

Employed Unemployed? On Shadow Employment During Transition*

Joanna Tyrowicz
University of Warsaw
National Bank of Poland
Rimini Centre for Economic Analysis

Stanisław Cichocki
University of Warsaw
National Bank of Poland

First version: April 2009

This version: March 2010

Abstract

Unregistered employment poses two types of challenges to the researchers: (i) reliably evaluating the wage differential between formally and informally employed and (ii) accounting for the push and pull factors in general and the effects of business cycle in particular. We address the former with the use of propensity score matching and analyse the evolution of the estimated average compensations and the differentials with reference to GDP and unemployment fluctuations. Using 13 years of quarterly labour force survey data from Poland on *de iure* unemployed but *de facto* employed individuals we find, that in-the-shadow compensations tend to be higher and procyclical. We also find considerable distributional heterogeneity.

Key words: undeclared employment, propensity score matching, transition

JEL Codes: O17, J22, P37

1 Introduction

Evidence on whether people in shadow employment are there due to tax evasion reasons or because of the limited access to formal labour market - is thus far mixed. Consistent with the first view, people with informal employment are actually experiencing higher wages than their formally employed counterparts' *ceteris paribus*. Conversely, if there barriers to access exist, one should expect "grey" wages to fall short of the "non-grey" ones, again, other things held equal. The empirical challenge lies not in the conceptualisation but rather in reliable implementation of the *ceteris paribus* device.

*Authors would like to thank two anonymous referees for extremely valuable comments. At various stages of working on this paper we have benefited greatly from comments by Edgar Feige, Karsten Staehr, Ryszard Kokoszcyński, Michael Pickhardt, Zbigniew Zólkiewski, Mateusz Pipień and participants of WIEM 2009 and EEFS 2009, XXIV AIEL conferences as well as the international workshop "The Shadow Economy, Tax Evasion and Social Norms", Muenster (Germany). We also appreciate comments from the participants of WNE UW, NBP and CEEERC Seminars. About the availability about the part of the data we were made aware by Gabriela Grotkowska, to whom we are also thankful. Part of the research was done when Joanna Tyrowicz was a Mellon Fellow at The Netherlands Institute for Advanced Studies, whose exceptional support is gratefully acknowledge. The remaining errors are ours.

Additional context to this ambiguity is provided by the process of economic transition and/or development. In many of the theoretical frameworks shadow activity - be it wage- or self-employment - is frequently considered one of the transitory phases, sort of tipping the toe, Bennett and Estrin (2007). On the other hand, research on developed economies seems focus more on the potential adverse effects of the taxing schemes and discrimination/social exclusion, Kopczuk (2001) and Fugazza and Jacques (2004). European transition economies and Poland among them are an interesting mixture of these two contexts. On one hand, because of the transition heritage, they were and still are characterised by high shares of shadow economy in the GDP. On the other hand, the current estimates are roughly three times lower than in mid-1990s. According to the latest available survey, about 9.6% of the labour force are active in shadow economy¹.

In this paper we use a set of 52 quarterly labour force surveys for the period 1995q1-2007q4. In each survey roughly 1-2% of the labour force reported having paid employment in the previous week despite being registered as unemployed. We apply propensity score matching to address the problem of considerable heterogeneity among regular and informal compensation - and workers. We demonstrate that characteristics of the informally employed resemble more the population of unemployed than that of the formally employed. Using this specific type of “grey” employees we sought “statistical twins” in each of the respective quarters for this very specific group. Subsequently, we have compared their self-reported earnings (both total wages and per hour compensations) among the two groups of “grey” and otherwise identical “non-grey” workers. Based on this data, we are able to identify the discrepancies in declared income for regular and this form of informal employment, while this differential is decomposable into a part attributable to the individual characteristics and a part attributable to informal employment.

The results suggest that with or without the correction for the heterogeneity - “grey” compensations behave procyclically. However, once the differences in productivity determinants are accounted for, the sign of the differential changes from relatively lower to relatively higher compensations. The differentials tend to increase in the periods of economy contraction. The evidence concerning the number of hours worked is mixed, but the differentials - if significant - tend to be small. It seems also that the access barriers are more influential in the bottom of the wages distribution.

The paper is structured as follows. Section 2 discusses relevant literature, especially the business cycle context of the informal employment studies. In section 3 we discuss in detail the nature of datasets, the informational properties of the wages data we dispose of and the empirical method applied in this study. Results are reported in section 4. We performed numerous sensitivity checks of the findings, taking into account different potential causes of the formal labour market entry barriers. Finally, in section 5 we conclude and provide some constraints concerning the interpretation of these findings but also some relevant policy implications.

2 Literature review

The literature on shadow economy is indeed surprisingly vast, if one considers how difficult to capture it is. First papers on this phenomenon developed simple definitions and basic statistical methods of

¹The estimates for the size of undeclared employment for Poland are mostly conducted by the statistical authorities, as of mid-1990s there was one dedicated survey conducted by the Ministry of Labour and Social Affairs.

measurement, Gutmann (1977). Tanzi (1983) as well as Frey and Weck-Hanneman (1984) focused on more sophisticated methods of grey economy measurement. In the course of research on shadow economy more attention was paid to causes of shadow economy Johnson, Kaufmann and Zoido-Lobaton (1998). In a widely cited paper, Schneider and Enste (2000) provide a comprehensive overview on definitions of underground economy, its size, causes, consequences and methods of measurement.

Interestingly, the literature which investigates the behaviour of shadow economy - or informal employment - during business cycle is also abundant. The traditional view states the labour market to be dualistic (formal and informal sector) with the informal being an inferior one. According to classical approach, the informal sector is in fact countercyclical: it expands during the downturns acting as an absorbent for increased unemployment, Fiess, Fugazza and Maloney (2007). Busato and Chiarini (2004) propose a two sector DSGE model and calibrate it for Italy showing that shadow employment follows this predicted pattern of the countercyclicity. This finding is also confirmed by Loayza and Rigolini (2006), who takes an extra mile to emphasise two stylised facts: (i) countries with larger unregistered employment are less countercyclical and (ii) in the long run higher informal employment is associated with the lower *per capita* GDP².

An indirect proof of countercyclical properties of unregistered employment is offered by Boeri and Garibaldi (2002). Using a search and matching model they show informal employment and unemployment to be two sides of the same coin. A reduction in the latter leads in this framework to a natural reduction in the former, as well. Repressing the informal sector causes unemployment to rise if measures providing alternatives to the unregistered employment are not implemented. Kolm and Larsen (2003) explores further the link the effect of repressions on a transmission between unemployment and shadow employment, demonstrating that intensified punishment leads to an increased wage demand in the informal sector. Consequently, lower wage demand in the formal sector occurs and as a consequence firms move from the informal sector to the formal one.

The hypothesis of countercyclical behaviour of informal employment is also confirmed by Carillo and Pugno (2004). They propose a general equilibrium model with the aim of capturing underground activities. The model has two stable equilibria: a “good” equilibrium with numerous firms, high productivity and output; a “bad” equilibrium with the opposite features. Carillo and Pugno (2004) view this two-equilibria case as consistent with the empirical finding of countercyclical behaviour of unregistered labour.

Despite convincing evidence and conceptual framework advocating the countercyclical type of relationship, also a different picture - *i.e.* procyclical behaviour of informal employment - has been demonstrated in the literature. Bovi (2006) estimates a bivariate VAR using data on shadow employment in Italy released by the ISTAT. He demonstrates a significant positive correlation between unregistered employment and output which loses momentum as time passes. He also forcefully argues that this aggregate picture is an outcome of diversified sectoral patterns - some of which prove procyclical, other countercyclical and in some cases no relationship may be pinned down. Consequently, there seems to be a huge complexity underlining the shadow employment. Chiarini and

²It should be noted that a big part of the literature dealing with informal employment and business cycle analyses the case of Italy. This is an effect of abundant and high quality official data, *i.e.* long time series, provided by Italian Statistical Office (ISTAT).

Marzano (2007) using the same data on unregistered employment but also data on tax evasion confirm the main findings of Bovi (2006).

Ihrig and Moe (2001) develop a simple partial equilibrium model of an agent's decision to work in the informal or formal sector. After estimating the model with the use of data for selected Asian countries they point to a negative correlation between GDP and informal employment and countercyclical behaviour of the latter one. Roca, Moreno and Sanchez (2001) introduce underground economy in a standard Real Business Cycle model and calibrate it for the US economy. Their main findings point to a negative relationship between the participation rate, underground economy and output fluctuations: low participation rate is associated with a bigger shadow economy and larger fluctuation in registered employment.

Fiess et al. (2007) question the traditional view of a dualistic structure of the labour market and the segmentation into formal and informal employment. Estimating a two sector labour market model for Argentina, Brazil, Colombia and Mexico they show the informal and formal sector to be rather integrated as numerous periods show strong co-movement between relative sector sizes. In fact, there has been a growing body of evidence against viewing the informal sector as inferior, which corroborates the expectation that informal employment should behave procyclically. An extensive research project summarised in four forceful papers - Bosch and Maloney (2006); Bosch and Maloney (2007); Bosch, Goni and Maloney (2007) and last but not least Bosch and Maloney (2008) - analyses explicitly the transition between sectors for various Latin American countries. These papers argue that: (i) transition into informality corresponds to more - and not less - to job-to-job transitions and also less disguised unemployment; (ii) flows from informal employment to formal employment and from formal employment to informal employment tend to be procyclical; (iii) as a consequence informal employment is countercyclical.

In the transition context, cross country studies show a much bigger size of shadow economy for Central and Eastern European countries (CEECs) than for established market economies of the EU15, Friedman, Johnson, Kaufmann and Zoido-Lobaton (2000). While the size of underground economy for EU-15 is estimated on average to be 18-19% of GDP it is about 31-32% of GDP for CEECs, Schneider (2007). Regarding unregistered employment in other CEE countries Renooy, Ivarsson, van der Wusten-Gritsai and Meijer (2004) conducted a study estimating its size from 9% of GDP in Estonia and Czech Republic to 30% of GDP in Bulgaria³, however the number of such analyses is still scarce. Poland was typically comprised in cross-country studies, while data usually originated from Central Statistical Office (CSO), which conducts biannual survey study on undeclared employment and shadow economy, Johnson et al. (1998), Schneider (2007).

Summarising, there is a large body of empirical studies analysing the response of the informal sector to the cyclical character of output and registered employment. While the consensus has not

³Research on the shadow economy in Poland is no different in this respect, while the first studies appeared even before the transition to the market economy. Bednarski, Kokoszczynski and Stopyra (1988) tried to measure its size, while already after the outbreak of transition Cassel, Jaworski, Kath, Kierczynski, Lutkowski and Paffenholz (1989) as well as Bednarski (1992) gave a detailed description of the shadow economy phenomenon. While this research coped predominantly with the consequences of centrally planned economy for the ways in which individuals exhibited economic activity, the comparability with analyses for the other countries, let alone Poland after the transformation is rather limited. The Institute for Market Research (1995) estimated that about 30% of adolescent Poles work in an undeclared fashion.

been reached on whether the link is pro or countercyclical, the respective empirical studies have been convincing in arguing the direction of the link between the *size* of the informal sector and output fluctuations. On the other hand, these analyses face obvious shortcomings. First of all, focusing on the *size* of shadow economy and/or unofficial employment, they rarely provide any insights into the verification of the determinants of this phenomenon⁴. These models typically presuppose wage formation patterns in both sectors, as well as a universal response to the changing conditions - as argued already by Bovi (2006) this is not necessarily true at a sectoral level, while differences may appear also with respect to workers (*e.g.* educational attainment, gender, age or occupation). In this paper we inquire the actual changes in wages in the informal sector as opposed to the formally employed, *ceteris paribus*.

Secondly, there has been little effort so far to analyse the mechanisms through which the size of the “grey” economy has shrunk in transition countries. Encompassing the period of 1995q1-2007q4 permits observing the realised changes in incentives for employees in both sectors of the economy. Based on theoretical and conceptual prediction, the fluctuations differential between the formal and informal compensations should be linked to the changes in employment, GDP as well as institutional events (*e.g.* changes in minimum wage, alternative costs, etc.).

Thus, we complement the literature in two major ways. Firstly, using micro-level evidence we inquire the evolutions of differentials in compensations for formal and informal employees in a transition economy. Secondly we trace the effects of the business cycle fluctuations and the institutional changes on these differentials, analysing the behavioural responsiveness of the “grey” economy compensation schemes to the external changes.

3 Data and empirical strategy

Inquiring the nature of selection into “grey economy” requires the use of individual level data and indeed this is the approach followed in this paper. However, instead of using a single data set, *i.e.* one point in time survey, we use a complete set of consecutive labour force surveys (LFS) conducted by the Central Statistical Office over the period of 1995q1 to 2007q4.

Because of the data limitations we analyse a specific category of “grey employees”. Namely, all LFS interviewees report independently their labour market status (at the beginning of the questionnaire), but also whether they are registered as unemployed with the local labour office (in the last section of the questionnaire). Using consecutive labour force surveys we identify individuals who officially report unemployment status but at the same time declare wage employment and actual compensations. Over the analysed horizon there were 10 994 individuals identified as informally employed, for whom wage, demographic and educational characteristics were also available. Using propensity score matching we identify “statistical twins” of these “grey” workers. Naturally, this is “informal” employment - even if there is any form or written or oral contract with the employer, no social security contributions are paid and the individual voluntarily *declares* at the local labour office that (s)he has no employment and is actively seeking a job.

⁴Kriz, Merikull, Paulus and Staehr (2008) discuss the determinants of employment in-the-shadow in Estonia using official, administrative data and surveys on “envelope wages”. Gerxhani (2003) focuses on institutions and social networks as a factor explaining shadow economy in Albania.

“Grey” workers as defined above do not exhaust the concept of informal employment. Typically, this wide category comprises also: (i) individuals for whom only part of the compensation is declared to the tax authorities⁵; (ii) individuals who are officially inactive (not registered as unemployed) and obtain earned income; (iii) individuals who dispose of an official job and have an additional, informal source of earned income; (iv) individuals who run their own/family company, either without registration or not reporting the corporate revenues. Especially the “grey” self-employment has received a lot of attention from the researchers, Maloney (2004). However, this subpopulation of the informally employed is particularly interesting for at least two reasons.

Firstly, although they are not receiving unemployment benefit⁶, they are allowed to apply for the social assistance⁷. Moreover, they entire family is automatically covered by the health insurance. Thus, the status of registered unemployed yields material - though frequently nonpecuniary - benefits.

On the other hand - and this is the second reason for which this group is particularly interesting - registering as unemployed despite having the ability to earn income, may be an indication of official labour market entry barriers. Namely, families of all formally employed are covered by the health insurance, while social assistance is accessible to all poorer households fulfilling income conditions - not only the labour market status. Thus, informally working individuals officially registered as unemployed may actually be the ones who are more constrained in accessing the official labour market than those who function entirely in the shadow, *e.g.* working inactive or unregistered entrepreneurs.

3.1 Data

We use 52 consecutive labour force surveys. Each set contains roughly 50 000 individuals. Surveys are collected quarterly on a representative sample of adult individuals (as of 2002 also individuals with age below 15 years of age are included) while the non-systematic refusals to participate in the survey are compensated by the weighting scheme⁸. Both the data and the weights are provided by the Central Statistical Office. The datasets from pre-1997 do not contain information on wages for employed declaring part-time contracts, but as of 1997 all employees declare wage income.

Analysing the aggregate evolutions, Figure 1, there seems to be a clear co-movement between formal and informal employment, as suggested in general in the literature. All of the shares in the upper left panel of Figure 1 are expressed in % of the labour force, which implies that the changes in these rates can be directly compared. The size of the informal sector decreases together with the formal sector, which would be consistent with the view emphasising the procyclical nature of the “grey” evolutions. However, this perception is based mainly on the profound labour market contraction between 1999 and 2005⁹. Over this horizon the unemployment doubled reaching nearly

⁵This form of tax evasion is only feasible from the employers’ initiative, as reporting obligation relies on their shoulders. It is not known how frequent this form of “grey” economy is, but a recent survey by the Ministry of Labour and Social Affairs reports that roughly 1/3rd of the informal employment takes this form, Krynska, Bednarski, Walewski and Pater (2007).

⁶Majority of the Polish unemployed do not receive it, since entitlement expires 12 to 18 months after registration, depending on the relative unemployment rate in the region of residence. In our sample, on 10 994 individuals only 2 reported receiving an unemployment benefit.

⁷This income support instrument is assigned based on the household equivalent revenue.

⁸Central Statistical Office reports the response rate to stick around 50% for the LFS, this percentage has not changed significantly over the horizon of the analysis.

⁹Lowering of the GDP growth rates covered only 2000-2002 period, whereas labour market contraction lasted till

23% towards the end of 2003. If the evolutions in the beginning and towards the end of the sample are concerned, it seems that post 2006 growth in wage employment has been associated with *decrease* in informal employment. On the other hand, the employment growth from mid 1990s is has been associated with historically highest size of the “grey” employment. The period covered by this study - at the most 52 quarterly observations - is too short to permit time series approach, but the diverging patterns of relationship between GDP already raise questions about the determinant of these dynamics.

Figure 1 depicts the evolution of the basic demographic and educational characteristics of the “grey” workers as opposed to the other participants of the labour force. Clearly, the characteristics of the Polish informal employees do not seem to deviate from what has been already found in the literature for the other countries. They tend to be younger than the formal wage earners. While they less frequently are female, they seem to have also slightly lower educational attainments than the employees. As of 2001 there is a clear change in the trend in tertiary or higher education, which is associated with the educational boom experienced by Poland. While prior to the boom the “grey” workers were slightly less frequently university graduates than the wage employed, the dramatic increase in tertiary enrolment seems to have equalised the initial differential.

Thirteen years of data cover both the up- and downturns in the economy and in the labour market, which might have affected the propensity - and opportunities - of individuals employed in the informal sector. Indeed, the unemployment rate varied between 8% and 21% based on the LFS (10% and 23% based on the registries). Also, this period captures the final stage of the transition from a centrally planned to a market economy, which implies that except for temporary swings; also longer term trends could be observed. Both the share of the informally employed in the labour force and their characteristics have witnessed considerable changes throughout this period, Figure 1.

Observing the data one can state that, while the overall trend concerning the increase in educational attainment is evident throughout the economy in general, the pace of this process among the “grey” workers has been initially smaller, with tertiary attainment picking up only as of 2001. Similarly, the ageing of the economically active population in general is slower than for the “grey” workers. These characteristics already suggest that there is some systematic selectivity into informal employment accounted for by the basic observables like the demographics or education. Consequently, our empirical strategy has to address this non-randomness.

3.2 Data on wages

Labour force surveys provide self-reported net values of wages on the main job (interviewee decides which of his possibly many jobs is the main one). Until the end of 1997, only full-time employees (self-reported) reported the wages, whereas in the remainder of the time span all employees reported net wages¹⁰. Both formally and informally employed reply to the same question, since the interviewers do not recognise those who declare registered unemployed status and contemporaneous wage

2005.

¹⁰Interviewees may refuse to reply to this question or answer “I don’t know”. In such cases the answer is coded separately. It is typical in the labour force surveys that interviewees round up the actual values of earnings, yielding in distributions with multiple local peaks.

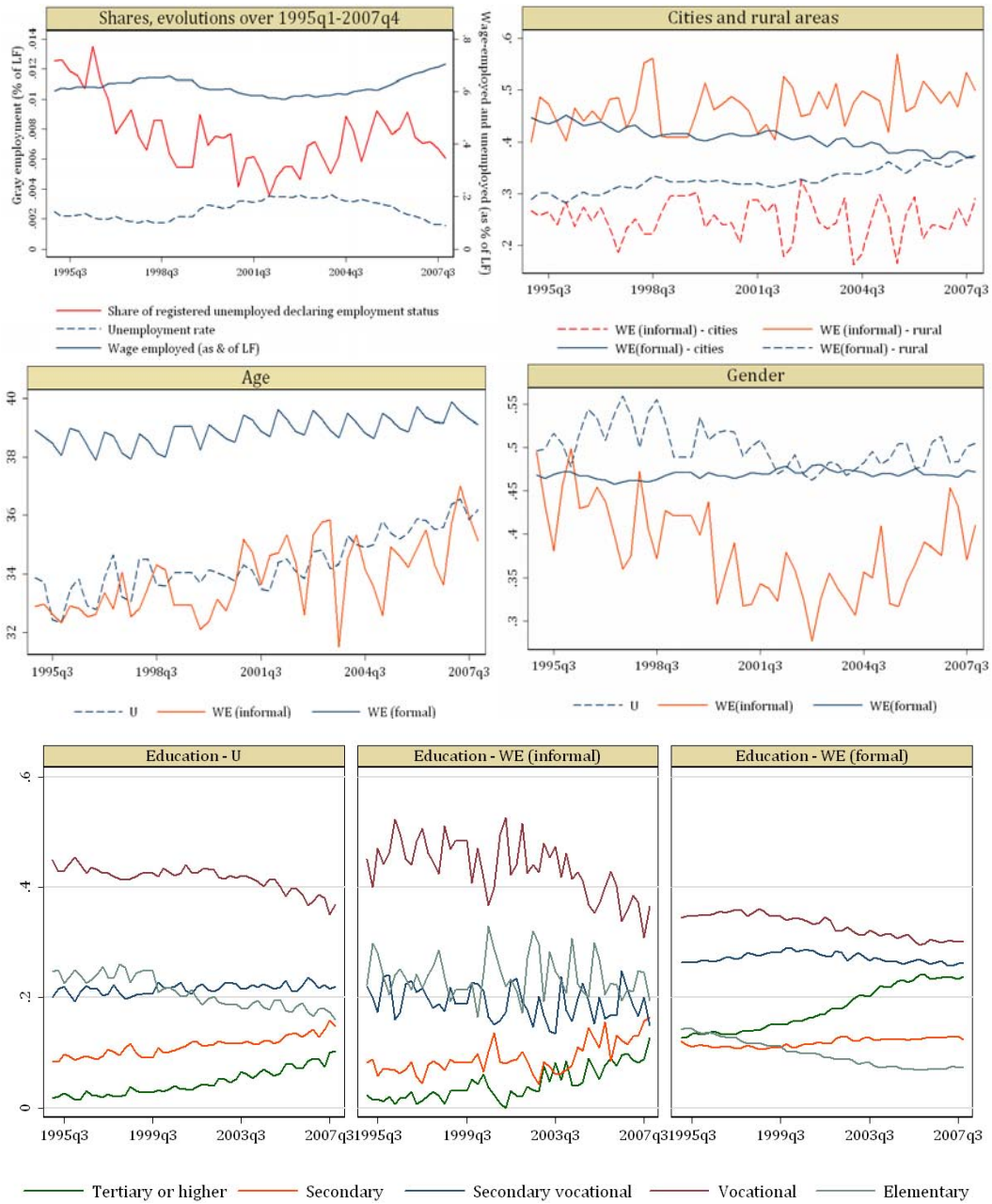


Figure 1: *Source:* Labour Force Surveys for 1995q1 - 2007q4. Notation: WE (informal) - wage employed in the informal sector, WE (formal) - wage employed in the formal sector, U - unemployed. Weights included in computing the averages.

employment as informal workers¹¹.

Along with the self-reported net wage revenues, interviewees report also the number of hours worked on this job. Here too, they may refuse to do so, while the number of individuals unable to provide an answer is extremely high in the context of this study - data for roughly 30% of the “grey” individuals is missing or inadequate¹². Consequently, relying on hourly wage would eliminate a large

¹¹Every two to four years Central Statistical Office performs an independent analysis of the shadow economy, with separate drawing scheme and survey. These data are not made available for research.

¹²Except for missing observations, also answers considerably exceeding the 40 hours per week norm are troublesome for the researcher. While the answer of 170 hours is probably relevant for monthly - not weekly - *status quo*, many of

part of the sample from the analysis.

In Poland, for all wage employed, social security and tax contributions are born roughly symmetrically by the employer and the employee. On the other hand, these liabilities are typically subtracted from the gross wage by the employer and paid to the authorities¹³. Thus, practically the whole effective tax burden lies on the employers. While it is not immediately obvious compare the net earnings in the formal and informal sector is adequate, with this institutional design of tax and social security contributions it seems grounded¹⁴.

Summarising, whenever we report wages in the reminder of this paper, we refer to self-reported net wage for the full-time employees (to permit comparability across datasets). Hourly wage is computed as a ratio between self-reported net wage and self-reported number of hours worked for both part-time and full-time employees across the whole analysed period of time¹⁵. We also test for the differences in the average number of hours worked by the officially and unofficially employed, without any corrections on the reported hour values.

3.3 Empirical strategy

Following the recent literature developments, we approach the issue of formal and informal compensations by the use propensity score matching. This technique allows evaluating to what extent “grey” workers (i) experience return to observable characteristics is different on formal or informal labour market or (ii) suffer from unequal treatment with respect to their individual characteristics by the employers. These methods provide proxies to verify whether shadow employment effectively increases the potential benefits for the individuals or posits the only available opportunity of guaranteeing wage income.

The chosen technique corresponds to certain *ex ante* beliefs about the nature of the non-randomness and unobserved heterogeneity. Namely, discrimination on the labour market can be defined to exist, if the actual mean earnings of members of specific groups are not identical to the mean which would be observed in a perfectly functioning labour market, without discrimination. In principle, if one applies Oaxaca (1973)-Blinder (1973) and Juhn, Murphy and Pierce (1993) decompositions on a Heckman (1979) corrected Mincerian wage equation should allow to find the *size and direction of discrimination conditional on the assumption that non-random selection is the only source of potential earnings differential*. Alternatively, the opposite *ex ante* belief about the nature of the unobserved heterogeneity may hold. Namely, if the *selection into “grey” economy was random conditionally on some individual characteristics*, decomposition techniques even with the Heckman (1979) correction would not necessarily bring about correct results. Consequently, one would not be able to use selection equation correction as reliable counterfactual in the second stage equation of the compensation, Rosenbaum and Rubin (1983), Heckman, Ichimura and Todd (1997), Heck-

the less extreme cases remain uninterpretable.

¹³Once the annual tax forms are submitted, if too much contributions were subtracted by the employer - revenues are returned to the workers.

¹⁴Part of the foregone social security contributions would have been transferred to the individualised pension accounts for individuals born after 1969, which is a consequence of the pension reform introduced in 1999. However, although there are not rigorous studies on the subject, it seems that the recognition of this fact is rather low, especially among lower skilled.

¹⁵For the pre-1997 datasets wages for only full time employees are available.

man, Ichimura and Todd (1998). Especially, if the at any stage the equation parameters are not statistically significant or robust - the decomposition can actually produce false findings.

The solution to this shortcoming of the parametric techniques is brought about by the propensity score matching. Relying on non-parametric estimates one is able to “create” the counterfactual, *i.e.* correct the calculations for the effect of choosing among the control group only those who “match” (are similar) the observed characteristics to the analysed group. By using the property of randomness within the matched groups, one is able to evaluate the average effect of a particular phenomenon (in this case: shadow employment wages) with respect to a reasonable benchmark. The quality of benchmark can actually be statistically verified *via* the post-matching balancing properties tests. Rosenbaum and Rubin (1983)¹⁶.

The critical element in propensity score matching lies in the conditional independence assumption construct¹⁷. With propensity score matching, the quality of estimation depends much on the data availability. In the case of this study, the pool for matching (the size of the control sample in the relation to the size of the analysed sample) is relatively large, so there is no need for imposing sampling with replacement. The algorithm will thus freely seek “twins”, maximising the resemblance between the treated and matched controls groups without additional constraints.

We apply kernel estimates of propensity scores with the kernel nearest neighbour matching, following Heckman, Ichimura, Smith and Todd (1998)¹⁸. We perform two types of analyses. The first one pools the observations from all quarterly datasets comparing and employs kernel nearest neighbour matching and additional stratification (Mahalonobis metric) based on the quarter of analysis. The stratification permits seeking the “statistical twins” in adjacent quarters, which is justifiable with datasets of such high frequency. As a consequence, statistical quality of matching may be enhanced. Such pooled results obviously do not permit time-dependant analysis, but give a broader view on the dominant character of the wage differential.

In the second approach - our preferred and elaborated more exhaustively in the empirical section - performs matching on a quarter-by-quarter basis. Here too kernel nearest neighbour matching is implemented. With the relative magnitudes of the samples (typically less than 500 individuals characterised as treated and over 20 000 pool of potential “twins”) the re-weighting instead of oversampling guarantees best statistical performance, Caliendo and Kopeinig (2008).

Although the set of variables is limited in this study, it comprises majority of the preferred determinants: age, gender, marital status, education, residence, occupation and industry are all accounted for (including the relevant interactions)¹⁹.

¹⁶Caliendo and Kopeinig (2008) discuss in detail recent development as well as guide through the process of adequate construct of this approach.

¹⁷In other words, for the reliability of the results it is important that the selection is solely based on observed characteristics and that all variables that influence belonging to the shadow economy and potential earnings are simultaneously observed. In practice it implies that there should be no other sources of systematic (i) selection and (ii) outcome.

¹⁸Alternatively, we could have used the oversampling technique. However, the choice of the over sampling magnitude is always arbitrary, while ten- or twenty-fold oversampling (as feasible in our sample) does not differ from the kernel approach in terms of statistical quality. Oversampling or sampling without replacement produced results qualitatively the same and quantitatively similar.

¹⁹We believe that relying on demographics (gender and age) as well as education and individual incentives (marital status) may be sufficient for the stability of propensity score matching approach and conformity with the conditional

- Age is a continuous variable expressing the age of individual at the moment of survey in years.
- Gender is coded to take the value of 1 for women.
- Education is a categorical variable with levels: elementary or lower; vocational, secondary vocational, secondary, tertiary or higher²⁰.
- Martial status has separate coding for singles (reference level), married, divorced/separated or widowed²¹.
- Residence is a categorical variable too²².

Since the rapid urbanisation associated with the brain drain is characteristic for the internal flows and educational patterns over the past two decades, we include additionally interaction terms for highly educated inhabitants of large cities and those who live in the rural areas and are characterised by elementary or lower education. We also include the interaction of gender with age and with tertiary education.

4 Results

This section reports the outcomes of the propensity score estimation applied to wages, hourly wages and number of hours worked for the formally and informally employed in Poland. We report the following results: (i) estimation properties and outcomes on pooled data; (ii) quarterly estimations for all observations; (iii) quarterly estimations taking into account potential heterogeneity along the wages distribution; (iv) quarterly estimations taking into account potential heterogeneity depending on physical barriers to participate in the labour market. The quarterly sample sizes and the synthetic balancing properties of all the estimations and are reported in the appendix²³.

Results reported in Table 1 confirm the initial conviction that actually the compensations of formally and informally employed differ. This assertion is confirmed by the statistical significance of the wages differential. Interestingly, when we compare pre and post-matching differential, the value is reduced by roughly a half. This suggests that approximately 50% of the net observed differential may be attributable to slightly lower endowments in terms of productivity determinants, while the remaining 50% follows from different compensating schemes in the formal and informal sectors. It also seems that informally employed work slightly shorter hours their formally employed counterparts. Here, the size of the differential changes between pre- and post-matching, which may suggest that some of the informal jobs are concentrated in general in occupations where working times are longer, but the informally employed tend to work less.

independence assumption. We verify this approach empirically by the use of t-tests, as suggested by Rosenbaum and Rubin (1983).

²⁰Some of the respective surveys comprised more categories, but for the sake of consistency between the quarters this is the most detailed categorisation.

²¹LFS does not contain information on whether an interviewee has children and/or other dependants.

²²The strata include: rural areas, towns under 2 thou inhabitants, towns under 5 thou inhabitants, towns under 10 thou inhabitants, towns under 20 thou inhabitants, towns under 50 thou inhabitants, cities under 100 thou inhabitants, cities above 100 thou inhabitants.

²³Detailed results for all 260 estimations available upon request. In the Appendix we include Figure 5 for referencing the fluctuations in GDP and unemployment flows.

Table 1: Propensity score estimation - pooled data

| Variable | Coefficient | Standard error | Z-statistic | Balancing |
|--|-------------|----------------|-------------|------------------|
| Gender | 0.66 | 0.055 | 12.01 | Yes |
| Marital status dummies (three and base) | Significant | | | Yes |
| Educational dummies (four and base) | Significant | | | Yes |
| Occupational dummies (eight and base) | Significant | | | All but one cat. |
| Residence dummies (eight and base) | Significant | | | Yes |
| Age | -0.025 | 0.0042 | -6.00 | Yes |
| Age squared | 0.0018 | 0.000054 | 3.33 | Yes |
| Education x gender | -0.052 | 0.012 | -4.44 | Yes |
| Age x gender | -0.014 | 0.00012 | -12.01 | Yes |
| Large city x tertiary education | 0.083 | 0.056 | 1.49 | No |
| Rural areas x primary or lower education | -0.049 | 0.025 | -1.99 | Yes |
| Industry dummies (sixteen and base) | Significant | | | Yes |
| Constant | -1.27 | 0.11 | -11.68 | |
| No of observations | 530 624 | | | |
| No of observations on support | 525,183 | | | |
| LR χ^2 (45) | 8 243.54*** | | | |

| Before matching | | | | |
|-----------------|------------|----------------|-----------------|--------------|
| Variable | Difference | Standard error | t - statistic | Significance |
| Wage | -306.34 | 7.10 | -43.11 | *** |
| Hours | 0.33 | 0.107 | 3.06 | *** |
| Hourly wage | -7.85 | 0.21 | -36.77 | *** |

| After matching | | | | |
|----------------|------------|----------------|-----------------|--------------|
| Variable | Difference | Standard error | t - statistic | Significance |
| Wage | -140.22 | 4.92 | -28.47 | *** |
| Hours | -0.79 | 0.14 | -5.63 | *** |
| Hourly wage | -2.87 | 0.14 | -20.25 | *** |

Notes: Results of a probit estimation, log likelihood = -26 211.513. Kernel density matching with Mahalonobis metric stratification along time. Standard estimators after matching do not take into account that the propensity score is estimated. Balancing properties signify the post-matching test for the equality of means - matching algorithm is considered balanced if post-matching differences are insignificant. Category 2 in occupation definition denotes specialists and middle management.

Nonetheless, we are reluctant to interpret these findings. Namely, the nominal wages have grown considerably over the analysed horizon. Also the characteristics of both working populations have undergone many structural changes, as depicted in Figure 1. Consequently, although the stratification made sure that only individuals from adjacent quarters were matched (*i.e.* matching quality is good), the mean computed on this sample is meaningless for wages due to the non-stationary character of this variable. Thus, we move to analysing the time evolution of the wage and hours differentials. Each of the figures depicts as continuous line (right scale) the evolution for the informal sector - be it wages, hourly wages or the number of hours worked. The left scale reports the size of the differential (expressed as % of the mean value recorded on the right axis). These differentials are only report if they were statistically significant. Red dots signify post-matching differential and green stand for the pre-matching ones, both as % of the mean in the informal sector. Lack of dot for a particular point in time is equivalent to the insignificance of the differential²⁴.

Analysing the graphs the attention will be focused on two main factors. First, the evolutions of differentials will be viewed through the prism of cyclical and unemployment fluctuations. Secondly,

²⁴We have universally imposed 5% threshold for evaluating the significance of the pre- and post-matching differentials.

the moments of institutional changes (*e.g.* revisions of the minimum wage or social assistance benefits) will be identified²⁵.

Overall, the findings point to a procyclical behaviour of the wage and the hourly wage in informal economy, Figure 4. This would be in line with the more recent approach in literature, Fiess et al. (2007), as discussed earlier in the paper. It is hard to judge on the cyclical behaviour of hours worked in shadow economy: they sharply decline during the economic downturn of 2001 but later move around an average of 39 hours. The differential in wages between formally and informally employed is negative before matching - it seems that “grey” workers earn less. However, after matching the differential turns out to be positive and points to consistently higher wages for informally employed as compared to the statistical “twins” in the formal economy. Also, roughly the same relative size of the differential should be noted: it is around 40-60% of the informal wage before and after matching. Similar conclusions can be drawn for hourly wages: the difference before matching is negative and of different size than the difference after matching which is in favour of the informally employed and of fairly stable relative magnitude over the entire horizon of the analysis.

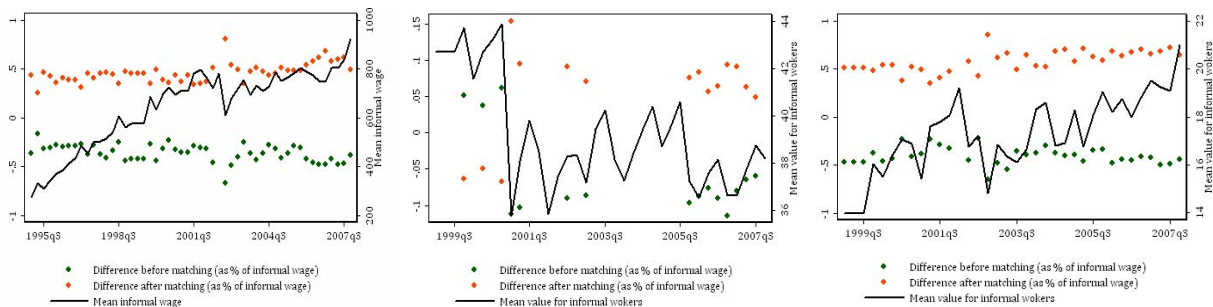


Figure 2: Results of PSM, total sample, 1995q1-2007q4

Since the relative size of the differential is stable, the cyclical behaviour of the differential is the same as that of wages or hourly wages - it tends to be procyclical as wages, but the growth rates is the roughly the same as that of compensations. An interpretation of the differential for hours is rather impossible due to the fact of insignificance of a large part of the differentials. However it should be noted that in the period before the 3rd quarter of 2001 the differential points to a larger amount of hours for the informally employed whereas the differential after matching shows less hours for them. This phenomenon is reversed after the 3rd quarter 2001: informally employed work more hours than their statistical twins.

The procyclical evolution of the differential for matched individuals over time points to some important conclusions about the nature of compensation schemes in formal and informal economy. Namely, it seems that larger increase in formal wages creates room for raising also the informal compensations *contemporaneously*. Only large, unexpected swings seem to be associated with the growing discrepancies (the differential increases during downturns - *i.e.* 3rd quarter of 2000, 2nd

²⁵Regarding the dependence between informal employment and minimum wage Loayza (1994) motivated by the behaviour of informal sector in Latin American countries constructs a model for an isolated urban economy showing that a too high minimum wage will cause all firms to be informal. Therefore a minimum wage above market levels will induce the formation of an informal sector. However, Doaldo, Karamarz, Machin, Manning, Margolis and Teulings (1996) underline that the effects (good and bad) of minimum wages are exaggerated. Empirical analysis for some European countries shows mixed evidence of the effect of minimum wage on employment: higher minimum wages reduce employment in some cases (particularly for young workers) and raise it in others (particularly for total employment).

quarter of 2002 - and decreases slightly with the economy expansion, as of *i.e.* 2005q2).

The patterns of informal wages and differential evolutions seem to be rather associated with the GDP fluctuations than with the unemployment movement. The lowering of GDP growth rates covered the period 2000-2002 whereas the the labour market contraction lasted from 1999 to 2005, Figure 5 in the Appendix. While this finding may seem puzzling, since we analyse informal employment - the analysis focuses on compensations patterns and not the *size* of unregistered employment. Thus, the results are not at odds with the expectations posed by both strands of theoretical research²⁶. While this is not a major concern for our research as we focus on wages in informal economy and not on the size of it - the finding suggests that the population used in this study definitely does not exhaust quantitatively the employment in the shadow economy.

Regarding the influence of minimum wage on informal wages it is hard to see any effects for the entire analysis horizon. There are two explanations for this finding. First, the increments in minimum wage were rather small ranging from about 5% to 10% nominally from year to year between 1996 and 2007. Therefore they had negligible effect on the wages in general, including the ones in the informal sector²⁷. The other explanation stems from the fact that the minimum wage has not been binding over the analysed horizon in the formal sector²⁸.

Finally, as far as the institutional framework is concerned, there have been considerable changes in the labour code regulations regarding subcontracting and compliance with employment regulations. Namely, it was said to be frequent that the employers would “push” employees into self-employment as subcontracting implied lower tax wedge and social security contributions than a standard employment agreement of the same net value. As of 2006 the changes in the legislation increased the punishment for the employers for such practices and controls were intensified. There were also minor changes in the legislation concerning the unemployment benefits and eligibility conditions. However, from observing the evolutions for the whole population it seems that none of the hikes or sudden drops in informal wages can be pinned down to these changes. Thus, we move further to inquire if there could be some distributional effects, whose occurrence could be blurred when analysing the whole population.

To address the potential heterogeneity in the behaviour of outcome variables along the wages distribution, we split the sample with reference to median wages, Figure 4. Estimation procedure

²⁶The size of informal employment - as defined in this paper - in the labour force has been presented in Figure 1. However, we would be reluctant to derive absolute instead of relative measures from this data. The weighting scheme applied by the Central Statistical Office is reliable for the whole survey, roughly 50 thou individuals. “Grey” workers, as identified in this study, amount to roughly 500 individuals in each sample, which makes the computation of aggregate stocks questionable. Please note that the shortcomings of CSO weights for this small population do not undermine the validity of propensity score matching. Namely, PSM in the version applied produces it’s own weights for the treatment and the comparison groups. These weights do not serve the purpose of assuring objective representativeness, but the representativeness within the control group with respect to the treatment group.

²⁷A large increase in the minimal wage by 20% was announced in mid 2007 to take place from the beginning of 2008. This announcement could have a strong influence on informal wages which is clearly visible in the data: a large increase in informal wage is observed for the 3rd and 4th quarter of 2007. However this period was also marked by visible increase in formal wages which can also affect the informal wage.

²⁸Many compensations and compensation limits are benchmarked to the minimum wage, which makes it a relevant reference point for the employers, employees and the government. However, in our study this indirect effect is non-detectable, as it concerns contributions of the employers to the employees social fund or the limits on pecuniary and nonpecuniary benefits for the employees for the tax purposes.

was the same, but matching was performed separately for “grey” earnings above and below median. The obtained results point to two conclusions: (i) below median differences are more frequently statistically significant than the above median differences; (ii) while the above median relative differential seems to be stable across time, for workers with lower compensations there is a clear increase - firstly occurring in the period of GDP and labour market contraction, but subsequently picking up as of 2005q3, which is a period of both relatively high GDP growth and gradual labour market improvement.

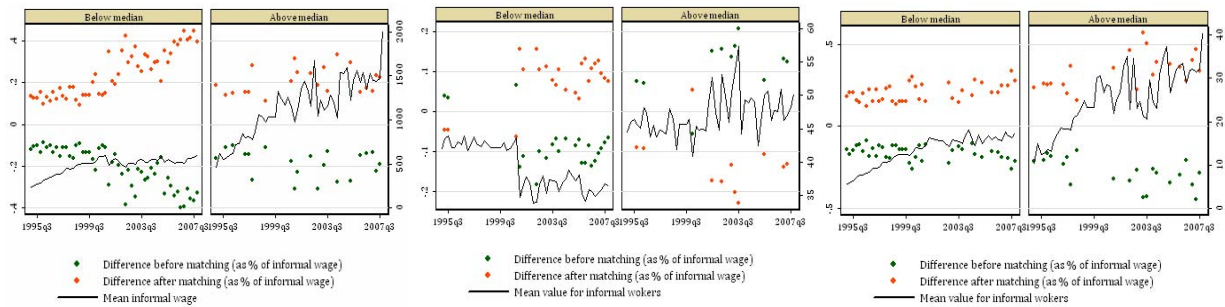


Figure 3: Results of PSM, separate matching for above and below the median for the informal workers, 1995q1-2007q4

Conclusions are different for the hourly wages. Namely, below median relative differential is stable across cycle. Note also, that the period of most rapid unemployment growth is associated with insignificance of the differentials. While this effect may be a statistical artefact, it may also indicate that massive unemployment increases raise demand for informal employment from the side of employees, which pulls down the wages offered by the employers. On the other hand, the above median hourly compensation show high variability and suggest a weak but procyclical behaviour of the post-matching relative differential. In this sense, below median earnings behave similarly to the whole sample, which is depicted in the right panels of Figure 4 and 4. However, the above median earnings show a puzzling story of both higher variability and cyclically growing differentials.

Hours worked for the below subsample exhibit similar behaviour as for the total sample: they sharply decline during the downturn of 2001 and latter move around an average of about 37 hours. For the above median subsample the number of hours worked is much larger in absolute values exceeding over 50 hours in some periods. It also behaves differently in time pointing to a rather countercyclical behaviour in some periods. An interesting result is obtained for the numbers of hours worked in the above median subsample: informally employed work less than their statistical twins in the formal economy contrary to the below the median and the whole sample²⁹.

The distribution of wages was considered with the rationale that maybe the effects are different for the lower and upper bottom of the distribution. This rationale was confirmed, but the reasons for such findings may in fact again be dual: labour market segmentation or strong effect of asymmetric tax wedge. Thus, we have selected to ask the same question but splitting the sample in a way that could allow to somehow capture potential access barriers. We have used the size of residence with the following rationale: in larger metropolies the costs associated with changing employment

²⁹Recall that the information on hours worked is missing or inconsistent in roughly 25-30% of the cases, which undermines the statistical quality of these findings. For this reason, we restrain from providing economic interpretation of this result.

are lower and seeking a formal alternative for an informal offer from a local employer is easier. In smaller towns, locally choice is more constrained while the costs of moving to another town may in fact be considerable. Thus, we have identified large cities (above 100 thou inhabitants) and small town (below 10 inhabitants) with rural areas as two subsamples in the analysis³⁰.

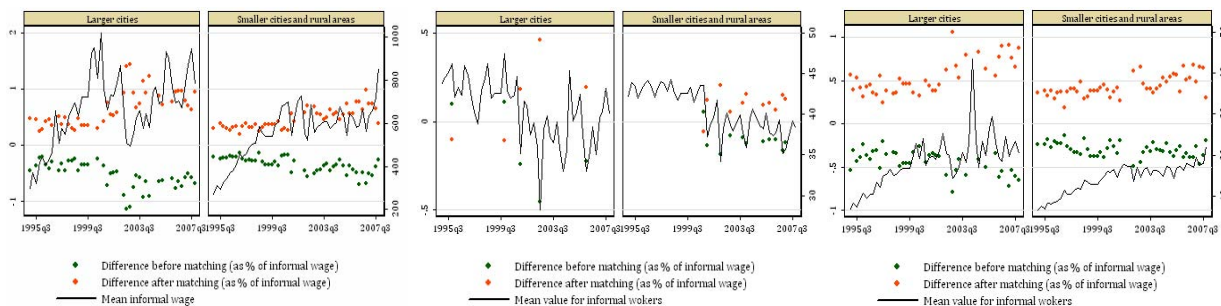


Figure 4: Results of PSM, separate matching for inhabitants of large cities *versus* smaller towns and rural areas, 1995q1-2007q4. *Note:* only inhabitants of the cities above 100 thou and towns below 10 thou included in the analysis.

Indeed, in metropolies wages and hourly wages for informally employed behave strongly procyclically, whereas in the smaller residences they are considerably lower and tend to have weakly acyclical patterns, Figure 4. On the other hand, the differentials are more variable and also more cyclical for the large cities, though the effect seems to be shifted in time from the period of GDP slowdown towards unemployment hike³¹. The sharp decline in the numbers of hours worked during the downturn in 2001 is similar to the cases discussed above, but the decreasing trend for smaller towns and rural areas is a novelty, suggesting lowering labour supply in the informal economy for this particular subsample. The results for the smaller towns and rural areas seem to suggest also a divergence in the size of pre- and post-matching relative differentials for the hourly wages. Informally employed when compared to the statistical twins from the formal economy - earn more and this tendency grows over time³². The differentials seem to show no cyclical properties in the peripheral areas.

The results of our analyses indicate consistently higher wages for informally employed when compared to the formally employed “statistical twins” for the whole sample as well as for various subsamples. The consistently higher wage for informally employed can serve as evidence for the tax evasion hypothesis: individuals engage into shadow economy because they receive higher earnings

³⁰It should be noted that - unlike the analysis along the wage distribution - we are not “splitting” the sample in this analysis, but selecting two subsamples, while part of the observations (neither metropolies nor smaller locations) was not included in the analyses. Accordingly, roughly 15-25% of the sample was not used in this analysis.

³¹This effect is an outcome of the distinguishing characteristics of informal employment - as defined in this paper - in the metropolies. Namely, the two types of informal employment dominate in the metropolies. Individuals officially unemployed in periphery locations temporarily come to metropolies seeking the opportunities to earn income in (i) the construction sector (especially renovations and individual constructions); and (ii) providing household services. While the demand for latter will be lower in the periods of high unemployment, the supply of the former should intensify over the economic slowdown, thus suppressing the wages in the initial stage. Since unemployment typically rebounds sooner in the metropolies, the demand for both types of services starts growing faster than the observable improvement in the unemployment rate.

³²As in previous cases, due to a large number of insignificant differentials it is hard to draw conclusion for the number of hours worked.

when compared to formal employment. This result can serve as an advice for policy makers to focus on a reduction in taxes and social security contributions when facing the problem of a reduction in shadow economy.

The results also point to the procyclical behaviour of informal wages which is in line with the recent literature, Bovi (2006), Busato and Chiarini (2004), or Fiess et al. (2007). However, the procyclical properties of the wages seem to suggest that the informal sector co-exists with the formal economy, responding similarly to GDP fluctuations. However, some distributional effects - mainly concerning the workers with below median wages and those inhabiting the more peripheral areas - suggest that in-the-shadow employment may not necessarily be a matter of individual choice, but rather the only available alternative.

5 Conclusions

Individuals without employment are in a very difficult situation - especially in transition countries, where benefit systems are of less generosity while activation policies of lower coverage than in the mature market economies, especially in Europe. The main question we attempt to address answer the question whether workers in-the-shadow would potentially earn higher after-tax income in formal sector (thus they participate in the informal sector for the lack of better alternatives) or would it be lower (suggesting tax evasion reasons³³). We use 52 consecutive labour force surveys for Poland covering the period of 1995q1-2007q4 and apply propensity score matching to obtain reliable estimates of the counterfactual formal wages for the “grey” employees. To the best of our knowledge, both the application of the propensity score matching and such use of large consecutive data sets is unique in the literature of the field.

The feasible identification of the unregistered employees narrowed the definition of the shadow employment in this paper to those who officially report unemployment but have a constant job (as defined by ILO) and earned wage income. We could not include the self-employed in the analysis as the data sets report no data on income for the self-employed. Neither could we comprise in the estimations officially inactive individuals who would nonetheless work informally. Consequently, we analyse a specific but at the same time crucial sub-group of the unregistered employees.

We find that in fact “grey” workers are paid more than the “statistical twins”. This finding is not surprising, since in terms of education, age and gender the “grey” workers are fairly similar to the population of the unemployed, which suggests that their employment odds are relatively lower in general, while tax wedge relatively higher. Consequently, labour demand curve may in fact be more elastic in the bottom of the distribution, making tax evasion opportunities very powerful incentives encouraging the avoidance of formalising employment contract. On the other hand, when distributional effects were explicitly accounted for, over the lower half of the wages distribution we find evidence that less earning workers located in smaller towns or rural areas may in fact be informally employed for the lack of alternatives.

Analysing the cyclical properties of the wages paid by the shadow economy as well as the differentials between unmatched and matched formal compensations, we find that procyclical behaviour

³³Please note, that this is equivalent to individual with certain before tax earnings potential in both sectors choosing the informal sector because of higher after tax earnings.

of wages and hourly wages in shadow economy keeps fairly the same pace as compensations in the formal sector. We identify important exceptions of weak acyclicity and interpret them as evidence of entry barriers. The evidence concerning the number of hours worked is rather mixed but the differentials - if significant - tend to be small indicating a slightly larger number of hours worked by the “grey” employees.

These results can underpin some important policy implications concerning the role of tax incentives in pushing those willing to work into shadow. The differential in the wage of informally employed and the statistical “twins” is quite large (roughly 40% to 60%), but in fact its magnitude coincides with the effective tax and social security contributions burden that would have been born by the employers, had they formalised and registered these contracts.

There are some interpretational limitations of this study, though. Firstly, LFS is not a reliable source for estimating the absolute size of informal employment. Thus, we are not analysing if the GDP or unemployment fluctuations have affected significantly the quantity of labour supplied and demanded informally. Secondly, the profound nature and the long-term effects of informality exhibits through the trajectories of the workers and the self-employed. Such analyses are rarely feasible for the reasons of data availability, but studies like this should be treated with adequate caution.

References

- Bednarski, M.: 1992, *Drugi obieg gospodarczy – Przeslanki, mechanizmy i skutki w Polsce lat osiemdziesiątych*, Wydawnictwa Uniwersytetu Warszawskiego, Warsaw.
- Bednarski, M., Kokoszczyński, R. and Stopyra, J.: 1988, Kształtowanie się rozmiarów drugiego obiegu w Polsce w latach 1977-1986, *Bank i Kredyt* (8-9), 11–18.
- Bennett, J. and Estrin, S.: 2007, Informality as a Stepping Stone: Entrepreneurial Entry in a Developing Economy, *IZA Discussion Papers 2950*, Institute for the Study of Labor (IZA).
- Blinder, A. S.: 1973, Wage Discrimination: Reduced Form and Structural Estimates, *Journal of Human Resources* **18**(4), 436–455.
- Boeri, T. and Garibaldi, P.: 2002, Shadow Activity and Unemployment in a Depressed Labor Market, *CEPR Discussion Papers 3433*, CEPR.
- Bosch, M., Goni, E. and Maloney, W.: 2007, The Determinants of Rising Informality in Brazil; Evidence from Gross Worker Flows, *IZA Discussion Papers 2970*, Institute for the Study of Labor (IZA).
- Bosch, M. and Maloney, W.: 2006, Gross Worker Flows in the Presence of Informal Labor Markets. Evidence from Mexico, 1987-2002, *Technical report*, World Bank.
- Bosch, M. and Maloney, W.: 2007, Comparative Analysis of Labor Market Dynamics Using Markov Processes: An Application to Informality, *IZA Discussion Papers 3038*, Institute for the Study of Labor (IZA).
- Bosch, M. and Maloney, W.: 2008, Cyclical Movements in Unemployment and Informality in Developing Countries, *IZA Discussion Papers 3514*, Institute for the Study of Labor (IZA).
- Bovi, M.: 2006, The Cyclical Behavior of Shadow and Regular Employment, *Working Paper 62*, ISAE.
- Busato, F. and Chiarini, B.: 2004, Market and underground activities in a two-sector dynamic equilibrium model, *Economic Theory* (23), 831–861.
- Caliendo, M. and Kopeinig, S.: 2008, Some Practical Guidance For The Implementation of Propensity Score Matching, *Journal of Economic Surveys* **22**(1), 31–72.
- Carillo, M. R. and Pugno, M.: 2004, The underground economy and underdevelopment, *Economic Systems* (28), 257–279.
- Cassel, D., Jaworski, W., Kath, D., Kierczyński, T., Lutkowski, K. and Paffenholz, H.-J.: 1989, *Inflation und Schattenwirtschaft im Sozialismus*, S+W Steuer- und Wirtschaftsverlag, Hamburg.
- Chiarini, B. and Marzano, E.: 2007, Structural and Cyclical Patterns of Underground Labour Input in Italy from 1980 to 2004, *Discussion Papers 3/2007*, D.E.S. (Department of Economic Studies), University of Naples “Parthenope”, Italy.

- Doaldo, J., Karamarz, F., Machin, S., Manning, A., Margolis, D. and Teulings, C.: 1996, The economic impact of minimum wages in Europe, *Economic Policy* (3), 319–372.
- Fiess, N. M., Fugazza, M. and Maloney, W. F.: 2007, Informal Labor Markets and Macroeconomic Fluctuations, *Proceedings of the German Development Economics Conference, Göttingen 2007 6*, Verein für Socialpolitik, Research Committee Development Economics.
- Frey, B. and Weck-Hanneman, H.: 1984, The Hidden Economy As An “Unobserved” Variable, *European Economic Review* **26**(1-2), 33–53.
- Friedman, E., Johnson, S., Kaufmann, D. and Zoido-Lobaton, P.: 2000, Dodging the Grabbing Hand: the Determinants of Unofficial Activity in 69 Countries, *Journal of Public Economics* **76**(3), 459–493.
- Fugazza, M. and Jacques, J.-F.: 2004, Labor Market Institutions, Taxation And The Underground Economy, *Journal of Public Economics* **88**(1-2), 395–418.
- Gerxhani, K.: 2003, Tax Evasion in Transition: Outcome or Institutional Clash? - Testing Feige’s Conjecture, *Aias working paper*, Amsterdam Institute for Advanced Labour Studies.
- Gutmann, P.: 1977, The Subterranean Economy, *Financial Analysts Journal* **33**(6), 24–27.
- Heckman, J., Ichimura, H., Smith, J. and Todd, P.: 1998, Characterizing Selection Bias Using Experimental Data, *Econometrica* **66**(5), 1017–1098.
- Heckman, J., Ichimura, H. and Todd, P.: 1997, Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme, *Review of Economic Studies* **64**, 605–654.
- Heckman, J., Ichimura, H. and Todd, P.: 1998, Matching as an Econometric Evaluation Estimator, *Review of Economic Studies* (65), 261–294.
- Heckman, J. J.: 1979, Sample Selection Bias as a Specification Error, *Econometrica* **47**(1), 153–61.
- Ihrig, J. and Moe, K. S.: 2001, Tax Policies and Informal Employment: The Asian Experience, *Asian Economic Journal* **15**(41), 369–383.
- Johnson, S., Kaufmann, D. and Zoido-Lobaton, P.: 1998, Corruption, Public Finances and the Unofficial Economy, *Technical report*, World Bank.
- Juhn, C., Murphy, K. and Pierce, B.: 1993, Wage Inequality and the Rise in Returns to Skill, *Journal of Political Economy* **101**, 410–442.
- Kolm, A.-S. and Larsen, B.: 2003, Wages, Unemployment, and the Underground Economy, *CEsinfo Working Paper 1086*, CESinfo.
- Kopczuk, W.: 2001, Redistribution When Avoidance Behavior Is Heterogeneous, *Journal of Public Economics* **81**(1), 51–71.
- Kriz, K., Merikull, J., Paulus, A. and Staehr, K.: 2008, Why Do Individuals Avoid Payroll and Income Taxation in Estonia, in M. Pickhardt and E. Shinnick (eds), *The Shadow Economy, Corruption and Governance*, INFER Advances in Economic Research, Edward Elgar.

- Krynska, E., Bednarski, M., Walewski, M. and Pater, K.: 2007, Raport z badania “Przyczyny pracy nierejestrowanej, jej skala, charakter i skutki społeczne” (Report from a survey “Causes of unregistered employment, scale character and social implications”) , *Technical report*, Ministry of Labour and Social Affairs, Center for Socio-Economic Research.
- Loayza, N.: 1994, Labor regulations and the informal economy, *Technical report*, World Bank.
- Loayza, N. V. and Rigolini, J.: 2006, Informality Trends and Cycles, *Technical report*, World Bank.
- Maloney, W. F.: 2004, Informality Revisited, *World Development* **32**(7), 1159–1178.
- Oaxaca, R.: 1973, Male-Female Wage Differentials in Urban Labor Markets, *International Economic Review* **14**(3), 693–709.
- Renooy, P., Ivarsson, S., van der Wusten-Gritsai, O. and Meijer, R.: 2004, Undeclared Work in an Enlarged Union: An In-Depth Study of Specific Items, *Final report*, European Commission, Directorate-General for Employment and Social Affairs, Brussels.
- Roca, J. C. C., Moreno, C. D. and Sanchez, J. E. G.: 2001, Underground economy and aggregate fluctuations, *Spanish Economic Review* (3), 41–53.
- Rosenbaum, P. R. and Rubin, D. B.: 1983, The Central Role of the Propensity Score in Observational Studies for Causal Effects, *Biometrika* **70**(1), 41–55.
- Schneider, F.: 2007, Shadow Economies and Corruption all over the World: New Estimates for 145 Countries, *Technical report*, University of Linz.
- Schneider, F. and Enste, D. H.: 2000, Shadow Economies: Size, Causes and Consequences, *Journal of Economic Literature* **38**(1), 77–114.
- Tanzi, V.: 1983, The Underground Economy in the United States: Annual Estimates, 1930-1980, *Staff papers*, IMF.

6 Appendix

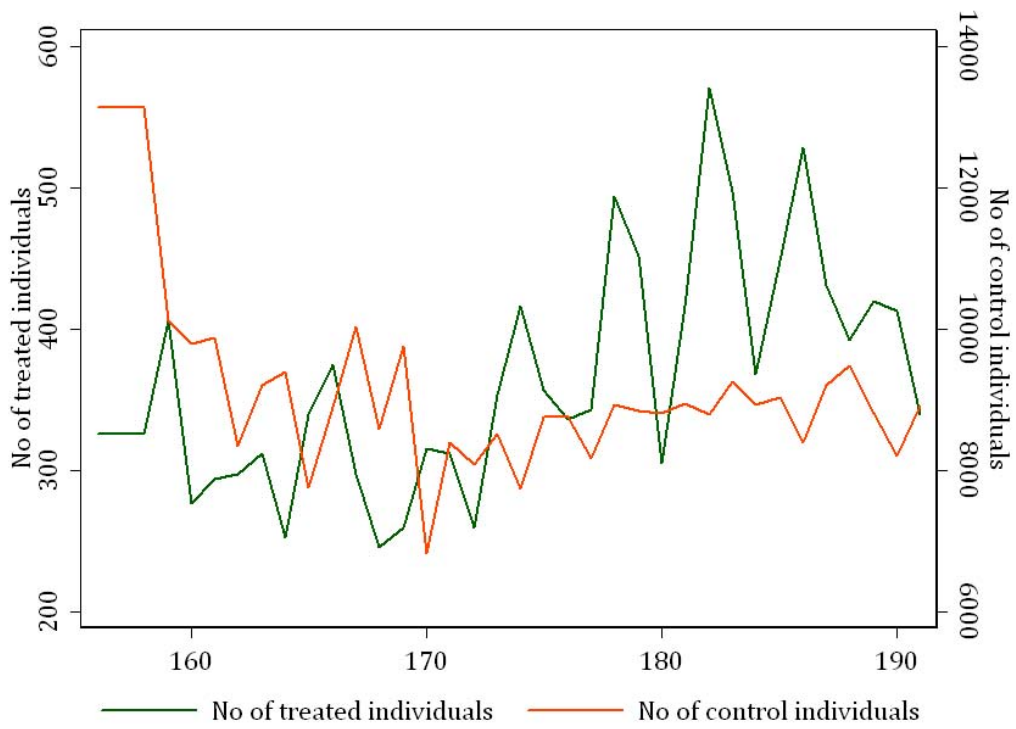


Figure 5: Sample size, 1995q1-2007q4

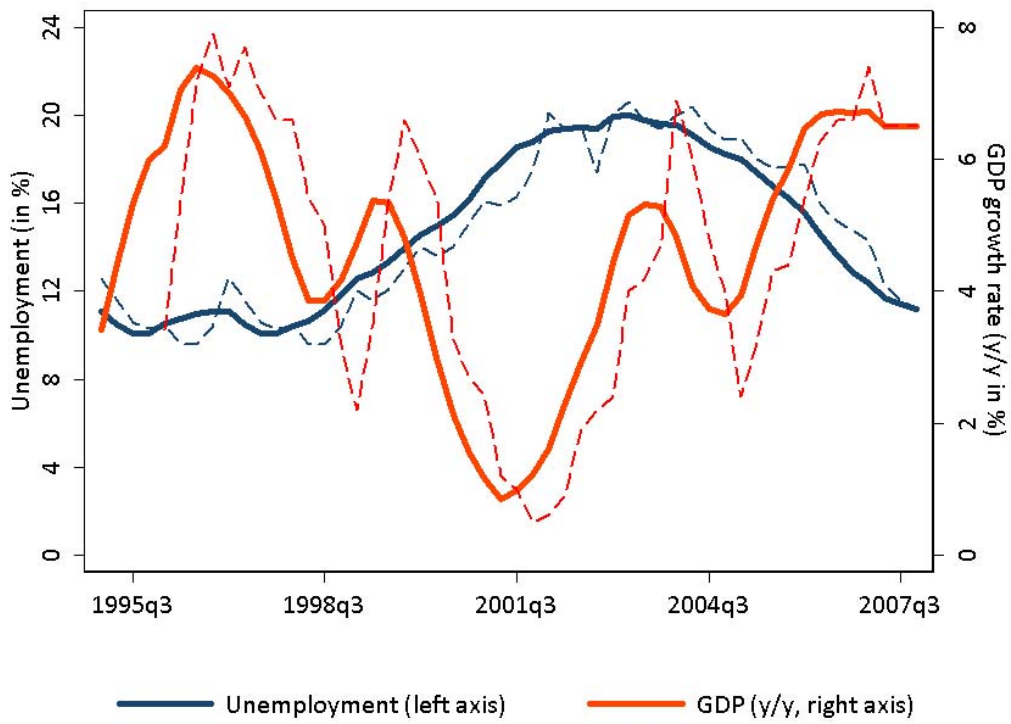


Figure 6: Business cycle in Poland, 1995q1-2007q4

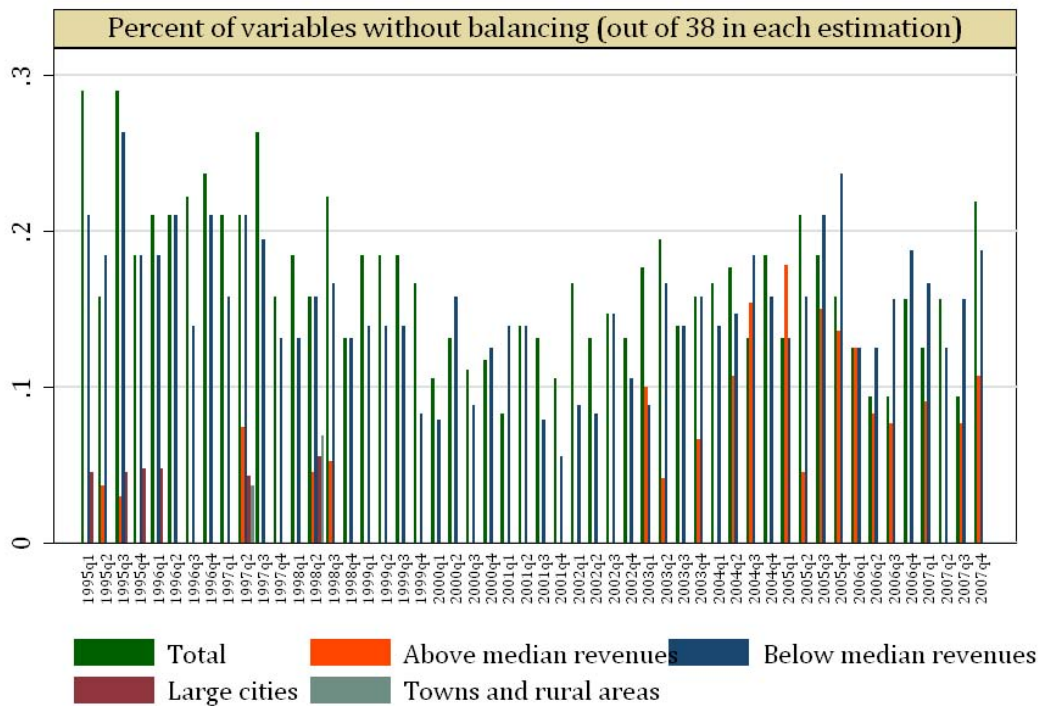


Figure 7: Balancing properties after matching for respective estimations, 1995q1-2007q4. *Notes:* In each estimation of the propensity score 38 variables were included. In some cases - percentage depicted above - the post-matching test for means equality reported significance rejection of the null hypothesis. Visibly, only total sample and above median sample can be considered troublesome, while even in these cases the number of variables vis-a-vis which the balancing was not obtained is as low as 4 to 8 variables in the majority of the quarters. Detailed logs available upon request.