
35–44	0.268	0.174	0.268	0.232	0.267	0.245
	[0.443]	[0.381]	[0.443]	[0.423]	[0.443]	[0.431]
45–54	0.274	0.217	0.285	0.154	0.287	0.151
	[0.446]	[0.414]	[0.451]	[0.361]	[0.452]	[0.359]
55–64	0.137	0.107	0.139	0.11	0.139	0.112
	[0.344]	[0.311]	[0.346]	[0.313]	[0.346]	[0.316]
65+	0.002	0.015	0.002	0.01	0.002	0.009
	[0.048]	[0.123]	[0.044]	[0.100]	[0.044]	[0.094]
Educational attainment:						
Primary or less	0.143	0.556	0.123	0.461	0.124	0.415
	[0.350]	[0.500]	[0.329]	[0.499]	[0.330]	[0.494]
Secondary	0.668	0.429	0.678	0.503	0.678	0.523
	[0.471]	[0.498]	[0.467]	[0.501]	[0.467]	[0.500]
Tertiary	0.189	0.015	0.199	0.037	0.198	0.062
	[0.391]	[0.123]	[0.399]	[0.188]	[0.399]	[0.241]
Part-time	0.02	0.313	0.013	0.182	0.011	0.183
	[0.141]	[0.466]	[0.115]	[0.387]	[0.104]	[0.387]
Industry:						
Agriculture	0.045	0.402	0.035	0.252	0.035	0.227
	[0.206]	[0.493]	[0.184]	[0.435]	[0.183]	[0.420]
Man/Min/Electricity	0.351	0.091	0.364	0.144	0.364	0.176
	[0.477]	[0.290]	[0.481]	[0.351]	[0.481]	[0.381]
Construction	0.103	0.409	0.088	0.341	0.09	0.293
	[0.304]	[0.495]	[0.284]	[0.475]	[0.287]	[0.456]
Trade/Services	0.134	0.047	0.131	0.131	0.13	0.139
	[0.340]	[0.212]	[0.337]	[0.338]	[0.336]	[0.347]
Hotels/Restaurants	0.03	0	0.025	0.06	0.025	0.058
	[0.170]	[0.000]	[0.158]	[0.239]	[0.157]	[0.234]

Table 6 Means and standard deviations of monthly earnings and explanatory variables by formality status using three alternative informality measures: october 2008 *(Continued)*

Transports	0.09	0.037	0.095	0.022	0.097	0.019
	[0.286]	[0.191]	[0.294]	[0.147]	[0.296]	[0.138]
Finance/Real estate	0.047	0.000	0.049	0.014	0.048	0.026
	[0.212]	[0.000]	[0.216]	[0.119]	[0.214]	[0.161]
Public sector	0.148	0.000	0.158	0.000	0.159	0.011
	[0.355]	[0.000]	[0.365]	[0.000]	[0.366]	[0.106]
Other	0.053	0.013	0.054	0.036	0.052	0.049
	[0.224]	[0.113]	[0.225]	[0.186]	[0.222]	[0.217]
Occupation:						
Legislators	0.043	0.000	0.046	0.000	0.047	0.000
	[0.203]	[0.000]	[0.210]	[0.000]	[0.211]	[0.000]
Professionals	0.106	0.000	0.113	0.006	0.113	0.02
	[0.309]	[0.000]	[0.317]	[0.076]	[0.317]	[0.140]
Technicians	0.142	0.013	0.149	0.033	0.148	0.06
	[0.349]	[0.113]	[0.357]	[0.179]	[0.355]	[0.238]
Clerks	0.069	0.053	0.073	0.029	0.073	0.037
	[0.254]	[0.225]	[0.261]	[0.169]	[0.260]	[0.188]
Service	0.13	0.000	0.127	0.119	0.126	0.125
	[0.336]	[0.000]	[0.333]	[0.325]	[0.332]	[0.331]
Skilled agriculture & fishery	0.007	0.09	0.004	0.061	0.004	0.053
	[0.082]	[0.288]	[0.062]	[0.239]	[0.063]	[0.225]
Craft & trade	0.252	0.325	0.249	0.3	0.25	0.289
	[0.434]	[0.471]	[0.433]	[0.459]	[0.433]	[0.454]
Plant/machine operators	0.156	0.025	0.16	0.081	0.161	0.076
	[0.363]	[0.157]	[0.366]	[0.273]	[0.368]	[0.266]
Elementary occupations	0.094	0.494	0.078	0.371	0.078	0.339
	[0.292]	[0.503]	[0.269]	[0.484]	[0.268]	[0.474]
Firm Size:						
Firm size 1–5	0.23	0.636	0.218	0.479	0.214	0.476
	[0.421]	[0.484]	[0.413]	[0.500]	[0.410]	[0.500]
Firm size 6–19	0.294	0.237	0.29	0.31	0.29	0.313
	[0.456]	[0.427]	[0.454]	[0.463]	[0.454]	[0.464]
Firm size 20–99	0.241	0.061	0.251	0.09	0.254	0.084
	[0.428]	[0.240]	[0.434]	[0.286]	[0.435]	[0.278]
Firm size 100+	0.17	0	0.181	0.012	0.181	0.032
	[0.376]	[0.000]	[0.385]	[0.111]	[0.385]	[0.175]
Firm size not sure: 10 or less	0.018	0.022	0.013	0.065	0.013	0.057
	[0.131]	[0.148]	[0.112]	[0.248]	[0.112]	[0.233]
Firm size not sure: 11 or more	0.047	0.044	0.048	0.043	0.048	0.038
	[0.212]	[0.207]	[0.213]	[0.204]	[0.214]	[0.191]
Geographical location:						
Urban	0.624	0.31	0.634	0.423	0.633	0.458
	[0.484]	[0.465]	[0.482]	[0.495]	[0.482]	[0.499]
Central Serbia	0.502	0.295	0.504	0.421	0.507	0.406

Table 6 Means and standard deviations of monthly earnings and explanatory variables by formality status using three alternative informality measures: october 2008 *(Continued)*

	[0.500]	[0.459]	[0.500]	[0.495]	[0.500]	[0.492]
Belgrade	0.217	0.085	0.223	0.123	0.221	0.153
	[0.413]	[0.281]	[0.416]	[0.329]	[0.415]	[0.360]
Vojvodina	0.28	0.62	0.274	0.455	0.273	0.441
	[0.449]	[0.488]	[0.446]	[0.499]	[0.445]	[0.497]
N	2,691	89	2,507	273	2,474	306

Notes: Estimations incorporate sampling weights. Values in brackets are robust Huber-White (Huber 1967; White 1980) standard errors
Source: Serbia Labor Force Survey (October 2008)

contributions, social assistance, and family benefits may help explain the generally found "stylized fact" that informal sector workers tend to be less educated and lower earning (or vice versa) than formal sector workers. From Tables 6 and 7, it can be seen that informal sector workers are indeed (much) worse off than formal sector workers in terms of human capital. For example, in October 2008 only 14.3% of workers in formally registered firms had completed primary education or less, while more than half of workers in firms that were not formally registered had completed primary or less (Table 6). On the other hand, almost 19% of workers in formally registered firms had completed tertiary education, while this was the case for only 1.5% of workers in firms that were not formally registered. From the bottom panel of Table 8, inclusion of these factors as explanatory variables make the earnings gap shrink, though they remain sizeable—hinting that worker characteristics are important but at the same time cannot explain away the earnings gap (this will be discussed further in the following sections).

Comparing Tables 6 and 7, however, it seems that the *composition* of the informal sector has changed following the crisis: for example, while, again, more than half of workers in firms that were not formally registered had completed primary or less, this had decreased to about 37% in October 2009. Similarly, the share of workers with completed secondary education increased from about 43% to about 56% and the share of workers with tertiary education from 1.5% to 6.6%. At the face of it, this is consistent with formal workers (having better characteristics than informal sector workers—in terms of education, for example) being pushed into informality due to the crisis. What is more likely going on, however, is that the change in the composition of the informally employed is due to an outflow from informality into unemployment, since the informal sector workers were likely the first workers to be fired—being a more flexible segment of the labor market, not regulated by employment protection legislation. Additionally, formal workers are more likely to receive unemployment benefits and would therefore probably rather flow into unemployment than into informality. It seems, therefore, that the outflow from informality into unemployment is biased, with the least endowed leaving over-proportionally into unemployment, therefore overall seemingly "improving" the endowments of the informally employed.

While the existence of substantively large formal-informal sector earnings gaps has now been established across all three informality measures and for both the pre- and post crisis period—and with sectoral human capital differences motivated as possibly accounting for at least some of this gap as well as the narrowing of the gap following

Table 7 Means and standard deviations of monthly earnings and explanatory variables by formality status using three alternative informality measures: october 2009

	(1) Formally registered?		*(2) Has labor contract?*		*(3) Receives pension benefits?*	
	Yes	No	Yes	No	Yes	No
Ln Monthly earnings	10.079	9.437	10.096	9.572	10.093	9.637
	[0.501]	[0.693]	[0.493]	[0.605]	[0.501]	[0.552]
Informality measure:						
Unregistered firm	0.000	1.000	0.006	0.315	0.009	0.254
	[0.000]	[0.000]	[0.076]	[0.466]	[0.094]	[0.437]
No labor contract	0.043	0.781	0.000	1	0.013	0.762
	[0.203]	[0.417]	[0.000]	[0.000]	[0.113]	[0.427]
No benefits	0.049	0.663	0.016	0.802	0.000	1.000
	[0.216]	[0.476]	[0.127]	[0.399]	[0.000]	[0.000]
Age cohort:						
15–24	0.064	0.112	0.058	0.179	0.056	0.202
	[0.245]	[0.317]	[0.233]	[0.385]	[0.229]	[0.403]
25–34	0.235	0.134	0.229	0.302	0.23	0.275
	[0.424]	[0.343]	[0.420]	[0.460]	[0.421]	[0.448]
35–44	0.266	0.318	0.27	0.235	0.271	0.216
	[0.442]	[0.469]	[0.444]	[0.425]	[0.445]	[0.413]
45–54	0.28	0.308	0.288	0.18	0.289	0.171
	[0.449]	[0.465]	[0.453]	[0.386]	[0.453]	[0.377]
55–64	0.15	0.112	0.153	0.092	0.151	0.12
	[0.357]	[0.317]	[0.360]	[0.291]	[0.358]	[0.326]
65+	0.004	0.017	0.004	0.011	0.003	0.017
	[0.063]	[0.129]	[0.062]	[0.105]	[0.058]	[0.129]
Educational attainment:						
Primary or less	0.125	0.371	0.113	0.399	0.114	0.371
	[0.330]	[0.487]	[0.317]	[0.491]	[0.318]	[0.484]
Secondary	0.668	0.563	0.671	0.583	0.671	0.584
	[0.471]	[0.500]	[0.470]	[0.495]	[0.470]	[0.494]
Tertiary	0.207	0.066	0.216	0.018	0.215	0.045
	[0.405]	[0.250]	[0.411]	[0.132]	[0.411]	[0.209]
Part-time	0.013	0.297	0.01	0.183	0.01	0.167
	[0.114]	[0.460]	[0.097]	[0.388]	[0.100]	[0.374]
Industry:						
Agriculture	0.036	0.313	0.033	0.194	0.032	0.203
	[0.186]	[0.467]	[0.179]	[0.396]	[0.175]	[0.403]
Man/Min/Electricity	0.346	0.163	0.354	0.137	0.35	0.206
	[0.476]	[0.372]	[0.478]	[0.345]	[0.477]	[0.405]
Construction	0.08	0.315	0.072	0.301	0.075	0.251
	[0.272]	[0.468]	[0.258]	[0.460]	[0.263]	[0.435]
Trade/Services	0.13	0.048	0.129	0.121	0.129	0.118
	[0.337]	[0.216]	[0.335]	[0.327]	[0.335]	[0.323]
Hotels/Restaurants	0.028	0	0.026	0.054	0.026	0.047
	[0.165]	[0.000]	[0.158]	[0.226]	[0.159]	[0.211]

Table 7 Means and standard deviations of monthly earnings and explanatory variables by formality status using three alternative informality measures: october 2009 *(Continued)*

Transports	0.109	0.015	0.111	0.043	0.112	0.035
	[0.312]	[0.123]	[0.314]	[0.203]	[0.315]	[0.185]
Finance/Real estate	0.049	0.043	0.051	0.025	0.049	0.047
	[0.217]	[0.204]	[0.220]	[0.157]	[0.217]	[0.213]
Public sector	0.168	0	0.174	0.004	0.174	0.015
	[0.374]	[0.000]	[0.379]	[0.066]	[0.379]	[0.124]
Other	0.054	0.103	0.051	0.122	0.053	0.078
	[0.225]	[0.306]	[0.219]	[0.328]	[0.225]	[0.269]
Occupation:						
Legislators	0.044	0.000	0.046	0.000	0.045	0.006
	[0.205]	[0.000]	[0.209]	[0.000]	[0.208]	[0.077]
Professionals	0.115	0.057	0.121	0.000	0.12	0.017
	[0.319]	[0.233]	[0.326]	[0.000]	[0.326]	[0.129]
Technicians	0.145	0.067	0.146	0.088	0.148	0.067
	[0.352]	[0.251]	[0.353]	[0.285]	[0.355]	[0.251]
Clerks	0.073	0.021	0.073	0.061	0.074	0.044
	[0.261]	[0.145]	[0.260]	[0.240]	[0.262]	[0.205]
Service	0.131	0.012	0.129	0.115	0.128	0.126
	[0.337]	[0.109]	[0.335]	[0.320]	[0.334]	[0.333]
Skilled agriculture & fishery	0.008	0.12	0.007	0.067	0.006	0.077
	[0.090]	[0.328]	[0.084]	[0.251]	[0.079]	[0.268]
Craft & trade	0.241	0.361	0.24	0.307	0.238	0.33
	[0.428]	[0.484]	[0.427]	[0.463]	[0.426]	[0.471]
Plant/machine operators	0.159	0.045	0.162	0.065	0.162	0.065
	[0.366]	[0.210]	[0.369]	[0.247]	[0.369]	[0.247]
Elementary occupations	0.084	0.317	0.076	0.296	0.077	0.269
	[0.277]	[0.469]	[0.266]	[0.458]	[0.267]	[0.445]
Firm Size:						
Firm size 1–5	0.223	0.589	0.216	0.481	0.215	0.482
	[0.417]	[0.496]	[0.412]	[0.501]	[0.411]	[0.501]
Firm size 6–19	0.282	0.229	0.28	0.3	0.283	0.249
	[0.450]	[0.424]	[0.449]	[0.459]	[0.451]	[0.434]
Firm size 20–99	0.245	0.038	0.25	0.078	0.25	0.092
	[0.430]	[0.192]	[0.433]	[0.269]	[0.433]	[0.289]
Firm size 100+	0.177	0.01	0.183	0.02	0.181	0.06
	[0.382]	[0.102]	[0.387]	[0.141]	[0.385]	[0.238]
Firm size not sure: 10 or less	0.018	0.104	0.016	0.09	0.017	0.069
	[0.134]	[0.307]	[0.125]	[0.287]	[0.129]	[0.255]
Firm size not sure: 11 or more	0.054	0.03	0.055	0.031	0.054	0.048
	[0.227]	[0.173]	[0.228]	[0.175]	[0.226]	[0.215]
Geographical location:						
Urban	0.624	0.411	0.629	0.457	0.629	0.462
	[0.485]	[0.496]	[0.483]	[0.500]	[0.483]	[0.500]
Central Serbia	0.49	0.504	0.492	0.475	0.495	0.433

Table 7 Means and standard deviations of monthly earnings and explanatory variables by formality status using three alternative informality measures: october 2009 *(Continued)*

	[0.500]	[0.504]	[0.500]	[0.501]	[0.500]	[0.497]
Belgrade	0.251	0.179	0.256	0.149	0.256	0.156
	[0.434]	[0.386]	[0.436]	[0.357]	[0.436]	[0.364]
Vojvodina	0.259	0.317	0.252	0.376	0.25	0.411
	[0.438]	[0.469]	[0.435]	[0.486]	[0.433]	[0.494]
N	2,510	67	2,411	166	2,403	174

Notes: Estimations incorporate sampling weights. Values in brackets are robust Huber-White (Huber 1967; White 1980) standard errors
Source: Serbia Labor Force Survey (October 2009)

the onset of the crisis—the objective of the main analysis of this paper is to now try to "explain"[13] these gaps in more detail in terms of, on the one hand, characteristics/endowments such as educational attainment and job characteristics and returns to these characteristics (three-fold division) and, on the other hand, observable and unobservable characteristics (two-fold division). While the empirical strategy underlying this approach is widely used, it still seems fruitful to review the main components in some detail, especially in terms of how it is tailored to the application pursued here—which, therefore, is the objective of the next section.

4 Estimation strategy and related issues

The starting point of the Blinder-Oaxaca approach to decompose earnings (or other) differentials is an OLS regression of the outcome in question, estimated separately across the two relevant groups (Blinder 1973; Oaxaca 1973); here, workers from the formal and the informal sector, respectively (suppressing subscripts for individual workers):

Table 8 Raw and regression-adjusted formal-informal sector earnings gap (Informal Sector Penalty) using three alternative informality measures: October 2008 and October 2009 (in logs and de-logged)

	October 2008			*October 2009*		
	(1) Not formally registered	*(2) No labor contract*	*(3) No pension benefits*	*(1) Not formally registered*	*(2) No labor contract*	*(3) No pension benefits*
(i) Raw Earnings Gaps:						
Formal sector	10.067***	10.102***	10.108***	10.079***	10.096***	10.093***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Informal sector	9.152***	9.437***	9.465***	9.437***	9.572***	9.637***
	[0.005]	[0.003]	[0.002]	[0.005]	[0.003]	[0.002]
IS penalty (Logs)	−0.914***	−0.665***	−0.643***	−0.642***	−0.524***	−0.456***
	[0.005]	[0.003]	[0.002]	[0.005]	[0.003]	[0.002]
IS penalty (De-logged)	−0.599	−0.486	−0.474	−0.269	−0.195	−0.211
(ii) Regression-adjusted Earnings Gaps:						
IS penalty (Logs)	−0.314***	−0.217***	−0.237***	−0.307***	−0.187***	−0.140***
	[0.004]	[0.002]	[0.002]	[0.004]	[0.002]	[0.002]
IS penalty (De-logged)	−0.474	−0.408	−0.366	−0.264	−0.171	−0.131
N	2,783	2,783	2,783	2,577	2,577	2,577

Values in brackets are robust Huber-White (Huber 1967; White 1980) standard errors
***: statistically significant at 1%. *Source:* Serbia Labor Force Survey (October 2008 and October 2009)

$$Y_{FS} = \beta_{FS} X + \varepsilon_{FS} \quad (1)$$

$$Y_{IS} = \beta_{IS} X + \varepsilon_{IS}, \quad (2)$$

where Y_{FS} and Y_{IS} are the logarithms of monthly earnings of informal and formal sector workers, respectively, X is a vector of workers' characteristics (education, experience, occupation, and so on); β_{FS} and β_{IS} are the returns to the workers' characteristics; and ε_{FS} and ε_{IS} are error terms.

As such, these regressions are—at least in this context—merely inputs into calculating the decompositions. However, it is potentially fruitful to consider these auxiliary regressions in and of themselves as separate and integral parts of the overall analysis, also. Both because the results from these regressions directly indicate the different returns to characteristics across informality status but also because their specification, most notably in terms of explanatory variables, will affect the subsequent decomposition results.

Human capital theory suggests that education and potential experience directly affect earnings through the impact on individuals' productivity in the labor market and also suggest additional factors that are potentially important determinants of earnings such as education, industry and sector of employment, firm size, part-time status, and location of residence.

Hence, the first part of the multivariate analysis will examine these relationships using ordinary least squares. This is done by first including only the informality measure (thus recovering the raw earnings gaps from Table 8), then adding all the controls, and finally adding a full set of interactions with the informality measure. One potentially important econometric issue here is that educational attainment may be endogenous. The main concern here is possible omitted variables bias. Preferences and ability, for example, are unobserved and at the same time also, at least to some extent, determine both educational attainment and labor market earnings. However, as there are not available in this dataset any variables that may potentially act as instruments, it does not appear feasible to try to address this problem using instrumental variables methods. The effect of any omitted variables will therefore be captured by the error term, possibly causing omitted variables bias. The same goes for part-time status, industry, and occupation, which nevertheless are included as explanatory variables due to their potential importance for informality status.[14] As a result, we must interpret any subsequent results with caution and hence not give them a causal interpretation but rather as merely reflecting associations with labor market earnings. Further, so as to allow for arbitrary heteroskedasticity, the estimations will be carried out using Huber-White standard errors (Huber 1967; White 1980).

Again, these earnings regressions formally are merely inputs into the decomposition analysis. Specifically, the decomposition analysis amounts to examining to what extent the observed earnings gaps across informality status are attributable to differences in the observable characteristics, to differences in the returns to those characteristics, and to the interaction of the two ("three-fold decomposition," see below for details) and, relatedly, to what extent the observed earnings gaps are due to observable and unobservable characteristics ("two-fold decomposition," see below for details). This analysis will comprise the second part of the

multivariate empirical analysis and will be pursued as an Oaxaca-Blinder type decomposition.

Formally, following the methodology of Oaxaca (1973); Blinder (1973), the difference in mean earnings for formal and informal sector workers, denoted R, can be decomposed into three parts (Jann 2008) using the empirical counterparts of equations (1) and (2) above[15]:

$$R = \bar{Y}_{FS} - \bar{Y}_{IS} = (\bar{X}_{FS} - \bar{X}_{IS})\hat{\beta}_{FS} + \bar{X}_{FS}\left(\hat{\beta}_{FS} - \hat{\beta}_{IS}\right) - (\bar{X}_{FS} - \bar{X}_{IS})\left(\hat{\beta}_{FS} - \hat{\beta}_{IS}\right) \quad (3)$$

This is a three-fold decomposition (Winsborough and Dickinson 1971), where the first term represents the "endowments effect" and explains the differences that are due to worker characteristics (such as education, sector of employment, occupation, etc). The second term reflects the "coefficients effect," which shows the differences in the estimated returns to formal and informal sector workers' characteristics. Lastly, the third term, the "interaction effect," accounts for the fact that differences in endowments and coefficients between formal and informal sector workers exist simultaneously. If formal and informal sector workers obtain equal returns for their characteristics, the second and the third part in equation (3) will equal zero and earnings differentials between formal and informal sector workers will be explained by the differences in endowments alone.

The above decomposition is formulated based on the prevailing earnings structure of formal sector workers, i.e., the differences in endowments and coefficients between formal and informal sector workers are weighted by the coefficients (returns) of formal sector workers. This seems reasonable for the application here since the formal sector is the dominating one, at least in an economic sense/size-wise. This is therefore also the approach pursued in the subsequent analysis. Alternatively, however, this equation could also be represented based on the prevailing earnings structure of informal sector workers; this will be explored further in the sensitivity analysis.

An alternative approach, prominent in the literature on wage discrimination, is based on the assumption that wage differentials are explained by a unifying ("non-discriminatory," in the wage discrimination literature) coefficients vector, denoted β^*, which is estimated in a regression that pools together both of the two groups under consideration (here, formal and informal sector workers). Then, the earnings gap can be expressed as:

$$R = \bar{Y}_{FS} - \bar{Y}_{IS} = (\bar{X}_{FS} - \bar{X}_{IS})\hat{\beta}^* + \bar{X}_{FS}\left(\hat{\beta}_{FS} - \hat{\beta}^*\right) - \bar{X}_{IS}\left(\hat{\beta}^* - \hat{\beta}_{IS}\right) \quad (4)$$

The above equation represents the so-called two-fold[16] decomposition:

$$R = Q + U, \quad (5)$$

where $Q = (\bar{X}_{FS} - \bar{X}_{IS})\hat{\beta}^*$ is the part of the earnings differential that is "explained" by sample differences assessed with common "returns" across the two groups, and the second term $U = \bar{X}_{FS}\left(\hat{\beta}_{FS} - \hat{\beta}^*\right) + \bar{X}_{IS}\left(\hat{\beta}^* - \hat{\beta}_{IS}\right)$ is the "unexplained" part not attributed to observed differences in formal and informal sector characteristics. The latter

part is often treated as discrimination in the literatures on gender and racial earnings gaps. It is important to note, however, that the "unexplained" part also captures all potential effects of differences in unobserved variables (Jann 2008). And, to be sure, in the application here it is difficult to talk about "discrimination," *per se*, as it does not seem that being an informal sector worker is an intrinsic characteristic, such as gender or ethnicity. Again choosing the formal sector earnings structure as the reference, (4) reduces to:

$$R = \bar{Y}_{\mathrm{FS}} - \bar{Y}_{\mathrm{IS}} = (\bar{X}_{\mathrm{FS}} - \bar{X}_{\mathrm{IS}})\hat{\beta}_{\mathrm{FS}} + \bar{X}_{\mathrm{IS}}\left(\hat{\beta}_{\mathrm{FS}} - \hat{\beta}_{\mathrm{IS}}\right) \quad (6)$$

Again, while the main analysis here takes the formal sector earnings structure as the reference, several different specifications for the baseline[17] specification, i.e., $\hat{\beta}*$ in (4), will be pursued in the sensitivity analysis as a robustness check.

The standard errors of the individual components in equations (3) and (4) above are computed using the Delta method by applying the procedure detailed in Jann (2008), which extends the earlier method developed in Oaxaca and Ransom (1998) to deal with stochastic regressors.

In addition to examining the overall composition of the established earnings gaps, it would seem instructive to perform detailed decompositions as well, whereby it would be possible to see which explanatory variables contribute the most to the three- and/or two-fold overall decompositions. An issue here is that while the overall decompositions are always identified, the results for categorical variables in detailed decompositions depend on the choice of the reference category (Oaxaca and Ransom 1999). A possible solution to this problem is to apply the deviation contrast transformation to the estimates before conducting the decomposition (Yun 2005); this is also the approach pursued here. Similar to the OLS regressions, the decomposition estimations also all allow for arbitrary heteroskedasticity (Huber 1967; White 1980). So as to condense the wealth of results obtained here—thereby easing the interpretation of the many results—the detailed decompositions are done groupwise rather than for each individual variable (for example, for education as a whole, rather than separately for primary or less, secondary, and tertiary education). Here, too, the focus will be on the case where the formal sector is taken as the reference sector, though the sensitivity analysis again will consider alternative specifications as well.

In addition to the sensitivity analysis using different specifications for the weight of the formal sector in the decomposition analysis and the potential endogeneity of part-time status, I perform several additional robustness checks. Two sensitivity analyses relates to the relevant estimation sample—on the one hand restricting the sample to only prime age (25–64 years of age) males, on the other adding females in the estimations. Despite the potential issues with using wage rates here, also discussed earlier, a sensitivity analysis with (hourly) wage rates as the dependent variable—instead of (monthly) earnings—is performed as well. Lastly, to help both examine more and justify better the negative selection into informal jobs referred to earlier, I estimate simple regressions of the determinants of informality status.

5 Results

This section reviews the main results from the formal-informal sector earnings analysis for the Serbian Labor Force Survey from October 2008 and October 2009, thereby

exploring whether the nature of the formal-informal sector earnings gap changed during the first year or so of the international financial crisis.[18] This is done in four main parts: (i) OLS Mincer earnings regressions, (ii) overall earnings decompositions (both two- and three-fold), (iii) detailed earnings decompositions (again both two- and three-fold), and (iv) sensitivity analysis.

5.1 Mincer earnings regressions

The Mincer earnings regressions are estimated for three different specifications for each of the three alternative informality measures: (i) only including the informality variable (thus recovering the raw formal-informal sector gaps from Table 3), (ii) adding controls, (iii) adding a full set of interactions with the informality measure.

For each survey, the results are remarkably consistent across the three informality measures. First, for all three measures the initial raw earnings gaps discussed previously decrease substantially when including controls (Table 8, bottom panel). For example, the gap decreases from 59.9 to 26.9% when using formal registration of the enterprise as the informality measure in October 2008.[19] This again suggests that at least some of the gap can be explained by observable characteristics (as will be further explored in the Oaxaca-Blinder decompositions shortly); the substantial increase in R^2 when moving from the specifications with only the informality measure to adding the controls also supports this. These findings again confirm the findings by Krstić and Sanfey (2011), where the raw formal-informal earnings gap decreases from about 43 percentage-points to about 22 percentage-points in 2007 once controls such as age, education, and industry are added.[20] These results again also confirm the predictions from the literature on segmented labor markets that the lower productivity, or "marginal," sector also receives the lower wages (Lewis 1954; Kuznets 1955)—though now even when controlling for the worse characteristics of these workers in terms of, for example, education.

Second, while the raw gap decreases substantially following the onset of the crisis, the gap adjusted for observable worker characteristics is remarkably stable. For example, while the raw gap decreases from 59.9% in October 2008 to 47.4% in October 2009, when using formal registration of the enterprise as the informality measure, the adjusted gap only decreases from 26.9 to 26.4%. In turn, this indicates that the economic disadvantage pertaining to informality in Serbia is much more constant when taking observable characteristics of formal and informal workers into account. Third, in line with previous research, there are substantial returns to education, all industries experience an earnings premium relative to the reference category of agriculture, and similarly all occupations experience an earnings premium relative to the reference category of elementary occupations, workers in urban areas experience (mostly) an earnings premium relative to workers from rural areas, and workers from regions outside that of the capital (Belgrade) experience an earnings penalty.

Lastly, while the many estimated coefficients make it hard to assess this precisely, the overall impression from the fully interacted model is that informal sector workers receive a penalty on many of their observable characteristics.

In essence, the decomposition analysis—to which I now turn—formalizes and condenses the overwhelming results from the Mincer regressions of the fully interacted models into more easily interpretable numbers.

5.2 Overall earnings decompositions

Table 9 presents the overall earnings decompositions for the three alternative informality measures across the pre- and post crisis surveys, with the three-fold decompositions in the upper half of the table and the two-fold decompositions in the lower half of the table.

Starting with the three-fold decomposition, a couple of results in Table 9 stand out particularly strongly. First, the endowments increase the formal-informal sector earnings gap overall, indicating that informal sector workers have relatively less favorable observable characteristics—that is, they are concentrated in worse paying sectors (including especially agriculture), have less education, and so on (this will be examined more closely when considering the detailed decompositions in the next sub-section). Second, the returns to characteristics increase the gaps in both substantive and statistical terms, indicating that formal sector workers have higher returns to characteristics overall.

Moving to the two-fold decompositions, informal sector workers, on average, have worse employment-related characteristics as indicated by the positive sign in the explained part—which in turn serves to increase the overall earnings gap. At the same time, the unexplained part (capturing all the factors that cannot be attributed to differences in observed worker characteristics) accounts for a sizable share of the formal-informal sector earnings differential. Indeed, for the no labor contract and the no pension benefits definitions of informality, the unexplained part is even larger than the explained part (Table 9). While there does not seem to be any evidence available on formal-informal sector wage or earnings decompositions *per se* for Serbia, these results are in line with those in Marcouiller et al. (1997)—for Mexico, El Salvador, and Peru—that on the one hand observables explain sizeable parts of the

Table 9 Overall earnings decompositions using three alternative informality measures: October 2008 and October 2009

	October 2008			*October 2009*		
	(1) Not formally registered	*(2) No labor contract*	*(3) No pension benefits*	*(1) Not formally registered*	*(2) No labor contract*	*(3) No pension benefits*
Three-fold:						
Endowments	0.553***	0.323***	0.262***	0.308***	0.288***	0.287***
	[0.003]	[0.002]	[0.002]	[0.002]	[0.001]	[0.002]
Coefficients	0.151***	0.186***	0.230***	0.096***	−0.162***	0.161***
	[0.011]	[0.004]	[0.003]	[0.006]	[0.004]	[0.004]
Interaction	0.211***	0.156***	0.151***	0.237***	0.397***	0.008**
	[0.012]	[0.004]	[0.004]	[0.007]	[0.005]	[0.004]
Two-fold:						
Explained	0.553***	0.323***	0.262***	0.308***	0.288***	0.287***
	[0.003]	[0.002]	[0.002]	[0.002]	[0.001]	[0.002]
Unexplained	0.362***	0.342***	0.381***	0.334***	0.236***	0.169***
	[0.004]	[0.003]	[0.003]	[0.004]	[0.003]	[0.002]
N	2,783	2,783	2,783	2,577	2,577	2,577

Values in brackets are robust Huber-White (Huber 1967; White 1980) standard errors. The decompositions treat the formal sector workers as the reference (Oaxaca 1973; Blinder 1973)

: statistically significant at 5%; *: statistically significant at 1%. *Source:* Serbia Labor Force Survey (October 2008 and October 2009)

formal-informal sector wage gap, while on the other hand a sizeable part is still left unexplained.

In turn, this is indicative of the presence of what in studies of gender and ethnic earnings differentials has been termed "discrimination" towards informal sector workers—which is here instead suggested to be interpreted as reflecting non-observable characteristics of informal sector workers such as relatively lower bargaining power and less access to personal and professional networks (again, informality clearly is not an innate characteristic such as gender or race). Notably, the relative shares of the explained and the unexplained parts of the formal-informal sector earnings gap have changed from pre- to post crisis in several cases, though not consistently so across all three measures. For the two preferred measures, no labor contract and no pension benefits, the unexplained share of the overall gap decreased substantially, from accounting for more than half the gap (51.4 and 59.3%, respectively) before the crisis, to accounting for less than half the gap (45 and 37.1%, respectively) one year into the crisis. In turn, this might reflect a decreasing importance of non-observable characteristics such as bargaining power and access to personal and professional networks, both of which are presumably lower among informal sector workers. Similarly, this implies that observable characteristics such as education are now relatively more important in explaining the overall earnings gap.

But how are the overall gaps—both two- and three-fold—explained by the endowment of and returns to the separate individual characteristics (or groups of characteristics), say, education and industry, for example, rather than by the endowment of and returns to individual characteristics *overall*? This is the object of the final main empirical analysis—the detailed earnings decompositions—following next.

5.3 Detailed earnings decompositions

While the overall earnings decompositions examined in the previous section already add to the story about the nature of the formal-informal sector gap in Serbia established in its "raw" form in Table 8, additional insights may be had from going one step further and additionally decomposing these overall decompositions into the contribution coming from the individual explanatory variables from the Mincer earnings regressions—which, again, can be done both for the two- and three-fold decompositions. To help better facilitate interpretation, however, results are reported in groups of individual variables (for example, aggregating up the contribution from all the education variables) rather than separately for all the individual variables (for example, for each educational level).

The results from the detailed three-fold decompositions (Tables 10 and 11) reveal that many of the observable characteristics work to widen the formal-informal sector gap rather than to narrow it. Most notably, education and part-time status are consistently important factors across all three informality measures and both time periods, though occupation also is quite (indeed, sometimes even more) important. For example, using no labor contract status as the informality measure, education and part-time status are associated with a widening of the earnings gap in October 2008 of about 7 and 5 percentage-points, respectively. Occupation, at almost 11 percentage-points, is even more important in terms of widening the gap—indicating once again how the dominance of lowly enumerated occupations such as elementary occupations (including unskilled agriculture and fishery), which dominate the informal sector, works to widen

Table 10 Detailed three-fold earnings decompositions: october 2008

	(1) Not formally registered			*(2) No labor contract*			*(3) No pension benefits*		
	Endowments	*Returns*	*Interaction*	*Endowments*	*Returns*	*Interaction*	*Endowments*	*Returns*	*Interaction*
Age cohort	0.032***	0.087***	−0.041***	0.031***	−0.009***	0.015***	0.029***	−0.001	0.007***
	[0.001]	[0.002]	[0.002]	[0.000]	[0.002]	[0.001]	[0.000]	[0.002]	[0.001]
Education	0.086***	0.086***	0.287***	0.071***	−0.010***	0.010***	0.061***	−0.009***	0.017***
	[0.001]	[0.003]	[0.004]	[0.001]	[0.002]	[0.002]	[0.001]	[0.002]	[0.002]
Part-time	0.154***	−0.340***	0.220***	0.051***	−0.258***	0.096***	0.021***	−0.334***	0.124***
	[0.002]	[0.005]	[0.004]	[0.001]	[0.005]	[0.002]	[0.002]	[0.005]	[0.002]
Industry	0.066***	−0.039***	−0.140***	0.021***	−0.016***	−0.026***	0.022***	−0.022***	−0.033***
	[0.002]	[0.004]	[0.004]	[0.001]	[0.002]	[0.002]	[0.001]	[0.002]	[0.002]
Occupation	0.163***	0.115***	−0.01	0.108***	−0.074***	−0.061***	0.098***	−0.056***	0.006***
	[0.001]	[0.003]	[0.008]	[0.001]	[0.002]	[0.003]	[0.001]	[0.002]	[0.002]
Firm size	0.044***	0.156***	−0.056***	0.030***	0.020***	0.105***	0.028***	0.017***	0.029***
	[0.001]	[0.006]	[0.003]	[0.000]	[0.002]	[0.002]	[0.000]	[0.002]	[0.002]
Urban	0.003***	0.017***	−0.044***	0.004***	0.013***	−0.021***	0.003***	0.017***	−0.023***
	[0.000]	[0.001]	[0.003]	[0.000]	[0.001]	[0.001]	[0.000]	[0.001]	[0.001]
Region	0.005***	−0.115***	−0.005	0.006***	−0.014***	0.038***	0.001**	−0.020***	0.022***
	[0.000]	[0.002]	[0.004]	[0.000]	[0.001]	[0.001]	[0.000]	[0.001]	[0.001]
Constant		0.185***			0.534***			0.638***	
		[0.012]			[0.006]			[0.006]	
N	2,783	2,783	2,783	2,783	2,783	2,783	2,783	2,783	2,783

Notes: Values in brackets are robust Huber-White (Huber 1967; White 1980) standard errors, computed according to Jann (2008). The decompositions treat the formal sector workers as the reference (Oaxaca 1973; Blinder 1973). **: statistically significant at 5%; ***: statistically significant at 1%
Source: Serbia Labor Force Survey (October 2008)

Table 11 Detailed three-fold earnings decompositions: october 2009

	(1) Not formally registered			*(2) No labor contract*			*(3) No pension benefits*		
	Endowments	*Returns*	*Interaction*	*Endowments*	*Returns*	*Interaction*	*Endowments*	*Returns*	*Interaction*
Age cohort	−0.001***	0.200***	−0.018***	0.022***	−0.150***	−0.020***	0.023***	−0.021***	−0.015***
	[0.000]	[0.004]	[0.002]	[0.000]	[0.004]	[0.001]	[0.000]	[0.007]	[0.001]
Education	0.053***	−0.061***	0.035***	0.070***	0.129***	0.114***	0.060***	0.005*	−0.001
	[0.001]	[0.003]	[0.003]	[0.001]	[0.003]	[0.005]	[0.001]	[0.003]	[0.004]
Part-time	0.064***	−0.110***	0.064***	0.019***	−0.241***	0.086***	0.040***	−0.080***	0.026***
	[0.001]	[0.005]	[0.003]	[0.001]	[0.004]	[0.002]	[0.001]	[0.004]	[0.001]
Industry	0.026***	0.171***	−0.118***	0.017***	−0.040***	0.150***	0.021***	−0.034***	0.055***
	[0.001]	[0.003]	[0.005]	[0.001]	[0.003]	[0.005]	[0.001]	[0.002]	[0.004]
Occupation	0.074***	−0.019***	0.028***	0.081***	−0.076***	0.023***	0.076***	−0.073***	−0.064***
	[0.001]	[0.003]	[0.006]	[0.001]	[0.002]	[0.003]	[0.001]	[0.002]	[0.004]
Firm size	0.068***	−0.037***	0.184***	0.052***	−0.001	0.040***	0.045***	−0.009***	0.011***
	[0.001]	[0.004]	[0.008]	[0.001]	[0.002]	[0.002]	[0.000]	[0.002]	[0.002]
Urban	0.011***	−0.020***	0.035***	0.011***	0.029***	−0.039***	0.011***	0.023***	−0.030***
	[0.000]	[0.001]	[0.002]	[0.000]	[0.001]	[0.001]	[0.000]	[0.001]	[0.001]
Region	0.014***	−0.037***	0.027***	0.016***	0.007***	0.043***	0.011***	−0.013***	0.026***
	[0.001]	[0.001]	[0.002]	[0.000]	[0.001]	[0.001]	[0.000]	[0.001]	[0.001]
Constant		0.01			0.181***			0.363***	
		[0.007]			[0.008]			[0.007]	
N	2,577	2,577	2,577	2,577	2,577	2,577	2,577	2,577	2,577

Notes: Values in brackets are robust Huber-White (Huber 1967; White 1980) standard errors, computed according to Jann (2008). The decompositions treat the formal sector workers as the reference (Oaxaca 1973; Blinder 1973). *: statistically significant at 10%; ***: statistically significant at 1%
Source: Serbia Labor Force Survey (October 2009)

the earnings gap. In comparison, industry "only" contributes a (still sizable) 2.1 percentage-points to the earnings gap.

The results from the detailed two-fold decompositions are mostly consistent with the results for the detailed three-fold decompositions, so that education and part-time status again are among the most consistently important contributors to widening the formal-informal sector earnings gap across all three informality measures (not shown here but available upon request).

5.4 Sensitivity analysis

As also discussed in the estimations strategy, a number of potential issues may remain regarding the results discussed previously. I will therefore now explore these in turn, ordered in terms of their (increasing) relative invasiveness in terms of altering various dimensions of the main analysis presented above.[21]

First, while the estimation strategy followed above erred on the inclusion of explanatory variables (as long their inclusion was theoretically justified), it is often possible to think of additional potentially relevant additional explanatory variables. One such variable is marital status, whose relationship with wages and earnings has led to the creation of an entire subfield of economics itself (see, for example, Antonovics and Town 2004). Hence, while the main focus here is on the association between informality status and earnings, controlling only for the "usual suspects" as according to standard human capital theory, it still seems useful to at least examine the robustness of the previous results when additionally including this variable—keeping in mind that this variable is arguably endogenous from a theoretical point of view. Starting with the results for the Mincer earnings regressions, augmenting the regressions with marital status[22] reveals a positive and statistically significant marriage premium, of about 7% for October 2008 and about 11% for October 2009—with the increase in the premium possibly being due to the marriages of higher earning males being more likely to survive the economic strains imposed by the financial crisis than those of lower earning males (which similarly seem more likely to get divorced due to the economic strains imposed by the financial crisis). Importantly, however, there is virtually no effect on the other estimated coefficients from including marital status. In particular, the estimated informality coefficients are identical, not differing until the third decimal. For the three- and two fold decompositions, there are also only minimal differences when augmenting with the marital status variable.

Second, while it appears reasonable to choose the formal sector earnings structure as the reference in this particular application, the choice of weights in the decomposition remains controversial and may potentially give rise to different results, depending on the weights used (Oaxaca and Ransom 1994). Among the different alternatives, one extreme, following Oaxaca (1973); Blinder (1973), assigns a weight of 1 to the endowments and returns of the "favored group"—here, the formal sector workers—in the decomposition. Again, this is also the specification used in the main analysis above (since the formal sector may be interpreted as the dominant sector in the application here, at least in an economic sense/size-wise). Another extreme, also included in Oaxaca (1973), assigns instead a weight of 1 to the informal sector. Since the seminal work in Oaxaca (1973); Blinder (1973), several alternatives in between these two extremes have been suggested, including

0.5 (Reimers 1983) and the relative share of the "favored" group (Cotton 1988) out of the overall sample.

The results turn out to be fairly robust across the different specifications of weights. For the overall three-fold decomposition, for example, the results are qualitatively similar whether the decomposition is performed using the formal sector earnings structure or the informal sector earnings structure, except for the case of the no labor contract informality measure for October 2009, where the returns help decreasing the earnings gap. For the overall two-fold decompositions the results are qualitatively similar across all the different weights,[23] though of course with some differences in magnitude. Turning to the detailed decompositions, the results are qualitatively similar for both the three-fold and two-fold decompositions. In terms of magnitude, however, education and part-time status turn out to be even more important in terms of explaining the formal-informal sector earnings gap than for the main result reported previously. Considering, for example, the two-fold decomposition in October 2009 using the no labor contract informality measure, the relative importance of education in explaining the overall earnings gap using the informal sector earnings structure as base is more than double that using the formal sector earnings structure from the main analysis. Similarly, part-time status accounts for more than five times the amount of the overall gap in this case relative to using the formal sector earnings structure from the main analysis.

Third, there might be some concern that part-time status, industry and occupation could be endogenous to earnings. Lacking any suitable instruments for these variables, this can be examined a bit more by excluding these variables from the analysis, however, and then checking whether that causes any major changes to the results. I do this by excluding each of these variables in turn and estimating the resulting reduced specification. As perhaps can be expected, the coefficient on informality in the Mincer regressions with the controls added becomes somewhat more negative—again consistent with many informal sector workers being part-time workers and with industry and occupation also being important determinants of the wage gap, so that the estimated informality coefficients pick this up whenever these three other factors are excluded. On the other hand, omitting these variables do not dramatically change the coefficients on most of the key explanatory variables—most notably the magnitudes of the estimated coefficients of the education variables are remarkably similar. The results of the decompositions analysis are also very similar. For example, for the two preferred measures, no labor contract and no pension benefits, the unexplained share of the overall gap again decreased substantially following the crisis. In turn, this indicates that part-time status, industry, and occupation are not obviously endogenous, at least not empirically—or, at least their potential (theoretical) endogeneity does not have serious empirical consequences. In any case, the results are fairly robust to whether these variables are included or not.

Fourth, while the previous analysis erred on inclusion (which also helped secure a relatively larger sample size) regarding age of workers in the estimation sample, it could be argued that prime-age males (25–64 years of age) might be of special interest. To examine whether the results are robust to the age specification, the analysis is therefore also carried out restricted to only males 25–64 years of age. The results again are quite similar, again most notably leading to very similar estimated

coefficients in the Mincer regressions and, therefore, also very similar to results in the decompositions analysis.

Fifth, while restricting the estimation sample in the main analysis to males ensures that gender-issues do not confound the analysis, thus enabling focusing more narrowly on the formal-informal sector earnings gap *per se*, it would still be interesting to see whether the results are robust to this restriction. I therefore added female workers to the estimation sample and reran the analysis. One difference, of course, is the larger sample sizes (4,860 vs. 2,783 and 4,601 vs. 2,796, for October 2008 and October 2009, respectively). The results are very similar to the main results, though not surprisingly with some differences in magnitude. The results from the Mincer regressions are also very similar to those from the main results, again showing a substantial decrease in the coefficient for informality status, once controls are added. At the same time, again, a sizeable fraction is left unexplained by observable characteristics. The decomposition analysis similarly again reveals that while a substantial fraction of the earnings gap can be explained by the less favorable characteristics of informal sector workers, a sizeable fraction (though somewhat smaller than for the main analysis, at least for the two preferred measures) is left unexplained. At the same time, however—in accordance with the main analysis—the unexplained part has decreased substantially following the crisis. For the two-fold decomposition for the two preferred measures, for example, the unexplained part for the no labor contract measure decreased from 44.5 to 38.8% of the overall gap from October 2008 to October 2009, while the similar figures for the no pension benefits measure are 49.0 and 30.6%, respectively.

Lastly, while it was argued that using earnings was preferable for the application here, some might argue that wage rates are preferable to earnings. The main argument here is that hours worked in the two sectors can be very different, which would then impose a bias of the earnings gap. To examine this possibility, the robustness of the results to the choice of dependent variable is explored here as the final sensitivity analysis. Rerunning the analysis for hourly wages instead of monthly earnings leads to almost identical sample sizes (the only difference is a drop of 3 observations for October 2008). The results are again quite similar to the main results, though not surprisingly again with some differences in magnitude. The results from the Mincer regressions again are very similar to those from the main specifications, again showing a substantial decrease in the coefficient for informality status, once controls are added. At the same time, again, a sizeable fraction is left unexplained by observable characteristics. The decomposition analysis similarly again reveals that while a substantial fraction of the earnings gap can be explained by the less favorable characteristics of informal sector workers, a sizeable fraction (though somewhat smaller than for the main analysis, at least for the two preferred measures) is left unexplained. At the same time, however—in accordance with the main analysis—the unexplained part has decreased substantially following the crisis. For the two-fold decomposition for the two preferred measures, for example, the unexplained part for the no labor contract measure decreased from 55.5 to 52.2% of the overall gap from October 2008 to October 2009 (so a smaller decrease than for the main analysis), while the similar figures for the no pension benefits measure are 59.8 and 39.9%, respectively.

In conclusion, the sensitivity analysis reveals that the results from the main analysis are quite robust in several dimensions. Certainly the main results are robust in qualitative terms, all revealing the existence of a substantively large formal-informal

earnings gap (or wage gap, for one of the sensitivity analyses) which, although a substantial fraction can be explained, at the same time leaves a sizeable fraction unexplained—a fraction which is found to have decreased substantially following the crisis.

Finally, to help both examine more and justify better the negative selection into informal jobs referred to earlier, simple (linear probability) regressions of the determinants of informality status are estimated (Table 12). These results should not be taken as causal but are merely meant to give an idea of the characteristics—or proximate determinants—that appears prevalent for informal sector workers. From these results, it appears that informal sector workers are not to be found among the prime-age adults: they tend to be either younger (15–24 years of age) or older (65 plus) than formal sector workers. Informal sector workers also tend to be less educated than formal sector workers. Considering the contract-based informality measure, for example, informal sector workers are about 7–8 percentage-points less likely to have completed tertiary education than formal sector workers. Informal sector workers also tend to be part-time, with the lowest estimated coefficient here being about 25 percentage-points—and most of the remaining ones being far higher. Informal sector workers also tend to come from all other industries than agriculture (the reference sector in the regressions for Table 12)—which certainly can be seen as a marginal sector, in line with Lewis (1954); Kuznets (1955). Informal sector workers are far less likely to come from more desired occupations such as legislators, professionals, technicians and service workers than from elementary occupations (the reference group)—where the latter again can be interpreted as a marginal occupation. Informal sector workers are more likely to come from small firms (five or less)—the caveat using this measure as a proxy for informality status *per se* notwithstanding. Lastly, there does not seem to be too strong geographical patterns related to informality status, though there is a slight tendency for informality to be more of a rural phenomenon than an urban one—though the estimated coefficient here is small, at about minus 0.3 to about minus 0.5 percentage-points. Altogether this supports the notion of "the" informal sector as being a sector best characterized as a marginal sector in terms of education and other individual and job characteristics.

6 Conclusions

This paper examines the formal-informal sector earnings gap in terms of its prevalence, magnitude, and determinants using a recent Labor Force Survey collected in Serbia over a period roughly spanning the first year of the recent international financial crisis using identical survey instruments—and thereby adds to our still fairly limited understanding regarding the formal-informal sector earnings gap of the former socialist regimes of Eastern Europe and Central Asia in general and of Serbia in particular. In so doing, specific emphasis is on trying to understand better both the extent of the formal-informal sector earnings gap and the factors driving this gap across multiple dimensions of informality—as well as whether any such patterns, if found to be present, have changed following the onset of the crisis.

Estimation of raw formal-informal sector earnings gaps and overall and detailed earnings decompositions leads to six main results: (1) the presence of a substantively large formal-informal sector gap (favoring the formal sector); (2) the gap appears to have decreased substantially overall following the crisis (though with

Table 12 OLS/LPM regressions of informality status determinants: october 2008 and october 2009

	October 2008			*October 2009*		
	(1) Not formally registered	*(2) No labor contract*	*(3) No pension benefits*	*(1) Not formally registered*	*(2) No labor contract*	*(3) No pension benefits*
25–34 years old	−0.024***	−0.082***	−0.089***	−0.011***	−0.061***	−0.098***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.002]
35–44 years old	−0.035***	−0.101***	−0.103***	0.006***	−0.076***	−0.114***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.002]
45–54 years old	−0.035***	−0.134***	−0.147***	0.001	−0.099***	−0.135***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.002]
55–64 years old	−0.035***	−0.110***	−0.118***	−0.012***	−0.101***	−0.127***
	[0.001]	[0.001]	[0.002]	[0.001]	[0.001]	[0.002]
65 years plus	0.076***	0.120***	0.092***	0.008***	−0.060***	−0.006
	[0.006]	[0.007]	[0.008]	[0.003]	[0.006]	[0.005]
Secondary	−0.024***	−0.083***	−0.079***	−0.004***	−0.070***	−0.069***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Tertiary	−0.029***	−0.075***	−0.063***	−0.007***	−0.079***	−0.067***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Part-time	0.251***	0.363***	0.431***	0.301***	0.407***	0.374***
	[0.002]	[0.002]	[0.002]	[0.003]	[0.003]	[0.003]
Manufacturing/ Mining/Electrical	−0.139***	−0.217***	−0.205***	−0.107***	−0.142***	−0.142***
	[0.001]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Construction	−0.057***	−0.018***	−0.032***	−0.039***	0.013***	−0.025***
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Trade/Services	−0.147***	−0.186***	−0.178***	−0.118***	−0.143***	−0.154***
	[0.001]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Hotel/Restaurants	−0.166***	−0.109***	−0.107***	−0.133***	−0.095***	−0.121***
	[0.002]	[0.003]	[0.003]	[0.002]	[0.003]	[0.003]
Transport	−0.130***	−0.221***	−0.228***	−0.110***	−0.138***	−0.152***
	[0.001]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Financial/Real Estate	−0.145***	−0.218***	−0.198***	−0.095***	−0.135***	−0.116***
	[0.001]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Public Sector	−0.133***	−0.220***	−0.219***	−0.113***	−0.152***	−0.163***
	[0.001]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Other Sector	−0.165***	−0.229***	−0.206***	−0.093***	−0.082***	−0.133***
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Legislators	−0.059***	−0.152***	−0.170***	−0.036***	−0.073***	−0.073***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Professionals	−0.064***	−0.155***	−0.159***	−0.021***	−0.065***	−0.066***
	[0.001]	[0.001]	[0.002]	[0.001]	[0.001]	[0.001]
Technicians	−0.074***	−0.158***	−0.151***	−0.047***	−0.080***	−0.084***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Clerks	−0.053***	−0.139***	−0.135***	−0.041***	−0.062***	−0.073***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Service	−0.074***	−0.135***	−0.134***	−0.042***	−0.071***	−0.059***

Table 12 OLS/LPM regressions of informality status determinants: october 2008 and october 2009 *(Continued)*

	[0.001]	[0.001]	[0.002]	[0.001]	[0.001]	[0.001]
Skilled	0.011***	0.090***	0.057***	0.070***	0.046***	0.120***
	[0.004]	[0.005]	[0.005]	[0.004]	[0.005]	[0.005]
Craft/trade	−0.050***	−0.104***	−0.104***	−0.024***	−0.064***	−0.051***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Plant/Machine operators	−0.079***	−0.145***	−0.151***	−0.043***	−0.096***	−0.092***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Firm size 6–19	−0.042***	−0.049***	−0.058***	−0.037***	−0.047***	−0.062***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Firm size 20–99	−0.056***	−0.097***	−0.117***	−0.050***	−0.079***	−0.080***
	[0.000]	[0.001]	[0.001]	[0.000]	[0.001]	[0.001]
Firm size 100 plus	−0.061***	−0.112***	−0.120***	−0.049***	−0.078***	−0.072***
	[0.000]	[0.001]	[0.001]	[0.000]	[0.001]	[0.001]
Firm size not sure:10 or less	−0.062***	0.130***	0.106***	0.058***	0.126***	0.079***
	[0.002]	[0.003]	[0.003]	[0.002]	[0.003]	[0.003]
Firm size not sure: 11 plus	−0.047***	−0.074***	−0.098***	−0.041***	−0.075***	−0.059***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Urban	−0.005***	−0.005***	−0.004***	−0.003***	−0.003***	−0.003***
	[0.000]	[0.001]	[0.001]	[0.000]	[0.001]	[0.001]
Central Serbia	−0.007***	0.005***	−0.007***	0	0.008***	0
	[0.000]	[0.001]	[0.001]	[0.000]	[0.001]	[0.001]
Vojvodina	0.012***	0.021***	0.015***	−0.004***	0.024***	0.029***
	[0.000]	[0.001]	[0.001]	[0.000]	[0.001]	[0.001]
Constant	0.298***	0.599***	0.625***	0.190***	0.416***	0.462***
	[0.002]	[0.002]	[0.003]	[0.002]	[0.003]	[0.003]
R^2	0.242	0.325	0.305	0.189	0.237	0.212
N	2,783	2,783	2,783	2,577	2,577	2,577

Notes: Estimations incorporate sampling weights. Values in brackets are robust Huber-White (Huber 1967; White 1980) standard errors. ***: statistically significant at 1%
Source: Serbia Labor Force Survey (October 2008 and October 2009)

some variation across informality measures); (3) however, when controlling for observable characteristics, the gap has not really changed that much following the crisis—in turn indicating persistency in the gap once observable characteristics have been controlled for, thus indicating that workers with relatively more favorable characteristics have been "pushed" into informality following the crisis; (4) both endowments and the returns to characteristics increase the earnings gap—indicating that formal sector workers are concentrated in better paying industries and occupations, have more education, and so on, and at the same time also have higher returns to their (already favorable) characteristics overall; (5) for the two preferred measures, no labor contract and no pension benefits, the unexplained share of the overall gap decreased substantially, thus indicating a decreasing importance of non-observable characteristics such as bargaining power and access to personal and professional networks in explaining the relative earnings position of informal

sector workers; and (6) pursuing detailed decompositions of the formal-informal earnings gap indicates that education and part-time status consistently are among the main drivers of the observed gap across the different alternative specifications of the two- and three-fold decompositions.

These results have strong policy implications. The finding of a substantial formal-informal sector earnings gap and the fact that most of the informal sector workers have less favorable personal characteristics (most notably, lower levels of education) indicates the existence of a highly marginal sector composed of marginalized workers. This of course is (or if not, should be) a policy concern, since these workers have families to feed and children to raise, just as formal sector workers. At the same time, it is also important to stress that these results do not point to an importance of informal sector wages and earnings "catching up" with formal sector wages and earnings. After all, the informal sector effectively undermines governance and society because it neither contributes to society (tax evasion) nor does it operate under its rules (regulatory evasion). Hence, the overall goal of the government would seem to be to bring the informal sector into the formal sector, whereas equalizing wages would rather encourage informal employment. Instead, policy interventions might include increased support and programs to workers as well as re- and up-skilling of workers. Another viable policy option might be to subsidize jobs for low skilled workers (many of whom, again, work in the informal sector) either via direct subsidies or via reducing social security contributions; indeed, this might even give larger employment effects than subsidizing jobs for workers of all skill levels (Lehmann 2010: 15).

Relatedly, the results are consistent with part-time work being a main driver of informality. Coupling this with the fact that part-time work carries a very high tax burden because of minimum social security contributions in Serbia (Koettl 2010), part-time work in Serbia simply doesn't pay off, in particular for low-wage earners—many of which, again, are informal sector workers, working part-time. Future policy might want to consider ways in which to try to make it more favorable for workers to switch from the informal to the formal sector—for example, by introducing a system of strongly graduated social security contributions, linked to part-time vs. full-time status. That is, an informal sector worker joining the formal sector would start off with relatively much lower contributions and could then later possibly "graduate" to a full-time formal sector job. If the net effect on such a scheme decreases the incentive to continue informal part-time work, it may well be revenue-neutral (or, perhaps, even lead to a net increase in contributions).

Endnotes

[1]For extensive reviews see, for example, Blunch et al. (2001); Fields (2007); Kanbur (2009); Lehmann (2010); Perry et al. (2006); Ruffer and Knight (2007). This is also related to the notion of segmented labor markets originating with Lewis (1954); Kuznets (1955) and extensively discussed and reviewed more recently for the case of developing countries in Fields (2009).

[2]Whereas Lehmann and Pignatti (2007) consider the case of Ukraine, I find the same to be the case for Serbia (see Section 3).

[3]As also exemplified by the several dimensions included in "the" definition of informal employment by The International Labour Organization (ILO): (i) employers and own-account workers who work in their own non-registered enterprises; (ii) contributing family

workers; (iii) members of informal producers' cooperatives; (iv) employees whose employment relationship is not subject to national labor legislation, income taxation, social protection or entitlement to certain employment benefits; and (v) own-account workers engaged in the production of goods exclusively for own final use by their household (Hussmanns 2004).

[4]The remainder of this section draws extensively on Koettl (2010).

[5]Somewhat surprisingly, the survey does not collect information on union membership, which appears to be a potentially relevant variable when examining earnings determinants.

[6]It is fairly easy to receive health insurance for free as an informal employee in Serbia: workers with no formal income on record can get free health insurance from the health insurance fund; hence, many of the informally employed have it, too. In addition, many informally employed might get co-insurance through their spouse. I therefore focus on pension benefits, only (the results are quite robust to whether or not health benefits are also included, however).

[7]This ranking also roughly accords with ILO's definition (Hussmanns 2004), also given earlier in footnote 3.

[8]As a robustness check, one of the sensitivity analyses will examine the sample where the youngest and oldest age categories are excluded from the estimation samples.

[9]The dynamics in the female labor market are potentially quite different from those of the male labor market—as exemplified by, for example, the substantially lower female labor force participation rate in many countries, including Serbia, as well as potentially quite different determinants of that participation.

[10]Using the formula: Gap in percent = 100*{exp[ln(earnings gap)] − 1}. Since the log-gap corresponds to running a regression of log earnings on a dummy of the informality measure one might suggest instead using Kennedy's (1981) bias correction for dummy variables in semi-logarithmic models—however, since the variance is so small here (and throughout the paper), the results are identical whether or not the correction is used (to the third decimal).

[11]Remembering again that due to the issue of "envelope payments" the gap is likely underestimated.

[12]The data examined were from the 2007 Living Standards Measurement Survey (LSMS) and the (composite) informality measure was based on (1) workers employed with no social contributions paid; (2) people employed in a private unregistered firm; and (3) the employed who work at home, from door-to-door, at the flea market and other similar places (the analysis also examined the 2002 LSMS but the gap here was virtually nil).

[13]With causality being an issue here due to obvious endogeneity concerns—in turn indicating that any findings should more appropriately be viewed as indicating association, rather than causality *per se*, and any conclusions similarly modified accordingly.

[14]Additionally, sensitivity analysis where these variables are not included is also performed as a robustness check.

[15]In the following, bars on top of variables denote mean values, while $\hat{\beta}$ denotes estimated coefficient values from equations (1) and (2) above.

[16]See Oaxaca (1973); Blinder (1973); Cotton (1988); Reimers (1983); Neumark (1988); Jann (2008) for different approaches—basically, these differ in the relative weights they attribute to the two groups in the decomposition.

[17]Also known as the "absence of discrimination" specification in the wage discrimination literature.

[18]The results for the Mincer regressions, spanning a total of 8 pages, are omitted here to conserve space (but are available upon request).

[19]The formal sector premium (or informal sector penalty) from the Mincer regressions reported here and in the following are all de-logged, again using the formula: earnings premium (penalty) in percent = 100*{exp[coefficent] – 1}. Due to the small variance here, the results are again identical whether or not Kennedy's (1981) bias correction is used instead (to the third decimal).

[20]The estimated gap for the 2002 LSMS, now controlling for worker characteristics, again was neither statistically nor substantively significant.

[21]The tables are omitted here for brevity but are available upon request.

[22]In practice, marital status is incorporated in the estimations by including three additional dummy variables: "Married," "Widowed" and "Divorced" (with "Single" being the reference/omitted category).

[23]Again, using the alternative approaches outlined in Oaxaca (1973); Blinder (1973); Cotton (1988); Reimers (1983); Neumark (1988); Jann (2008).

Competing interests

The IZA Journal of Labor & Development is committed to the IZA Guiding Principles of Research Integrity. The author declares that he has observed these principles.

Acknowledgements

I thank Johannes Koettl, Hartmut Lehmann, Klara Peter, David Ribar, Indhira Santos, Kostantinos Tatsiramos, and participants at the IZA/World Bank workshop on *Institutions and Informal Employment in Emerging and Transition Economies* and at the Annual Meetings of the European Society for Population Economics (ESPE) for helpful comments and suggestions. I am also thankful to an anonymous referee for many valuable comments and suggestions. Remaining errors and omissions are my own. The data were kindly provided by the Statistical Office of the Republic of Serbia (RSO). The assistance of Vladan Božanić, RSO, to help me understand the data better is highly appreciated. The findings and interpretations are those of the author and should not be attributed to the Statistical Office of the Republic of Serbia.

Responsible editor: Hartmut Lehmann

Received: 4 January 2015 Accepted: 15 July 2015
Published online: 27 August 2015

References

Antonovics K, Town R (2004) Are all the good men married? Uncovering the sources of the marital wage premium. Am Econ Rev 94(2):317–321

Arandarenko M (2007) Transition from Education to Work: Serbia Country Report. Working Paper, European Training Foundation (ETF), Turin

Babović M (2008) The Position of Women on the Labour Market in Serbia. Belgrade, Gender Equality Council, Government of the Republic of Serbia and United Nations Development Programme, Serbia

Blinder AS (1973) Wage discrimination: reduced form and structural estimates. J Hum Resour 8:436–455

Blunch N-H, Canagarajah S, Raju D (2001) The Informal Sector Revisited: A Synthesis across Space and Time. Social Protection Discussion Paper No. 0119. Human Development Network, World Bank, Washington, D.C

Cotton J (1988) On the decomposition of wage differentials. Rev Econ Stat 70:236–243

Fields GS (2007) Employment in Low-Income Countries: Beyond Labor Market Segmentation? In: Pierella P, Pieter S (eds) Employment and Shared Growth: Rethinking the Role of Labor Mobility for Development (Directions in Development). World Bank Publications, Washington, DC

Fields GS (2009) Segmented Labor Market Models in Developing Countries. In: Kincaid H, Ross D (eds) The Oxford Handbook of the Philosophy of Economic Science. Oxford University Press, Oxford

Hart K (1971) Informal Income Opportunities and Urban Employment in Ghana. In: Jolly R et al (eds) Third World Employment: Problems and Strategy. Harmondsworth, Penguin

Huber PJ (1967) The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions. In: Le Cam LM, Jerzy N (eds) Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability Vol. 1. University of California Press, Berkeley, CA

Hussmanns R (2004) Measuring the Informal Economy: From Employment in the Informal Sector to Informal Employment. Working Paper No. 53. Policy Integration Department, Bureau of Statistics, International Labour Organization, Geneva

Jann B (2008) The Blinder-Oaxaca decomposition for linear regression models. Stata J 8(4):453–479

Kanbur R (2009) Conceptualising informality: regulation and enforcement. Indian J Lab Econ 52(1):33–42

Kennedy PE (1981) Estimation with correctly interpreted dummy variables in semilogarithmic equations. Am Econ Rev 71(4):801
Koettl J (2010) Does Formal Work Pay in Serbia? The Role of Labor Taxes and Social Benefit Design in Providing Disincentives for Formal Work. Technical Note, World Bank, Washington, DC
Kogan I (2011) When informal is normal… on the role of credentials and contacts for the job entry in Serbia. Res Soc Stratif Mobil 29(4):445–458
Krstić G, Sanfey P (2011) Earnings inequality and the informal economy: Evidence from Serbia. Econ of Transition 19(1):179–199
Kuznets S (1955) Economic growth and income inequality. Am Econ Rev 45:1–28
Lehmann H (2010) Policies to Combat Informality and to Broaden the Tax Base: Lessons for Transition Countries. Background paper to a World Bank project on informality in the Balkans. Department of Economics, University of Bologna, Bologna
Lehmann H (2015) Informal employment in transition Countries: empirical evidence and research challenges. Comp Econ Stud 57(1):1–30
Lehmann H, Pignatti N (2007) "Informal Employment Relationships and Labor Market Segmentation in Transition Economies: Evidence from Ukraine," IZA Discussion Paper No. 3269. Institute for the Study of Labor (IZA), Bonn
Lehmann H, Zaiceva A (2015) Redefining informality and measuring its determinants: evidence from the Russian labour market. J Int Dev 27(4):464–488
Lewis WA (1954) Economic development with unlimited supplies of labour. Manch Sch 22:139–191
Macias V (2009) Informal Employment: The Case of Bosnia and Herzegovina, Macedonia and Serbia. Background Paper, Human Development Sector Unit, Europe and Central Asia Region, World Bank, Washington, DC
Marcouiller D, de Castilla VR, Woodruff C (1997) Formal measures of the informal-sector wage gap in Mexico, El Salvador, and Peru. Econ Dev Cult Change 45(2):367–392
Neumark D (1988) Employers' discriminatory behavior and the estimation of wage discrimination. J Hum Resour 23:279–295
Oaxaca RL (1973) Male–female wage differentials in urban labor markets. Int Econ Rev 14:693–709
Oaxaca RL, Ransom MR (1994) On discrimination and the decomposition of wage differentials. J Econom 61:5–21
Oaxaca RL, Ransom MR (1998) Calculation of approximate variances for wage decomposition differentials. J Econ Soc Measurement 24:55–61
Oaxaca RL, Ransom MR (1999) Identification in detailed wage decompositions. Rev Econ Stat 81:154–157
OECD (2008) Serbia: A Labour Market in Transition. Organization for Economic Co-operation and Development (OECD), Paris
Perry G, Maloney W, Arias O, Fajnzylber P, Mason A, Saavedra-Chanduvi J (eds) (2006) Informality: Exit and Exclusion. World Bank Latin America and Caribbean Studies, Washington, D.C
Reimers CW (1983) Labor market discrimination against Hispanic and Black Men. Rev Econ Stat 65:570–579
Ruffer T, Knight J (2007) Informal Sector Labor Markets in Developing Countries. University of Oxford, Oxford
White H (1980) A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48(4):817–830
Winsborough HH, Dickinson P (1971) Components of Negro-White Income Differences. Proceedings of the American Statistical Association, Social Statistics Section: 6–8
Yun M-S (2005) A simple solution to the identification problem in detailed wage decompositions. Econ Inq 43:766–772