Accepted and Forthcoming in
Socio-Economic Planning Sciences

Measuring participation in undeclared work in Europe using survey data: a method for resolving social desirability bias

Maria Felice Arezzo[a*], Ioana A. Horodnic[b], Colin Williams[c] and Giuseppina Guagnano[d]

[a] *Department of Methods and Models for Economics, Territory and Finance, Sapienza university of Rome, Rome, Italy, mariafelice.arezzo@uniroma1.it ; [b] Faculty of Economics and Business Administration, Alexandru Ioan Cuza University of Iași, Iași, Romania, ioana.horodnic@gmail.com; [c] Management School, University of Sheffield, Sheffield, United Kingdom, C.C.Williams@sheffield.ac.uk; [d] Department of Methods and Models for Economics, Territory and Finance, Sapienza university of Rome, Rome, Italy, giuseppina.guagnano@uniroma1.it*

* Corresponding author. Sapienza University of Rome, Department Memotef, Via del Castro Laurenziano, 9 – 00161 Roma; mariafelice.arezzo@uniroma1.it

Abstract

Measuring participation in undeclared work using surveys has been criticized for under-estimating the level of engagement due to social desirability bias that leads to an under-reporting of "bad" behavior. Until now, few studies have sought to quantify the amplitude of this bias in surveys of undeclared work. The aim of this paper is to fill this gap by using the most appropriate methodologies for estimating the probability of misleading responses in such surveys. Reporting data from special Eurobarometer survey no. 498 conducted in 2019 and involving 27,565 respondents in EU-27 countries and the UK, only 3.5 per cent openly admitted to participating in undeclared work. The results of a Probit model with correction for misclassified cases (i.e., those undertaking undeclared work but declaring that they do not) reveals that nearly a quarter (23.3 per cent) of the respondents undertaking undeclared work refused to openly admit this during the survey, due to the social desirability bias. The estimated overall proportion of undeclared workers is 17.3 per cent. We obtained this value by correcting for both misclassification and the additional source of negative bias due to the large imbalance in the data (i.e., observations in one class are much lower than the other). The outcome of this new advanced approach in analysing undeclared work is that survey estimates can now report its size and determinants in a more accurate manner than has been previously the case.

Keywords: word; undeclared work, shadow economy, social desirability bias, binary choice models, misclassification

1. Introduction

Despite the expectation in the twentieth century that undeclared work would vanish as economies modernized, this has not been the case. It persists globally and it is even the case that the proportion of the workers who have their main employment in the informal economy has increased in the past two decades or so (ILO, 2020). The COVID-19 pandemic has further compounded this trend. Many companies operate in "survival mode", destabilizing national economies further (European Commission, 2020, p.1). Evaluating the extent of participation in undeclared work is essential. It helps in understanding its impact on national economies and in designing effective policy measures to address this issue. However, estimating the extent of undeclared work is methodologically challenging because of its hidden nature. It is also challenging methodologically to disaggregate undeclared work from criminal activities that together constitute the shadow economy. Indeed, the shadow economy is the focus of many of the estimation methods developed so far. The consensus in both the academic literature and policy community is that undeclared work covers all the paid activities which are legal in nature but hidden from the state agencies (e.g., tax authorities, labour law authorities and social security agencies) for the purpose of evading the legal contributions due (European Commission, 2007; OECD, 2012; Williams, 2014). The shadow economy, on the other hand, covers all hidden economic activities, including the criminal economy (e.g., drugs, prostitution, gambling; smuggling and other illegal activities) in addition to undeclared work and all other forms of tax evasion (Schneider and Buehn, 2018; Schneider and Enste, 2000).

Previous literature concerned with measuring the size of undeclared work can be divided into two lines of research, namely indirect and direct methods of measurement. Both groups have their specific strengths and shortcomings, and currently, most methods have limited applicability for capturing the size of undeclared work.

Starting with the indirect methods, the essence of these methods is that they assume that observable phenomena are signs of the hidden part of the economy. The most prominent include the discrepancy methods, the monetary transaction approach, the currency demand approach, the electricity consumption method, and MIMIC estimation procedure (Burgstaller et al., 2022; Feld and Larsen, 2012; Schneider, 2021; Schneider and Buehn, 2018). However, using these methods for estimating the size of undeclared work is problematic. These methods capture the size of shadow economy (including all forms of tax evasion as well as the criminal activity) and cannot capture the share which is undeclared work. In addition, one of the most important shortcomings of the indirect methods is that they tend to over-estimate the size of shadow economy (Burgstaller et al., 2022).

Direct methods rely on surveys, experiments or on tax audits (Burgstaller et al., 2022; Doerr et al., 2022; Schneider and Buehn, 2018) and are, therefore, based on information directly provided by the population. These methods specifically focus on undeclared work and can distinguish it from the broader shadow economy. Direct methods are considered more adequate for investigating issues related to the workers` profile and their reasons for undertaking undeclared work along with the typed and structure of undeclared work (Eurofound, 2013; Feld and Larsen, 2012; Franić et al., 2022; Williams et al., 2017). A major methodological issue when using audits is the lack of representativeness of the data due to the non-randomness of the sample audited. Sound statistical analysis requires careful adjustment (Arezzo and Guagnano, 2019). Tax audits and experiments are biased due to sample representativeness issues, making them unsuitable for estimating tax evasion or undeclared work size. Surveys, while commonly used, face challenges such as social desirability bias in responses (Burgstaller et al., 2022; Franić et al., 2022; Schneider and Buehn, 2018; Williams et al., 2017). It is very likely that many respondents do not reveal their participation in undeclared work when responding to surveys because they do not want to or fear to confess their fraudulent behavior. So far, only a few attempts have been made to quantify the amplitude of the bias related to the social desirability of responses in surveys measuring involvement in undeclared work (Kirchner et al., 2013; Trappmann et al., 2014). The aim of this paper is to fill this research gap, and to provide a methodological solution for obtaining reliable estimates of undeclared work from sample surveys. To achieve this, a novel approach based on the combination of existing methodologies is here used. The core part relies upon a statistical model that explicitly allows for untruthful responses and that estimates the probability of misleading answers along with the other usual model parameters (Hausman et al., 1998). In addition, the underestimation of the probability of undeclared work due to the imbalance of the sample was corrected using a method proposed by King and Zeng (2001).

In the next section, a short overview is provided of the methods used to estimate the size of undeclared work, highlighting their strengths and weaknesses. The third section then provides the statistical background of our approach and, based on a simulation study, a comparison between the estimates obtained from the proposed methodology and those from a classical unadjusted model. Using the data of Eurobarometer 2019 survey, the fourth section reports the results. The fifth and final section then draws conclusions by discussing the results and the policy implications of the findings.

2. Methods to estimate the size of undeclared work

Due to its very nature of being hidden, estimating the size of undeclared work represents a challenging task and, despite numerous methods being developed, there is no method without limitations and shortfalls (Burgstaller et al., 2022; Schneider, 2021; Schneider and Buehn, 2018). This section briefly reviews the main methods and their strengths and weaknesses.

2.1. Indirect methods

Indirect methods, often called indicator methods, use various indicators to capture the size of the shadow economy. These methods are mainly macro in nature and the most known include: the discrepancy methods (e.g., the income gap method, the gap between the actual and official labour force), the monetary transaction approach, the currency demand approach, the electricity consumption method, and MIMIC estimation procedure (Burgstaller et al., 2022, Feld and Larsen, 2012; Schneider, 2021; Schneider and Buehn, 2018).

The most common discrepancy methods include the income gap and the discrepancy between the official and the actual labour force. The income gap method relies on the assumption that people do not fully declare their activity and income if they purchase more goods and services than they could officially afford based on their declared income (Feld and Larsen, 2012). The main shortcoming of this method is that any error in the measurement of the expenditure results in over-estimating the real size of the shadow economy (Schneider and Buehn, 2018).

One of the established methods for evaluating the gap between the actual and official labour force is the Labour Input Method (LIM), developed by the Italian National Institute of Statistics. This method estimates the gap between the figures about the labour derived from national surveys (self-reported) and the official figures on employment reported by the private sector (from documents to fiscal agencies, social security offices etc.) (De Gregorio and Giordano, 2016; Elek and Köllő, 2019; Franić et al., 2022; Williams et al., 2017). The main advantage of this method is that it only comprises estimates of undeclared work, not the shadow economy. However, one of the shortcomings of this method is that the estimates can be affected by cases when persons have an official job but undertake undeclared work as well (Schneider and Buehn, 2018).

The monetary transaction approach has been developed by Feige (a full description is available in Feige, 1996). It is based on Fisher`s equation of the theory of money assuming that, over the time, there is a constant relationship between a country's official GNP and the volume of transactions (Feld and Larsen, 2012; Schneider and Buehn, 2018). This method is highly influenced by the accuracy of measuring the transactions (which is difficult especially for cash transactions) and includes all transactions, including those from criminal economy (Schneider and Buehn, 2018).

The currency demand approach assumes that undeclared activities use cash payments to not leave traces for the authorities. As such, the higher the demand for cash, the higher the shadow economy (Ahumada et al., 2008; Feld and Larsen, 2012; Schneider and Buehn, 2018). However, there is some criticism of this method. Isachsen and Strøm (1985) argue that the transactions in the shadow economy can occur without cash payments. Additionally, assumptions like equal money velocity between declared and shadow economies (Hill and Kabir, 1996) and the choice of a baseline year without a shadow economy (Schneider and Buehn, 2018) affect the accuracy of estimates. Dybka et al. (2022) solve part of the measurement errors associated with the currency demand models.

The electricity consumption methods start from the assumption that the consumption of electric power represents a good physical indicator to measure the economic activity covering both, the declared economy and the shadow economy. The estimate of the shadow economy, therefore, is obtained by subtracting the estimates of the declared economy from this overall estimation of the consumption of electric power (Feld and Larsen, 2012; Schneider and Buehn, 2018). There are two main econometric methods preminently used to this end, namely Kaufmann - Kaliberda method (1996) and Lackó method (1998). These methods face criticism because some (undeclared) economic activities do not require high electricity consumption, and technological advancements have reduced overall electricity use. As a result, estimates can be biased. Moreover, these methods include the entire shadow economy, encompassing criminal activities, which could impact the accuracy of the results (Schneider and Buehn, 2018).

Finally, the MIMIC method treats the shadow economy as an unobservable variable and takes into account multiple variables (e.g., related to the economy, the labour force or the financial markets) for obtaining the estimates of the shadow economy. The strength of this method is represented by the multitude of variables considered for the model. Its criticism is related to the dependency of the variables included in the model, the lack of stability of the coefficients in respect with sample size as well as the need of using calibration techniques for obtaining absolute figures (Feld and Larsen, 2012; Schneider and Buehn, 2018).

In sum, the most important shortcoming of the indirect methods is the fact that they over-estimate the size of shadow economy and, implicitly, the size of undeclared work (Burgstaller et al., 2022). Furthermore, using indirect approaches, apart from the

LIM estimates, it is not possible to capture the size of undeclared work from the total size of shadow economy (Burgstaller et al., 2022). Shadow economy has a larger scope than undeclared work and include the criminal activities and other forms of illegal activity. Other shortcomings of indirect methods include the sensitiveness to their assumptions and calculation, the "double-counting" issue (i.e., the use of causal factors such as regulation, unemployment, taxation which are also motivations for undertaking undeclared activities), and the fact that some of them (e.g., MIMIC method) only provide relative coefficients instead of absolute values, making the method highly influenced by the calibration procedure used (Burgstaller et al., 2022; Schneider, 2021). As such, these indirect methods are rather suitable for estimating the shadow economy and their applicability for estimating the undeclared work is limited.

2.2 Direct approaches

Moving to the direct approaches, although they can capture the size of undeclared work unlike most of the indirect methods, they too have weaknesses. The main direct methods include surveys, estimations based on self-reported data for tax auditing (or other compliance tools) as well as experiments (Burgstaller et al., 2022; Doerr et al., 2022; Schneider and Buehn, 2018).

Tax audits or other compliance tools can be used for estimating the size of undeclared work. In some instances, tax audits are conducted on random samples of individuals, as seen in the U.S. Internal Revenue Service (Feld and Larsen, 2012). However, usually, tax audits focus on random samples of taxpayers rather than random sample of individuals (Schneider and Buehn, 2018). As such the estimation can be biased and, to provide reliable estimates, the researcher must come up with credible selection-bias corrections (Arezzo and Guagnano, 2019). Another shortcoming of this approach is that tax audits reflect only the proportion that is related to tax fraud and do not include other aspects which are relevant for undeclared work (e.g., income hidden from social protection agencies etc.).

Moving to the use of experiments, previous literature has focused mainly on individual tax evasion. The main advantage of using experiments is the possibility of manipulation of the exogenous variables and the identification of the causal effects of tax evasion. The primary drawback lies in the limited external validity of laboratory experiments, often conducted on students. Additionally, both laboratory and field experiments suffer from limited generalizability (Burgstaller et al., 2022). Although there are recent applications of experiments (e.g., Bjørneby et al., 2021; Burgstaller and Pfeil, 2022; Doerr and Necker, 2021; Fochmann et al., 2021; Hallsworth, 2014), this method is not suitable for estimating the magnitude of tax evasion or the size of undeclared work due to its limited generalizability.

Finally, for analysing the issue of undeclared work, sample surveys are commonly used as they enable undeclared work to be differentiated from other forms of tax evasion and the broader shadow economy (Burgstaller et al., 2022). These sample surveys are commonly conducted on individuals, but they are also utilized to analyse company managers. This approach assumes that the managers are in a unique position and have good knowledge for estimating various types of undeclared activities undertaken in their industry (see for example, Putnins and Sauka, 2015). This method is considered a more appropriate method to investigate issues related to the nature of undeclared work such as who undertakes undeclared work, their motivations, and the types and structure of undeclared work (Eurofound, 2013; Feld and Larsen, 2012; Franić et al., 2022; Williams et al., 2017). Indeed, there are certain limitations of using this method to estimate the size of undeclared work. One of the drawbacks is the sensitivity on how the questions are formulated as well as where they are placed in the questionnaire (Burgstaller et al., 2022; Feld and Larsen, 2012). In addition, due to cultural issues, there are difficulties in building a cross-national survey which can allow comparison between countries (Renooy et al., 2004; Schneider and Buehn, 2018). However, the main challenge when using surveys for measuring the size of undeclared work is the social desirability bias (Burgstaller et al., 2022; Franić et al., 2022; Schneider and Buehn, 2018; Williams et al., 2017). Indeed, there is a widespread view that many respondents will be dishonest in their answers as they do not want to, or they fear to, confess their fraudulent behavior.

In sum, the direct methods are mainly biased due to under-estimation issues. Using tax audits or experiments, the estimates only capture certain aspects related to undeclared work (e.g., those related to tax evasion). Sample surveys have three main advantages. Firstly, they allow one to differentiate undeclared work from other aspects of the shadow economy or tax evasion. Secondly, they are the most suitable method to investigate the structure of undeclared work, the profile of the individuals undertaking undeclared work and their motivations. Thirdly, the sample surveys allow one to generalize the findings as they usually employ random samples. However, the social desirability issue limits their application for estimating the size of undeclared work. Indeed, very few studies have sought to quantify the amplitude of the bias related to social desirability in responses on surveys of undeclared work (Kirchner et al., 2013; Trappmann et al., 2014), and this research is limited to a narrow geographic and cultural setting (i.e., Germany). Hence, this gap. Below, attention turns towards a method for identifying those respondents in surveys who decline to report their undeclared work.

3. Material and methods

3.1 Data and variables

To estimate the level of involvement in undeclared work in lived practice, data from special Eurobarometer survey no. 498 conducted in 2019 is here used. The survey interviewed 27,565 individuals aged at least 15 years living in one of the EU-27 countries and the UK (the survey was conducted before the UK left the EU). To ensure representativeness, a multi-stage random sample design was applied in each state and the number of sampling points were drawn with probability proportional to population size and to population density. The sample sizes vary from a minimum of 505 in Malta to a maximum of 1565 in Germany.

Given the sensitive nature of the topic, questions on their different types of involvement in undeclared work were asked gradually. First, the attitudes of the respondents towards the acceptability of certain behaviors involving different forms of undeclared work were investigated. Second, this was followed by questions on whether they had purchased undeclared goods and services whose production included undeclared work and third and finally, whether they had participated in undeclared work themselves.

The binary 0/1 variable of interest (i.e., supply of undeclared work) is based on the question “Apart from a regular employment, have you yourself carried out any undeclared paid activities in the last 12 months?”, and equals 1 in the case of a positive answer. Out of 27,565 interviewed, only 961 answered yes, 26,139 answered no, 312 refused and 153 answered don’t know. This leads to an estimate of 3.5% of the EU population involved in undeclared work which would rises to 5.2% if those who refused or who said don’t know were included as engaging in such work. As it is not universally true that non-respondents are undeclared workers, we cannot conclude that the above figure is a reliable estimate of the population prevalence, which requires a more appropriate methodology.

To comprehend the effect of individual determinants on participation in undeclared work, a modified probit model (see subsection 3.3) is specified. Among the independent variables, the set of controls were chosen according to the previous related literature (Williams and Horodnic, 2015a, 2015b; Williams and Horodnic 2017). In particular, the following regressors were included:

- Female: a 0/1 variable with 1 for women.
- Age: age of the respondent at the interview, quantitative.
- Urban: a 0/1 variable with 1 if the respondent lives in a town of any size, and 0 otherwise.
- Occupation: a categorical variable indicating employment condition of the respondent (possible categories: unemployed, self-employed, employed, retired or inactive).
- Financial problems: a categorical variable indicating the degree of difficulties in paying bills (possible categories: most of the time, from time to time, almost never/never).
- Country: a categorical variable indicating the country of residence of the respondent, among the 27 EU States and the UK involved in the survey.
- Detection risk: a categorical variable measuring the individual perceived probability of being detected when perpetrating fraudulent behavior (possible categories: very high, fairly high, fairly small, very small).
- Expected sanction: a categorical variable corresponding to the expected sanction in the event of being caught undertaking undeclared work (possible categories: tax or social security contributions, tax or social security contributions plus a fine, prison).
- Tax morale: a quantitative variable, constructed by averaging the answers to questions on various non-compliant behaviors. Lower levels of this variable indicate higher levels of tax morale.

The missing data in the dependent variable were cancelled out due to their modest amount (1.7%). Regarding missing data in the covariates (see Table 1), we first checked for significant differences in their number between respondents who reported undeclared work and those who did not. We found none and then excluded that the missing mechanism depends on the binary variable of interest. We then proceeded with multiple imputation, using the method developed by Stekhoven and Bühlmann (2012) that easily allows the treatment of categorical and continuous variables simultaneously and that does not require any assumptions on the distribution of the variables.

3.2 A new method to measure the involvement in undeclared work

When an ad-hoc survey is available, it is very easy to estimate the percentage of people who undertake undeclared work. In what follows we will refer to this population prevalence as π. For the binary variable Y observed without error (where Y=1 if the respondent work off-the-book and Y=0 otherwise), an unbiased estimate of the population prevalence is the sample proportion, possibly weighted:

$$\hat{\pi} = Pr\widehat{(Y_i = 1)} = \left(\frac{\sum_{i=1}^{n} y_i}{n}\right). \qquad (1)$$

More in-depth analysis is possible. As stated above, a strength of direct methods based on random surveys is the possibility to quantify the effect of each determinant on involvement in undeclared activities, the others held constant. By specifying an appropriate statistical model (typically a binary response model), we can estimate how much a certain characteristic (for example being a woman or having occasional financial problems) impacts on the probability of undertaking undeclared work. Based on a regression model, we can estimate the population prevalence π along with the regression parameters.

The typical model for a binary dependent variable is as follows. Let Y^* be a latent (i.e., unobservable) continuous variable which, in our case, is the propensity to undertake undeclared work; let X_i be a random vector of q characteristics measured on individual i, relevant to explain the latent variable, U_i^S be the utility to undertake undeclared work, $U_i^{\bar{S}}$

be the utility of not undertaking undeclared work and β be a vector of unknown regression coefficients. Then:

$$Y_i^* = U_i^S - U_i^{\bar{S}} = X_i\beta + \varepsilon_i$$

If the utility U_i^S outweigh the utility $U_i^{\bar{S}}$, then we observe *Y=1*, otherwise *Y=0*. Note that, unlikely Y^* which is latent, *Y* is observable.

The probability of observing someone undertaking undeclared work can be modelled as follows:

$$\pi_i = Pr\left(Y_i = 1 | X_{1i}, X_{2i}, \dots, X_{qi}\right) = \mathrm{F}\left(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_q X_{qi}\right) \quad (2)$$

where *F(·)* is a generic cumulative density function (cdf), typically gaussian or logistic (probit and logit model respectively).

Once model (2) is estimated, important information can be retrieved: the signs of the beta coefficients denote an exposure (if positive) or a protection (if negative) to the risk of undertaking undeclared work and the net effect of each covariate can be estimated allowing to change and taking constant all the other predictors. Another advantage of using a model such as (2) is that it allows the estimation of *individual* probabilities π_i of having worked undeclared. The average of such estimated individual probabilities, that is $\hat{\pi} = \frac{1}{n}\sum \hat{\pi}_i$, is itself an estimate of the population prevalence π.

The log likelihood of model (2) is:

$$logL(\beta) = \sum_{i=1}^{n} y_i log\mathrm{F}(\beta' X_i) + (1 - y_i) log[1 - \mathrm{F}(\beta' X_i)] \quad (3)$$

As the literature has correctly pointed out, it is very likely that *Y* is misclassified. Misclassification of a binary variable means that an observation with a true value of 0 is observed as 1 or an observation that is truly a 1 is observed as a 0. Therefore, the observed binary variable Y differs from the true one. A relevant consequence is that neither the estimate in equation (1) nor that of equation (2) are consistent. In particular, the sample proportion of observed 1's underestimates π since there are some false 0's. On the other hand, it is unlikely to observe a false 1, i.e. a person who does not work off-the-book declares to have done so, and hence to overestimate π. Moreover, when models like (2) are used in presence of a misclassified binary dependent variable, the parameter estimates become inconsistent, making it impossible not only to have a trustworthy estimates of the betas, but also reliable estimate of the percentage of people working in the undeclared economy. But the advantage of working with a regression model like (2) is that it allows to deal with misclassified dependent variables. In the literature, there are two main approaches. The first requires additional data to verify the reliability of responses. Among the others, we may mention the proposals of Chua and Fuller (1987) and Poterba and Summers (1995). The first, for a J-level dependent variable, considers a parametric model that includes all possible $J(J-1)$ misclassifications, but requires at least three additional surveys to re-interview the original respondents; therefore, it is hardly feasible. The second, based on a conditional logit procedure, also considers all possible misclassifications and requires re-interviewing respondents to measure possible discrepancies between the different sets of responses.

The second approach, introduced by Hausman et al. (1998) and Abrevaya and Hausman (1999), directly includes the probability of misclassification into the model specification. As explained in greater detail in the next section, they expressly provide that the dependent variable can be observed as a 0 when in lived practice it is a 1, or that it can be observed as a 1 when it is in lived practice a 0. Each of these misclassification events is associated with a probability which is estimated jointly to the other parameters of the binary choice model.

In the following section we show a formalization of the methodology to obtain asymptotically unbiased estimator of this unknown quantity and therefore of the overall true proportion of 1's.

3.3 The model

Let ${}_oY_i$ be the observable error-prone dichotomous variable, and ${}_tY_i$ the corresponding unobservable true response. Referring to the discrepancy between ${}_oY_i$ and ${}_tY_i$, we may distinguish two types of misclassification. According to Hausman et al. (1998), the corresponding probabilities can be defined as:

$$\alpha_0 = Pr\big({}_oY_i = 1 \big| {}_tY_i = 0\big) \qquad (4a)$$

$$\alpha_1 = Pr\big({}_oY_i = 0 \big| {}_tY_i = 1\big) \qquad (4b)$$

i.e., α_0 is the probability that a true 0 is recorded as a 1 and α_1 is the probability that a true 1 is misclassified as a 0 and these are the unknown elements required to provide a consistent estimate of π. As previously noted, in our case study α_0 is expected to be nearly 0, while α_1, which is the probability measure of the social desirability bias, is expected to be greater than 0. These probabilities depend on the value of ${}_tY_i$ but, conditioning on ${}_tY_i$, they are independent of the covariates X_i. The model is identifiable as long as $\alpha_0 + \alpha_1 < 1$ (Hausman et al., 1998).

Referring to ${}_oY_i$, the corresponding two probabilities can be so defined:

$$Pr\big({}_oY_i = 1|X_i\big) = Pr\big({}_oY_i = 1|X_i,\ {}_tY_i = 1\big)Pr\big({}_tY_i = 1|X_i\big) +$$

$$+Pr\big({}_oY_i = 1|X_i,\ {}_tY_i = 0\big)Pr\big({}_tY_i = 0|X_i\big) = (1-\alpha_0-\alpha_1)Pr\big({}_tY_i = 1|X_i\big) + \alpha_0$$

and

$$Pr({}_oY_i = 0|X_i) = 1 - Pr({}_oY_i = 1|X_i) = 1 - \alpha_0 - (1 - \alpha_0 - \alpha_1)Pr({}_tY_i = 1|X_i)$$

Recalling equation (2), in general we have $Pr({}_tY_i = 1|X_i) = \mathrm{F}(\beta' X_i)$.

We assume that $\mathrm{F}(\cdot) = \Phi(\cdot)$, i.e., the cdf of a normal random variable. To estimate the entire vector of parameters, $\theta = (\beta, \alpha_0, \alpha_1)$, we must extend the likelihood function in (3), taking into account the possible misclassifications. The log likelihood then becomes:

$$logL(\theta) = \sum_{i=1}^{n} \left\{ y_i \left(log\left(\alpha_0 + (1 - \alpha_0 - \alpha_1)\Phi(\beta' X_i)\right) \right) + (1 - y_i) log\left(1 - \alpha_0 - (1 - \alpha_0 - \alpha_1)\Phi(\beta' X_i)\right) \right\} \quad (5)$$

Note that by maximizing (5) we obtain consistent estimates of the whole vector of parameters $\theta = (\beta, \alpha_0, \alpha_1)$. Plugging in the consistent estimate $\hat{\beta}$ in equation (2), we obtain the individual probabilities $\hat{\pi}_i$. However, in presence of observed rare events as in our case, these individual probabilities are underestimated. King and Zeng (2001) proposed an *ad hoc* adjustment that we implemented.

3.4 Simulation study

To gain insight into the improvement of the estimation method, we conducted a simulation study comparing the estimates from equation (3) and (5). To this end, we assessed several simulation scenarios corresponding to the following choices:

n= {10000, 15000, 30000}; α_1= {0.05, 0.15, 0.25}; β_1= {0.1, 2}; β_2= {0.05, 1.5}

The other parameters, fixed across the scenarios, were:

α_0=0; β_3= -1.2; β_0= -0.5.

The covariates were generated as follows: $X_1 \sim lognormal(0,1)$, $X_2 \sim Bin(n, \frac{1}{3})$ and $X_3 \sim Unif(0,1)$. We replicated 100 times each scenario.

In Figure 1 the main results. The boxplots synthetize the distribution of the relative sampling error of the estimators ($\frac{\hat{\beta}_{kj}-\beta_k}{\beta_k}; k = 0,1,2,3 ; j = 1,2, \dots, 100$), for both the corrected and unadjusted probit (Corr_probit and Probit, respectively).

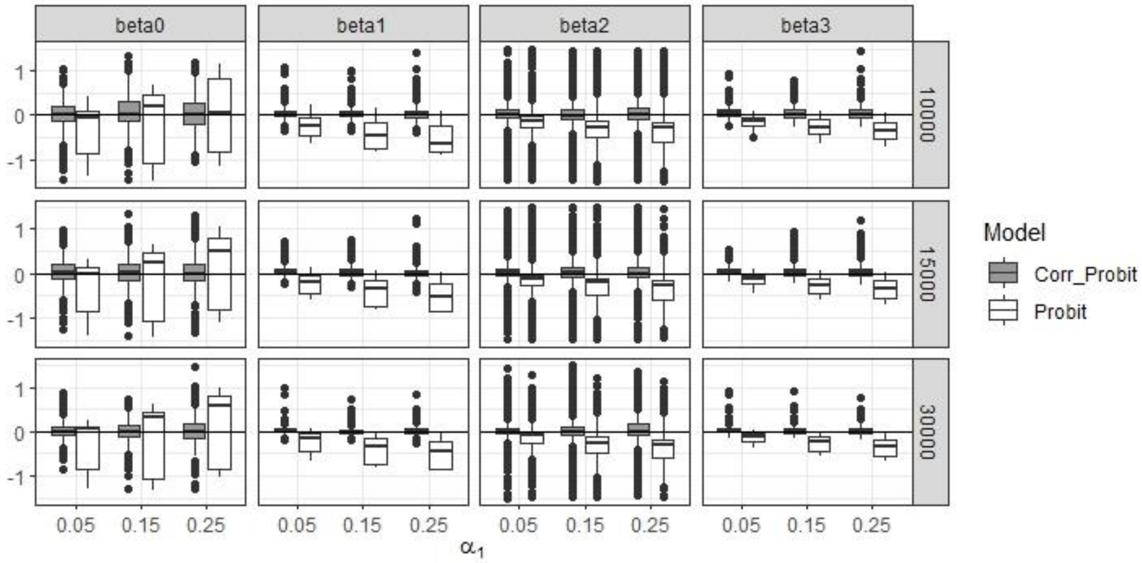

Fig 1 – Boxplots of the relative sampling errors of the estimators for each regression parameter, in adjusted (grey) and unadjusted (white) model, over the simulation scenarios.

The distributions of the adjusted estimates are always centered on 0. On the contrary, uncorrected estimates are distorted with a bias that increases with α_1 and does not decrease even with increasing sample size.

Considering the same transformation as before, Figure 2 displays the distributions of the relative sampling errors of $\hat{\alpha}_1$. It appears that the distribution is always centered on 0, and as α_1 and n increase, it is more and more concentrated on such value.

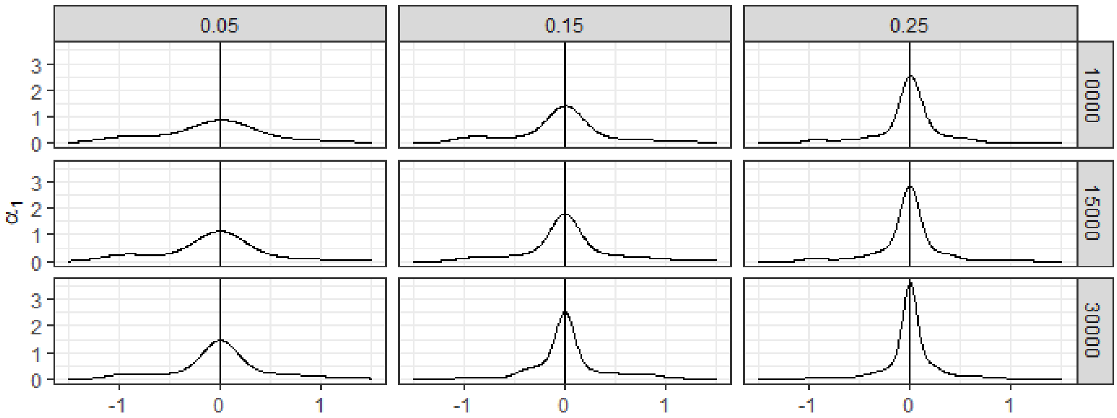

Fig 2 – Distribution of the relative sampling errors of $\hat{\alpha}_1$ over the simulation scenarios.

4. Results

Table 1 reports some descriptive statistics. When directly asking respondents whether they have conducted undeclared work or not in the past 12 months, only 3.5 per cent admitted doing so. This low number is not surprising considering the social desirability issues of admitting illegal behaviors. Indeed, the issue of social desirability has been documented previously as affecting the sincerity of responses in relation to participation to undeclared work and represents, as explained above, the main shortcoming of using representative surveys for estimating the size of undeclared work (Burgstaller et al, 2022; Franić et al., 2022; Kirchner et al., 2013; Schneider and Buehn, 2018; Trappmann et al., 2014; Williams et al., 2017).

Variable	**Definition**	**Mean/Mode**	**Missing** (N=27,565)
Supply of undeclared work (dependent variable)	Dummy for participation in undeclared paid activities in the last 12 months	0.035	465
Female	Dummy for female	0.546	0
Age	Respondent age	51.52	0
TM	Tax morale	2.481	1505
Urban	Dummy for living in a town of any size	0.657	0
Occupation	Occupation of the respondent	Employed (43.69%)	0

Financial problems:	Difficulty in paying bills	No financial problems (67.08%)	406
Detection risk:	Individual perception of detection risk	Fairly small detection risk (38.25%)	2838
Expected sanctions:	Individual evaluation of sanctions if caught	Normal tax or social security contributions due, plus a fine (57.76%)	3010
Country	Nationality of the interviewee	Germany (5.68%)	0

Table 1. Variables in the models: definition and descriptive statistics

To address the issue of social desirability and provide a method of using this representative survey to estimate the size of undeclared work, we conducted a Probit regression analysis. Table 2 presents the results of both Probit regression with and without correction. All quantitative analysis in the study was performed using the software R (the code is available as supplementary material). The results for the relationship between the dependent variable (i.e., undertaking undeclared work) and the independent variables are similar for the two Probit regressions, with and without correction. This result further reinforces previous findings which suggest that representative surveys (even if using models without correcting for social desirability issues) are suitable for identifying the population groups more likely to conduct undeclared work or their motivation of doing so (as, for example, in this case, the perception of the sanctions applied for undertaking undeclared work or their tax morality) (Eurofound, 2013; Feld and Larsen, 2012; Franić et al., 2022; Williams et al., 2017). As such, the finding is that in Europe, women are less likely to participate in undeclared work. Similarly, the likelihood to participate in undeclared work reduces with age and with a high tax morale. Contrariwise, the likelihood to participate in undeclared work is higher for the unemployed compared to other employment statuses as well as for those having financial difficulties most of the time. These results are in line with previous research which concludes that undeclared work is conducted by marginalized groups, such as those without employment or in financial distress (Brill, 2011; Taiwo, 2013; Slavnic, 2010; Williams and Horodnic, 2015). Analyzing the perception towards the sanctions applied for undertaking undeclared work, the finding is that the likelihood to participate to undeclared work decreases with the perception of a higher risk of detection and a higher applicable sanction for undertaking such work, adding to the literature that supports the deterrence approach in reducing participation in undeclared work (Feld and Frey, 2002; Kluge and Libman, 2017; Mas'ud et al., 2015; Mazzolini et al., 2017; Slemrod et al., 2001).

Turning to the key issue in this paper, Table 2 shows that the proportion of those who reported no involvement in undeclared work is almost a quarter of those whose actual status is irregular. ($\widehat{\alpha_1}$=0. 233). This result gives a quantification of the social desirability bias and shows that it is not negligible and that it cannot be neglected when analysing participation in undeclared work.

	Probit			Probit with misclassification correction		
	Coef	Se	pvalue	Coef	se	pvalue
(Intercept)	-0.473	0.123	0.00	-0.385	0.129	0.00
Female	-0.269	0.032	0.00	-0.284	0.035	0.00
Age	-0.012	0.001	0.00	-0.012	0.001	0.00
Tax Morale	0.076	0.006	0.00	0.084	0.006	0.00
Urban	-0.024	0.034	0.49	-0.027	0.038	0.48
Occupation (Ref. Cat.: Unemployed)						
Self-employed	-0.084	0.072	0.24	-0.099	0.079	0.21
Employed	-0.364	0.057	0.00	-0.388	0.062	0.00
Inactive	-0.397	0.068	0.00	-0.419	0.074	0.00
Retired	-0.536	0.074	0.00	-0.608	0.082	0.00
Financial problems (Ref. Cat: Most of the time)						
From time to time	-0.291	0.052	0.00	-0.313	0.056	0.00
Almost never/never	-0.501	0.050	0.00	-0.537	0.055	0.00
Detection risk (Ref. Cat: Very small)						
Very high	-0.169	0.071	0.02	-0.188	0.079	0.02
Fairly high	-0.244	0.045	0.00	-0.262	0.049	0.00
Fairly small	0.006	0.041	0.88	0.009	0.044	0.84
Expected sanctions (Ref. Cat: Tax or social security contributions)						
Tax or social security contributions plus a fine	-0.042	0.034	0.22	-0.040	0.037	0.29
Prison	-0.215	0.085	0.01	-0.228	0.094	0.02

alpha0		0.001	0.001	0.25
alpha1		0.233	0.014	0.00
logLik	-3615.56	-3615.43		

Table 2. Coefficient estimates of a probit model with and without correction[1].

The point estimate of the overall proportion of participation in undeclared work, i.e. $\hat{\pi}$, is found by averaging the estimates of individual probabilities found through the model as in equation (2). As previously stated, when the sample is highly imbalanced, the probabilities $\hat{\pi}_i$ are sharply underestimated. In such cases, a correction is recommended (King and Zeng, 2001). Following such approach, we obtained the point estimate $\hat{\pi} \cong 0.173$. The 95% interval estimate, whose bounds embody the uncertainty of the population proportion, is [0.170; 0.176]. The standard errors were computed using the delta method (see appendix A for details). This estimate appears realistic if compared with figures related to estimates of the whole shadow economy using MIMIC estimates, which was estimated to represent 20% of the EU28 GDP (Medina and Schneider, 2018). Finally, as expected, the estimate for α_0 is not statistically significant.

5. Discussion

This paper has reviewed the methods for estimating participation in undeclared work and shown that indirect methods tend to over-estimate participation and direct methods to under-estimate participation. The latter is due to social desirability bias that leads respondents not to report their engagement. The aim of this paper has been to provide a methodological solution that considers this social desirability bias in responses in sample surveys and provides more reliable estimates of participation in undeclared work. For doing so, this paper used data from special Eurobarometer survey no. 498

[1] Due to the lack of space, the estimates corresponding to EU countries are not included in the table, but are available upon request.

conducted in 2019 in EU-27 countries and the UK. Only 3.5 per cent openly admitted to participating in undeclared work. The results of a Probit model with correction for both misclassified cases (i.e., those undertaking undeclared work but declaring that they do not) and imbalance reveals the following: 1) nearly a quarter (23.3 per cent) of the respondents undertaking undeclared work refused to openly admit it during the survey; 2) overall, the percentage of people involved in undeclared work is approximately 17.3%. As such, the methodological advancement of this paper is that it provides a useful tool for researchers and policymakers using the sample surveys to consider social desirability bias when estimating participation in undeclared work. The outcome of this new methodological advancement is that survey estimates of the level of undeclared work can now report this in a more accurate manner than has been previously the case. Nevertheless, a limitation of this study is the absence of comprehensive insights into the specific characteristics of individuals prone to offering dishonest responses about their involvement in undeclared work. Future research endeavours could delve deeper into identifying distinct clusters of individuals predisposed to dishonest reporting. These findings could subsequently inform tailored measures for tackling undeclared work, including targeted inspections or awareness initiatives highlighting the advantages of formal employment.

Acknowledgment

This work was supported by the Sapienza University of Rome under the Visiting Professor Call 2020 and under the Ateneo funding projects nr P1181643624426C/2018.

Declarations of interest: none

Appendix A: derivation of standard errors for $\hat{\pi}_i$

The standard errors are computed with the well-known delta method (Greene, 1993). In the general form, it states that it is possible to approximate the asymptotic behavior of a function of asymptotically normal random variables. The result holds even when the expected value and the variance of the function are unknown.

Let us consider a generic function $h(\vartheta)$; the delta method states that its variance is:

$$Var(h(\vartheta)) = J_h' \Omega_\vartheta J_h \qquad \text{(A1)}$$

where Ω_ϑ is the variance and covariance matrix of ϑ, and $J_h = \frac{\partial[h(\vartheta)]}{\partial\vartheta}$ is the Jacobian.

Since we estimate the population prevalence π by averaging the individual estimated probabilities $\hat{\pi}_i = \Phi(X_i\hat{\beta})$, it follows that in our case $h(\vartheta) = \sum_{i=1}^{n} \Phi(X_i\hat{\beta})/n$. The Jacobian of $h(\vartheta)$ requires the first derivative of the standard normal cdf $\Phi(\cdot)$:

$$J'_h = \frac{\partial\left[\sum_{i=1}^{n} \Phi(X_i\hat{\beta})/n\right]}{\partial\hat{\beta}} = \frac{1}{n}\sum_{i=1}^{n} \varphi(X_i\hat{\beta})\, X_i$$

where $\varphi(\cdot)$ is the standard normal density function. Finally, the variance covariance matrix of equation (A1) is the variance covariance matrix of $\hat{\beta}$.

References

Abrevaya, J., Hausman, J.A., 1999. Semiparametric Estimation with Mismeasured Dependent Variables: An Application to Panel Data on Employment Spells. *Annales D'Economie et de Statistique*, 56, 243–75.

Ahumada, H., Alvaredo, F., Canavese, A., 2008. The monetary method to measure the shadow economy: The forgotten problem of the initial conditions. *Economics Letters* 101 (2), 97-99.

Arezzo, M.F., Guagnano, G., 2019. Is it possible to detect Tax Evasion using administrative data? A proposal based on a Binary Choice Model Affected by a Severe Censoring Mechanism. *International Journal of Statistics & Economics* 20 (2), 1-15.

Bjørneby, M., Alstadsæter, A., Telle, K., 2021. Limits to third-party reporting: evidence from a randomized field experiment in Norway. *Journal of Public Economics* 203, 104512.

Brill, L., 2011. *Women's Participation in The Informal Economy: What Can We Learn from Oxfam's Work?* Manchester: Oxfam.

Burgstaller, L., Feld, L.P., Pfiel, K., 2022. Working in the shadow: Survey techniques for measuring and explaining undeclared work. *Journal of Economic Behavior and Organization* 200, 661–671.

Burgstaller, L., Pfeil, K., 2022. You Don't Need an Invoice, Do You? An Online Experiment on Collaborative Tax Evasion. Unpublished Manuscript, Walter Eucken Institute.

Chua, T.C., Fuller, W.A., 1987. A model for multinomial response error applied to labor flows. *Journal of the American Statistical Association*, 82, 46–51.

De Gregorio, C., Giordano, A., 2016. The heterogeneity of undeclared work in Italy: some results from the statistical integration of survey and administrative sources. *Rivista di Statistica Ufficiale* 2, 99-129.

Doerr, A., Hartmann, C., Sajons, C., 2022. Unregistered work among refugees: Evidence from a list experiment in Germany. Working papers 2022/01. Faculty of Business and Economics - University of Basel.

Doerr, A., Necker, S., 2021. Collaborative tax evasion in the provision of services to consumers: a field experiment. *American Economic Journal: Economic Policy* 13 (4), 185-216.

Dybka, P., Olesiński, B., Rozkrut, M., Torój, A., 2022. Measuring the model uncertainty of shadow economy estimates. *International Tax and Public Finance*, https://doi.org/10.1007/s10797-022-09737-x

Elek, P., Köllő, J., 2019. Eliciting permanent and transitory undeclared work from matched administrative and survey data. *Empirica* 46 (3), 547-576.

Eurofound (2013). *Tackling Undeclared Work in 27 European Union Member States and Norway: Approaches and Measures since 2008*. Dublin: Eurofound.

European Commission (2007). Stepping Up the Fight Against Undeclared Work. Brussels: European Commission.

European Commission (2020). COVID-19: the impact on undeclared work and key challenges. Available at: https://ec.europa.eu/social/main.jsp?langId=en&catId=1299&furtherNews=yes&newsId=9702, Accessed on 22.04.2022.

Feige, E.L., (1996). Overseas holdings of U.S. currency and the underground economy. In: Pozo, S. (Ed.). Exploring the Underground Economy (pp. 5-62). W.E. Kalamazoo: Upjohn Institute for Employment Research.

Feld, L.P., Frey, B.S., (2002). Trust breeds trust: How taxpayers are treated. *Economics of Governance* 3, 87-99.

Feld, L.P., Larsen, C., (2012). *Undeclared Work, Deterrence and Social Norms: The Case of Germany*. Berlin: Springer.

Fochmann, M., Fochmann, N., Kocher, M.G., Müller, N., (2021). Dishonesty and risk-taking: compliance decisions of individuals and groups. *Journal of Economic Behavior and Organization* 185, 250-286.

Franić, J., Horodnic, I.A., Williams C.C., (2022). *Extent of undeclared work in the European Union*. Brussels: European Commission.

Greene, W.H. (1993). *Econometric Analysis*, Prentice Hall.

Hallsworth, M. (2014), The use of field experiments to increase tax compliance. *Oxford Review of Economic Policy* 30 (4), 658-679.

Hausman, J.A., Abrevaya, J., Scott-Morton, F.M., (1998). Misclassification of the dependent variable in a discrete-response setting. *Journal of Econometrics*, 87, 239–269.

Hill, R., Kabir, M., (1996). Tax rates, the tax mix, and the growth of the underground economy in Canada: what can we infer? *Canadian Tax Journal/Revue Fiscale Canadienne* 44, 1552-1583.

International Labour Organization (2020). *World Employment and Social Outlook*. Trends 2020. Geneva: ILO.

Isachsen, A.J., Strøm, S., (1985). The Size and Growth of the Hidden Economy in Norway. *Review of Income and Wealth* 31 (1), 21-38.

Kaufmann, D., Kaliberda, A., (1996). Integrating the unofficial economy into the dynamics of post socialist economies: a framework of analyses and evidence. In: Kaminski, B. (Ed.). Economic Transition in Russia and the New States of Eurasia (pp. 81-120). London: M.E. Sharpe.

King, G., & Zeng, L. (2001). Logistic Regression in Rare Events Data. Political Analysis, 9(2), 137-163.

Kirchner, A., Krumpal, I., Trappmann, M., von Hermanni, H., (2013). Measuring and explaining undeclared work in Germany–an empirical survey with a special focus on social desirability bias. *Zeitschrift fur Soziologie* 42 (4), 291-314.

Kluge, J.N., Libman, A., (2017). Sticks or carrots? Comparing effectiveness of government shadow economy policies in Russia. IOS Working Papers No. 364. Institute for East and Southeast European Studies.

Lackó, M., (1998). The Hidden Economies of Visegrad Countries in International Comparison: A Household Electricity Approach. In: Halpern, L., Wyplosz, C. (Eds.). Hungary: Towards a Market Economy. Cambridge (Mass.): Cambridge University Press.

Mas'ud, A., Manaf, N.A.A., Saad, N., (2015). Testing assumptions of the 'slippery slope framework' using cross-country data: Evidence from sub-Saharan Africa. *International Journal of Business and Society* 16 (3), 408-421.

Mazzolini, G., Pagani, L., Santoro, A., (2017). The deterrence effect of real-world operational tax audits. DEMS Working Paper Series No. 359. Milan: Department of Economics, Management and Statistics, University of Milan – Bicocca.

Medina, L., Schneider, F., (2018). Shadow economies around the world: what did we learn over the last 20 years? IMF Working Papers 17, 1-76.

OECD (2012), Reducing Opportunities for Tax Non-Compliance In The Underground Economy. Paris: OECD.

Poterba, J.M., Summers, L.H., (1995). Unemployment benefits and labor market transitions: a multinomial logit model with errors in classification. *Review of Economics and Statistics*, 77, 207–216.

Putnins, T., Sauka, A., (2015). Measuring the Shadow Economy Using Company Managers. *Journal of Comparative Issues* 43 (2), 471-490.

Renooy, P., Ivarsson, S., van der Wusten-Gritsai, O., Meijer, E., (2004). *Undeclared Work in an Enlarged Union – An Analysis of Undeclared Work: An In-Depth Study of Specific Items*. Brussels: European Commission.

Schneider, F., (2021). Do Different Estimation Methods Lead to Implausible Differences in the Size of the Non-Observed or Shadow Economies? A Preliminary Answer. CESifo Working Paper, No. 9434.

Schneider, F., Buehn, A.(2018). Shadow economy: estimation methods, problems, results and open questions. *Open Economics* 1, 1-29.

Schneider, F.; Enste, D.H., (2000). Shadow Economies: Size, Causes, and Consequences. *Journal of Economic Literature* XXXVIII, 77-114.

Slavnic, Z., (2010). Political economy of informalization. *European Societies* 12 (1), 3-23.

Slemrod, J., Blumenthal, M., Christian, C., (2001). Taxpayer response to an increased probability of audit: Evidence from a controlled experiment in Minnesota. *Journal of Public Economics* 79, 455-483.

Stekhoven DJ, and Bühlmann P (2012) Missforest-non-parametric missing value imputation for mixed-type data. Bioinformatics, 28, 112–118.

Taiwo, O., (2013). Employment choice and mobility in multi-sector labour markets: theoretical model and evidence from Ghana. *International Labour Review* 152 (3-4), 469-492.

Trappmann, M., Krumpal, I., Kirchner, A., Jann, B., (2014). Item sum: a new technique for asking quantitative sensitive questions. *Journal of Survey Statistics and Methodology* 2 (1), 58-77.

Williams, C.C., (2014). Out of the shadows: a classification of economies by the size and character of their informal sector. *Work, Employment and Society* 28 (5), 735-753.

Williams, C.C., Bejakovic, P., Mikulic, D., Franić, J., Kedir, A., Horodnic, I.A., (2017). *An Evaluation of the Scale of Undeclared Work in the European Union and its Structural Determinants: Estimates using the Labour Input Method*. Brussels: European Commission.

Williams, C.C., Horodnic, I.A., (2015a). Rethinking the marginalisation thesis. *Employee Relations* 37 (1), 48-65.

Williams, C.C. and Horodnic, I.A. (2015b). Evaluating the prevalence of the undeclared economy in Central and Eastern Europe: an institutional asymmetry perspective. *European Journal of Industrial Relations* 21 (4), 389-406.

Williams, C.C. and Horodnic, I.A. (2017), Evaluating the illegal employer practice of under-reporting employees' salaries. *British Journal of Industrial Relations* 55 (1), 83-111.