Analyzing the Determinants of the Shadow Economy With a "Separate Approach". An Application of the Relationship Between Inequality and the Shadow Economy

ROBERTO DELL'ANNO[*]

University of Salerno, Fisciano (Sa), Italy

Summary. — This paper suggests a "separate" approach to analyze the determinants of the shadow economy (SE). It is applied to investigate the relationship between inequality and the SE on a cross-section of 118 countries. We disentangle the effect of inequality on the SE ratio by estimating both direct and indirect effects on both the numerator and denominator of the ratio separately. We find that an increase in inequality increases the SE ratio. This positive correlation is primarily due to a reduction in the official GDP rather than an increase in the SE.
© 2015 Elsevier Ltd. All rights reserved.

Key words — shadow economy, inequality, separate approach, unobserved economy

1. INTRODUCTION

The shadow economy (SE) is a subject of considerable interest, and the literature on the analysis of its determinants is particularly extensive (see Friedman, Johnson, Kaufmann, & Zoido-Lobaton, 2000; Schneider, 2011; Schneider & Enste, 2000 for an overview). This paper aims at contributing to this issue by proposing an alternative approach to estimate the influence of a potential determinant on the SE ratio. The basic intuition of this research is to demonstrate that estimating the influence of an explanatory variable on a dependent variable measured as a ratio (hereinafter, "ratio approach") may be not conclusive because it blurs the impact of the explanatory variable on the denominator (e.g., official economy) with the impact that it has on the numerator (e.g., SE). Accordingly, we propose calculating the overall effect, estimating both direct and indirect impacts on both the official and on the unobserved Gross Domestic Product (GDP) separately (hereinafter "separate approach"). An empirical application of this approach is conducted to explore the relationship between income inequality and the SE.

The paper consists of two parts. In the first "methodological" part, we address the issue of the different approaches to define (and measure) the SE and introduce the "separate approach". The second part of the paper applies the proposed methodological hints to investigate the relationship between the income distribution and the SE. Over the past two decades, several research works empirically supported the hypothesis that income inequality and the SE are positively correlated (e.g., Ahmed, Rosser, & Rosser, 2007; Chong & Gradstein, 2007; Rosser, Rosser, & Ahmed, 2000, 2003). We verify that this result is empirically validated both by utilizing the "ratio approach" and by applying the "separate approach".

In sum, the paper contributes to the existing literature in several ways. Following the order in which they are presented in the article, we attempt to reconcile the definitions of the SE utilized in economic research with the Non-Observed Economy (NOE) concept adopted by national statistical institutes; because the "ratio approach" may cause misinterpretation of the actual influence of an explanatory variable on a ratio variable, we propose estimating both the direct and indirect effects of the explanatory variable on the numerator (i.e., SE) and

denominator (i.e., official GDP) disjointedly; we provide a method to calculate the effect of a determinant on the SE ratio by controlling for the double counting of a part of the SE in the SE ratio; and concerning the relationship between inequality and the SE, we find that the overall impact of inequality on the SE ratio is positive and higher than the effect estimated by the "ratio approach".

The paper is organized as follows. Section 2 addresses the definition of the SE and introduces the "separate approach". Section 3 provides theoretical background on the interactions among inequality, official GDP and the SE. Section 4 describes the database, econometric models, and hypotheses and reports the empirical outcomes. Section 5 concludes.

2. DEFINING AND ANALYZING THE SHADOW ECONOMY IN EMPIRICAL RESEARCH

(a) *Defining the shadow economy*

We discuss two general approaches to define and measure the SE. On the one hand, the national accounting system (SNA) employs the label NOE to refer to "*all productive activities that may not be captured in the basic data sources used for national accounts compilation*" (UNECE, 2008, p. 2). Following the Eurostat's (2005) "*Tabular approach to exhaustiveness*", the SNA classifies seven sources of *non-exhaustiveness* for GDP estimates: (N1) Producers deliberately not registered to avoid tax and social security obligations; (N2) Producers deliberately not registered as a legal entity or as an entrepreneur because they are involved in illegal activities; (N3) Producers not required to register because they have no market output; (N4) Legal persons or (N5) registered entrepreneurs not surveyed due to a variety of reasons; (N6) Producers deliberately misreporting to evade taxes or social security contributions; and (N7) Other statistical deficiencies. For

---

[*] I am indebted to anonymous reviewers and editor for providing insightful comments and directions for additional work which has resulted in this paper. The usual disclaimer applies.

analytical purposes, OECD (2014) proposes a simplification of this classification in four types of NOE adjustments. It defines N1 + N6 as *Underground production*, N2 as *Illegal production*, N3 + N4 + N5 as *Informal sector production* (including those undertaken by households for their own final use) and N7 as *Statistical deficiency*.

The second approach to define the SE is prevalent in economic research. Here, the adjectives informal, shadow, hidden, second, black, unrecorded, unofficial, unobserved, etc., are often utilized synonymously with terms such as economy, sector, market, and GDP. However, these labels refer to distinct phenomena and should be used appropriately (Bagachwa & Naho, 1995; Feige, 1990; Feige & Urban, 2008) to avoid misunderstandings. In this literature, a plurality of macro-econometric methods to estimate the SE is proposed. Among these methods, the Multiple Causes Multiple Indicators (MIMIC) approach and the currency demand approach are becoming dominant. Attempting to systematize the common definitions in this area of research, we identify two recent studies as benchmarks for the two most common sources of macro-econometric estimates of the SE, i.e., Buehn and Schneider (2012) for the MIMIC method and Alm and Embaye (2013) for the currency demand approach. The two studies adopt two different mainstream definitions of the SE. They differ in dealing with illegal activities in the SE. Specifically, Buehn and Schneider (2012, p. 141) define the SE as including all market-based legal production of goods and services that are deliberately concealed from public authorities to avoid payment of taxes or social security contributions, to avoid having to meet certain legal labor market standards, and to avoid complying with certain administrative procedures or statistical questionnaires. Following Smith (1984), Alm and Embaye (2013, p. 512) employ a somewhat broader definition of the SE that includes "*all market-based goods and services (legal or illegal) that escape inclusion in official accounts*". In other words, while Buehn and Schneider (2012) include "*all market-based legal production*", Alm and Embaye also consider market-based illegal production.

Aiming to find a *trait d'union* between the most used labels in economic research (i.e., SE) and in national accounting system (i.e., NOE), we distinguish four types of GDP aggregates: recorded observed economy (Gdp^{RO}); recorded non-observed economy (Gdp^{RNOE}) and unrecorded non-observed economy (Gdp^{UNOE}). Given the foregoing definitions, we can label the total economic activity as $Gdp^T = Gdp^{RO} + Gdp^{NOE}$ and the official (published) GDP as $Gdp^{off} = Gdp^{RO} + Gdp^{RNOE}$, where the total NOE is given by $Gdp^{NOE} = Gdp^{RNOE} + Gdp^{UNOE}$.

Combining this classification with the seven sources of non-exhaustiveness for GDP estimates proposed by the Eurostat's (2005) *Tabular approach to exhaustiveness*, we obtain a precise definition of the estimates of the SE ratio calculated by Alm and Embaye (2013) utilizing a modified Currency demand approach (SE_{Curr}^{Macro}) and Buehn and Schneider (2012) utilizing MIMIC modeling (SE_{MIMIC}^{Macro}).

A preliminary explanation is required here. Following Alm and Embaye's definition literally, we should include in the numerator only the GDP that "*escapes inclusion in the official accounts*", i.e., $Gdp^{UNOE(N1+N6)}$. However, the currency demand approach estimates a linear transformation of this value. In the last stage of the currency demand approach, the amount of the unobserved GDP is obtained by multiplying the stock of currency used to escape taxes and administrative

burdens (C^*) by the velocity of money (V). Considering that the velocity of money is the ratio between the nominal (official) GDP and money supply, what a researcher obtains by multiplying C^* by V is inevitably an estimate of the unobserved GDP that includes an additional share of the NOE in the same proportion – that we denote by b – in which the recorded NOE is included in the official GDP (hereinafter "currency demand bias"). Accordingly, we include $(1 + b)$ in the numerator of SE_{Curr}^{Macro}.

$$SE_{Curr}^{Macro} \equiv \frac{(1+b)Gdp^{UNOE(N1+N2+N6)}}{Gdp^{RO} + Gdp^{RNOE(Total)}} \qquad (1)$$

where b is the proportion of Gdp^{RNOE} on Gdp^{off} ($Gdp^{RNOE} = bGdp^{off}$).

With reference to the MIMIC estimates of the SE ratio, the numerator of SE_{Mimic}^{Macro} follows (1) because of the calibration of the MIMIC model to the currency demand method. [1]

$$SE_{Mimic}^{Macro} \equiv \frac{(1+b)Gdp^{UNOE(N1+N6)}}{Gdp^{RO} + Gdp^{RNOE(Total)}} \qquad (2)$$

This issue might be easily solved if the estimates of the imputed NOE were officially published and homogeneously estimated at the national level. However, this is not the normal case because national statistical offices do not regularly divulge the size of NOE adjustments in the official statistics. Moreover, for the countries where these data are available, these adjustments should be cautiously employed for cross-countries comparisons because of the differences in methodologies and practices followed by offices in estimating the NOE (OECD, 2014; UNECE, 2008).

In general, assuming no measurement errors, the differences between the macro-econometric and statistical national accounting methods may be explained both by divergences in the coverage of the NOE types and by the factor $(1 + b)$. For instance, the discrepancy between Alm and Embaye's (2013) estimates and the size of adjustments in national accounting (SE^{SNA}) should be equal to the imputed unobserved GDP yield by unregistered producers because they have no market output (N4 + N5), statistical discrepancies (N7) and unrecorded NOE for underground and illegal production divided by the official economy multiplied by the factor $(1 + b)$ (i.e., $SE_{curr}^{Macro} - SE^{SNA} = (1 + b)Gdp^{UNOE(N1+N2+N6)}/Gdp^{off}$). Again, the discrepancy between the SE ratio obtained by Buehn and Schneider's (2012) MIMIC specification and those obtained by the currency demand should be equal to the proportion of unobserved economy due to illegal activities (N2). However, given that the estimates obtained by currency approach calibrate the Buehn and Schneider's (2012) MIMIC model, we cannot extrapolate N2 by comparing these two sources of data. Hence, in the following, we will assume that the difference between SE_{MIMIC}^{Macro} and SE_{Curr}^{Macro} only depends on measurement errors. Concerning the consequence of this assumption, OECD (2014) states that N1 + N6 adjustments for NOE activities almost always represent the most significant part of the adjustments for non-exhaustiveness, reaching as much as 80% of all adjustments in some countries; therefore, we could suppose that our simplification does not significantly affect the results. In sum, the MIMIC and currency demand estimates of the SE approximately measure the following ratio:

$$SE^{Macro} \equiv (1+b)\frac{Gdp^{UNOE}}{Gdp^{RO} + Gdp^{RNOE}} \qquad (3)$$