



UNIVERSITY OF CYPRUS
ECONOMICS RESEARCH CENTRE



Economic Policy Papers

Income under-reporting and the shadow economy in Cyprus: Evidence from household survey data

E. Andreou, G. Papadopoulos and G. Syrichas
Economics Research Centre, University of Cyprus

No. 04-20

November 2020

Publication Editor: Angela Shekersavva

ERC Sponsors (in alphabetical order)

Central Bank of Cyprus

Directorate General for European Programmes, Coordination and Development

Economics Department, University of Cyprus

Ministry of Finance

University of Cyprus

Disclaimer: the views expressed in the Economic Policy Papers and Economic Analysis Papers are of the authors and do not necessarily represent the ERC.

Υποεκτίμηση Εισοδημάτων και η Παραοικονομία στην Κύπρο: Στοιχεία από Έρευνες Νοικοκυριών

Έ. Ανδρέου, Γ. Παπαδόπουλος, Γ. Συρίχας

Περίληψη

Οι εκτιμήσεις για την υποβολή αναφορών εισοδήματος με βάση τα δεδομένα των νοικοκυριών καθώς και το μέγεθος της παραοικονομίας ενδιαφέρουν τόσο τους ακαδημαϊκούς, καθώς και τους υπεύθυνους χάραξης πολιτικής. Για το σκοπό αυτό, έχουν αναπτυχθεί διάφορες μέθοδοι με τη χρήση μικροοικονομικών δεδομένων καθώς και μακροοικονομικών δεδομένων. Αυτή η μελέτη εφαρμόζει τη μικροοικονομική (έμμεση) μέθοδο που πρότειναν οι Pissarides και Weber (1989) σε δύο βάσεις δεδομένων σε επίπεδο νοικοκυριού και υπολογίζει την υπό-αναφορά εισοδήματος για την Κύπρο μετά την παγκόσμια χρηματοπιστωτική κρίση και πριν από την κρίση στην εγχώρια οικονομία το 2012/13. Χρησιμοποιούμε δεδομένα τόσο από την Έρευνα Προϋπολογισμού Νοικοκυριών (HBS) όσο και από την Έρευνα Οικονομικών και Καταναλωτικών Νοικοκυριών (HFCS), δεδομένου ότι και οι δύο αυτές έρευνες πραγματοποιήθηκαν τον ίδιο χρόνο στην Κύπρο και παρέχουν παρόμοιες αλλά συμπληρωματικές πληροφορίες για τα νοικοκυριά. Η ενημερωμένη ανάλυσή μας για το 2009 δείχνει ότι η υπο-αναφορά εισοδήματος παρέμεινε σταθερή τουλάχιστον από το 2002/3 όταν η μελέτη των Pashardes και Polycarrou (2008) υπολόγισε ότι κυμαίνεται από 38% έως 44%.

CONTENTS

<i>Abstract</i>	1
1. <i>Introduction</i>	2
2. <i>Method and Data</i>	3
3. <i>Empirical Results</i>	8
4. <i>Conclusions</i>	12

Income under-reporting and the shadow economy in Cyprus: Evidence from household survey data

E. Andreou, G. Papadopoulos, G. Syrichas

Abstract

Income under-reporting estimates based on households data as well as the size of the shadow economy are of interest to academics as well as policy makers. For that purpose, various methods have been developed using micro/survey data as well as aggregate/macro data. This study applies the micro (indirect) method proposed by Pissarides and Weber (1989) on two household-level databases and estimates income under-reporting on Cyprus in the wake of the Global Financial Crisis and before its peak on the domestic economy in 2012/13. We use data from both the Household Budget Survey (HBS) as well as the Household Finance and Consumption Survey (HFCS), given that both of these surveys were conducted in the same year in Cyprus and provide similar yet complementary information on households. Our updated analysis for 2009 shows that income under-reporting has remained stable at least since 2002/3 when the study of Pashardes and Polycarpou (2008) estimated it to be in the range of 38% to 44%.

1. Introduction

Income under-reporting estimates based on households data as well as the size of the shadow economy are of interest to academics as well as policymakers. For that purpose, various methods have been developed using micro/survey data as well as aggregate/macro data. For a survey of the literature, the methods employed and their strengths and weaknesses, the recent studies by Schneider (2014) and Schneider and Buehn (2018), among others, provide excellent references.

For Cyprus in particular, several studies have estimated income misreporting and the size of its black economy. Recent studies by Medina and Schneider (2018) using the MIMIC approach estimated the size of the shadow economy in Cyprus during the period 1991-2015 to be on average 31,3% of its GDP. In addition, Andreou, Andreou, Michael and Syrighas (2020) use macro methods for the period 1995-2018 to estimate the shadow economy in Cyprus to be in the ranges of 11%-18% of the GDP following the currency demand method and 26%-34% of GDP according to the energy consumption approach.

This study complements the above results by applying a micro (indirect) method proposed by Pissarides and Weber (1989) on two household-level databases and estimates income under-reporting on Cyprus in the wake of the Global Financial Crisis and before its peak on the domestic economy in 2012/13. An earlier study by Pashardes and Polycarpou (2008) applied the Pissarides and Weber (1989) approach to estimate income under-reporting in Cyprus using the Household Budget Survey (HBS). Their results suggest that income under-reporting of the self-employed was around 40%, corresponding to an approximate size of the shadow economy of 8% of GDP in 2003. The objective of this paper is to provide an updated estimate of the income misreporting and the size of the shadow economy in Cyprus using the Pissarides and Weber model on data from both the Household Budget Survey (HBS) as well as the Household Finance and Consumption Survey (HFCS), given that both of these surveys were conducted in the same year in Cyprus and provide similar yet complementary information on households. The survey wave of 2009 has been used since it is common in both databases thus allowing the comparison and cross-checking of the results. Also, using data for 2009 is guided by the fact that the impact of the crisis of 2012/13 that hit Cyprus would not bias our empirical results and provide a more direct comparison of both household surveys with the 2003 study by Pashardes and Polycarpou (2008).

2. *Method and Data*

To estimate households' income misreporting we follow the method described in the seminal work of Pissarides and Weber (1989) (henceforth PW). The PW method estimates income misreporting by dividing the population into two groups and comparing their reported food consumption in relation to their reported income. In particular, the household population is split into salaried worker (SW) households and self-employed (SE) ones.¹ One of the key assumptions is that SW households accurately report their income when surveyed. This is reasonable since their income is already taxed and in general, the respective authorities have accurate records of it. On the contrary, SE households might misreport their income in the fear that this information might be revealed to the tax authorities to which they might have also misreported their income. The second key assumption is that both household groups report accurately their food expenditures. Due to its innocuous nature, households do not have a motive to misreport such information since there is no concern of leading to a possible reveal of income misreporting.

The intuition behind the method, assuming similar preferences, is that households from the two groups should exhibit the same food consumption for the same income. If this holds, the log-linear Engle curves (log food expenditure versus log income) for each household group should have the same intercepts.² Alternatively, any discrepancy is attributed to SE households under-reporting their income, given that SW households are assumed to report it accurately.

In particular, the following regression model is estimated on the pooled household sample:

$$c_i = \beta y_i + \gamma SE_i + \alpha X_i + \varepsilon_i \quad (1)$$

where i is the subscript denoting individual households, c_i is the logarithm of monthly food expenditures, y_i the logarithm of income, SE_i a dummy discriminating the two types of households by their employment status (salaried worker or self-employed), taking the values 1 for SE and 0 for SW, and ε_i the error term. In addition to income, the model specification in equation (1) includes the vector X_i of control variables which also affect consumption. The coefficient γ captures the vertical distance between the

¹ More precisely, households in which the head is a salaried worker and households in which the head is self-employed.

² The elasticity of consumption with respect to income is assumed to be the same in the two household groups.

two groups' Engle curves with $\gamma > 0$ indicating that the self-employed households under-report their income. The degree of under-reporting is expressed as the ratio γ/β . An alternative way to present under-reporting is by the number by which a self-employed household's reported income needs to be multiplied to arrive at the true figure. This number is simply obtained as: $k = \exp(\gamma/\beta)$ and can be used to get an estimate of the size of the shadow economy if one knows the aggregate figure of income from self-employment as a share of the GDP.

There are two points worth highlighting with respect to this method. The first one is that this approach yields an estimate of income under-reporting of the self-employed *relative* to the reference group, i.e. salaried worker households. By assuming that the latter truthfully report their income, one can obtain an absolute figure for the self-employed households' under-reporting. However, it may well be that salaried worker households also under-report their income, thus the results should be cautiously interpreted.³ The second one refers to the fact that Equation 1 represents a consumption function and therefore the income variable refers to the permanent, disposable income. However, survey data contain information about households' current income. Since the original study of Pissarides and Weber (1989), the literature has employed several methods to circumvent this issue. In the original paper, the authors impose some additional assumptions and treat income as endogenous which allows the estimation of upper and lower bounds for income under-reporting. This has been the standard way of dealing with this issue in the subsequent literature (Schuetze, 2002; Lyssiotou, 2004; Johansson, 2005; Kim et al., 2009; Kukk et al. 2018), along with the simplified approach which completely disregards that problem and provides a point estimate (Feldman and Slemrod, 2007; Engström and Holmlund, 2009; Martinez-Lopez, 2013; Hurst et al., 2014; Kukk and Staehr, 2014).

Our empirical analysis uses data from the Household Budget Survey (HBS) conducted by the Statistical Service of Cyprus and the Household Finance and Consumption Survey (HFCS) coordinated by the European Central Bank which employs a harmonized methodology and definitions. The survey wave of 2009 has been used in particular, since it is common in both databases thus allowing the comparison of the two cross household sections that have different sample characteristics and coverage at the same reference year. In addition, the empirical analysis for the survey data in 2009 is guided by the fact the impact of the crisis of 2012/13 that hit Cyprus could alter

³ This could be particularly true in Cyprus. Anecdotal evidence suggests that a significant number of salary workers have a second job in the informal market. Employees both in the private and public sector provide services in education, tourism agriculture, construction and maintenance in exchange for extra untaxed cash.

the empirical results and allows a comparison with the Pashardes and Polycarpou (2008) results which refer to the 2003 HBS.⁴

The HBS's household sample covers about 1% of the total population in Cyprus and contains variables on their demographic characteristics, income and its sources, expenditures, employment, loans and residence. A detailed description of the database is presented in Polycarpou (2018). The HFCS is a micro-level dataset which provides a wealth of information on euro area households. It contains harmonized information on households' demographics, balance sheet and consumption variables in the euro area. A specific aspect of the HFCS database is the oversampling of wealthy households in order to ensure their adequate representation in the sample. The final household weights are taking this into account along with adjustments for the sampling design and item non-response. The weights are calibrated using external benchmarks to obtain results which are consistent with information known over the total population. Non-response is treated with multiple imputation methods, estimating missing observations conditional upon information that could credibly explain them.

The key variables used in the analysis are households' food expenditure, their income and their employment status; whether household's head is a salaried worker or self-employed. Additional variables which could affect food consumption include demographic information such as the head's age, number of children and their age, variables related to the household's economic and financial position such as their real assets, whether they rent or own their accommodation and the employment status of the spouse. In 2009 the global had only started affecting Cyprus.

The sample selection criteria are as in Pissarides and Weber (1989). Only households with two adults who are a couple and a working household head are considered in the empirical analysis. In order to eliminate potential outliers, observations have been trimmed at the bottom 1st percentile of food expenditure's distribution. Tables 1 and 2 present the descriptive statistics of the main, as well as some of the additional variables and provide an overview of the relevant information from the two databases. At the bottom of each table, the ratio of each variable's average, as well as median figures between the two groups, are presented. These ratios provide a more direct picture of how the two groups compare to each other for every specific variable. Interestingly the mean and median ratios are similar for all variables suggesting a symmetric

⁴ In 2009 the global had only started affecting Cyprus. It was the only economy in the euro area which recorded positive growth in the first quarter of 2009 before ending the year with a relatively small contraction in the order of 1,7% of GDP.

distribution, except in the case of the number of children (for which the mean ratio between the two groups is equal to 1 whereas the median ratio is 0).⁵

TABLE 1
Descriptive statistics of the HBS database's sample for 2009

	Statistic	Monthly food expenditure (EUR)	Annual income (EUR)	HH's age	# of children	# of cars	# of rooms
Salaried workers	Mean	466.4	52788.6	45.3	0.7	2.1	5.9
	Median	433.2	47738	46	0	2	6
	Std.	233.3	26445.8	10.4	1	0.9	1.5
	Min	101.9	14616	24	0	0	1
	Max	1686.2	323078	85	6	6	14
	N	1117	1117	1117	1117	1117	1117
Self-employed	Mean	506.6	45284.9	48.1	0.8	2.1	5.8
	Median	449.7	42217	49	0	2	6
	Std.	315.1	21560.8	10.8	1	0.9	1.5
	Min	131.4	14665	27	0	0	2
	Max	2888.3	180610	76	4	5	15
	N	195	195	195	195	195	195
	Means ratio	0.92	1.17	0.94	0.88	1	1.02
	Median ratio	0.96	1.13	0.93	0	1	1

Note: Monthly food expenditure is the 1/12th of annual expenditure on food and non-alcoholic beverages. Income is household's total, net income.

⁵ Different levels of trimming of 2% were also considered and the empirical results remained robust.

TABLE 2

Descriptive statistics of the HFCS database's sample for 2009

	Statistic	Monthly food expenditure (EUR)	Annual income (EUR)	HH's age	# of children	# of cars	Size of HMR (m²)
Salaried workers	Mean	877.6	56735.2	40	1	1.9	173.2
	Median	800	47197.8	36	1	2	166
	Std.	467.3	46003	10.8	1	0.7	83.8
	Min	250	0	23	0	0	35
	Max	4000	448547.9	72	6	7	604
	N	335	335	335	335	335	335
Self-employed	Mean	986.6	41428.8	47.4	0.8	2	221.5
	Median	961	35824.5	46	0	2	200
	Std.	589.1	34477.8	13	0.9	0.8	109.5
	Min	300	600	26	0	1	80
	Max	3500	285806.1	65	3	5	560
	N	63	63	63	63	63	63
	Means ratio	0.89	1.37	0.84	1.25	0.95	0.78
	Median ratio	0.88	1.32	0.78	-	1	0.83

Note: Monthly food expenditure refers to the expenditure on food and non-alcoholic beverages consumed at home and outside home. Income is household's total, gross income.

A key observation in both datasets is that average (or median) food expenditure is higher in self-employed households, while their average (or median) reported income is lower than the salaried worker ones. This is in-line with the intuition that, given their higher food expenditure, the reported income figures by the self-employed may not reflect their actual one. An interesting observation is that income figures in the HFCS are lower than the respective ones from the HBS, for the self-employed group of households. This holds even though the former are reported in gross while the latter in net terms. However, the dispersion in the HFCS is substantially higher, with the related coefficients of variation in both groups being twice as large compared to the HBS. Moreover, households in the HFCS sample are younger, the salaried worker ones have slightly more children and they are almost identical in terms of car ownership. The proxy for households' main residence (HMR) size is different between the two datasets, therefore the figures cannot be directly compared (HBS: number of rooms, HFCS: surface). Nevertheless, the data from the HFCS suggest that the average/median self-employed households' main residence is larger than their counterparts', whereas the HBS doesn't indicate the difference to be considerable. An important aspect of the two datasets is the size of the respective samples; in both employment groups, the HBS sample is about three times larger than the HFCS's.

While the salaried workers' group is sufficiently populated in both databases, the self-employed group's size in the HFCS is only 63 individuals which could affect the estimation results.

3. *Empirical Results*

The estimation follows the original approach of PW. It is based on the estimation of the regression model in Equation 1 which represents the monthly food expenditures on income, a dummy identifying households' employment status and a set of controls. In addition to income, we include a vector of demographic and financial position control variables, X_i , that affect households' consumption such as the household head's age, the number of kids and their age, a dummy if the household lives in rented accommodation and a dummy if they own it. Note that we consider the logarithmic transformation which eliminates observations with zero food expenditure or income. To address endogeneity issues due to a possible correlation between food expenditure and income to common shocks, the model is estimated with 2SLS. Income is instrumented with variables such as the number of cars, the proxy of the household's main residence size (number of rooms or square meters), a dummy indicating whether the household's head has tertiary education and dummies related to the spouse's employment status. Table 3 reports the estimation results of Equation 1 on the two datasets.

TABLE 3

IV estimates of Equation 1 on the HBS and HFCS databases

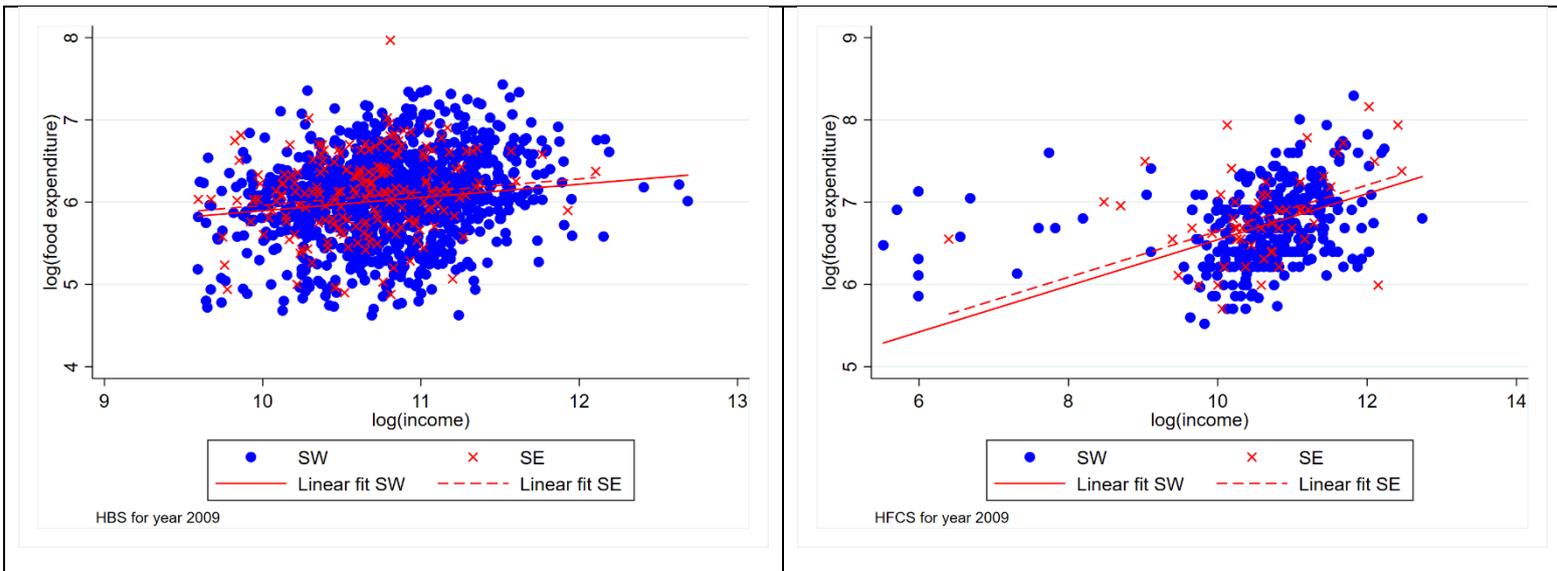
Variable	HBS	HFCS
LNCY	0.161*** (0.055)	0.280** (0.139)
SED	0.065** (0.037)	0.105 (0.106)
AGEH	0.008*** (0.002)	0.014*** (0.004)
AGEHSQ	-0.0003*** (0.000)	-0.0003 (0.0004)
NCH	-0.034 (0.034)	0.019 (0.087)
NCHSQ	0.031*** (0.010)	0.015 (0.025)
KIDS2	0.115*** (0.044)	-0.054 (0.053)
OOM	0.010 (0.016)	-0.066 (0.066)
REN	-0.013 (0.049)	-0.316*** (0.099)
Constant	4.288*** (0.590)	3.741** (1.501)
Endogeneity test (p-value)	0.768	-
Instrument strength test (F-statistic)	134.922	-
Overidentifying restrictions test (p-value)	0.024	-

Standard errors in parentheses. */**/** denotes significance at 10%/5%/1% level. LNCY is the log of net and gross income in the HBS and the HFCS databases respectively.

The two most important coefficients for the estimation of income misreporting are those of LNCY (β) and SED (γ). We find that in both datasets, β is statistically significant, while for γ this holds only for the case of the HBS. This suggests that in the HFCS, the two groups statistically exhibit the same average propensity to consume. This could indeed be the result of the small sample, given the high regressor-to-observations ratio in the case of self-employed households in the HFCS database. Model diagnostics in the case of the estimation on the HBS dataset suggest that LNCY is exogenous, therefore estimation via OLS could be used instead of 2SLS. However, the high value (> 10 as suggested by Stock et al. (2002)) of the F-statistic indicates that the instruments used are strong. Finally, the test for overidentifying restrictions suggests that instruments are valid at the 1% significance level.

The scatter plots in Figure 1 shows the relationship between income and food expenditure for the two household groups from each database.

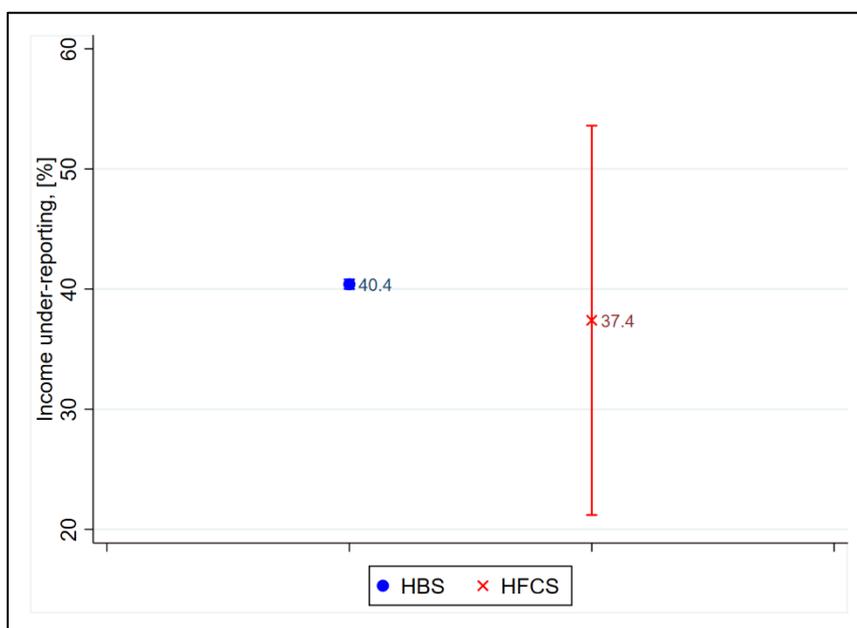
Figure 1: Scatter plot of food expenditure versus income and the respective Engle curves by household group for the HBS (left) and HFCS (right) datasets



In Figure 1, the differences described in Tables 1 and 2 are evident. Data in the HBS are less dispersed compared to the HFCS ones. In the right panel, the HFCS dataset shows a region with a few low income – low food expenditure households which seem to be separated from the rest observations. Moreover, there are apparent differences in sample sizes with the HBS being denser and more populated compared to the HFCS. This prohibits excluding the possibly outlying observations from the latter database since it would further reduce the sample. In addition, this issue points to a possible explanation of the fact that, while the magnitude of the two γ coefficients is comparable in the two datasets, it is statistically significant only in the HBS.

The results of income under-reporting's estimates from the two datasets are presented in Figure 2.

Figure 2: Estimates of income under-reporting of the self-employed using the HBS (blue circle, left) and the HFCS (red X, right) databases for 2009



The results in Figure 2 show two main findings. The first is that income under-reporting by the self-employed in Cyprus in 2009 is estimated to be about 40% in both databases or alternatively, one needs to multiply their reported income by about 1.5 in order to get the true figure. This result is very close to the one obtained by Pashardes and Polycarpou (2008) for 2002/3. The respective estimates of income under-reporting were found to be between 38% and 44%. Therefore, it seems that income under-reporting (and arguably, the size of the shadow economy) has remained unchanged in the two years 2003 and 2009. The second finding is the range of the estimates. For the case of the HBS, it is so narrow, that it is barely visible on the plot. The upper bound is 40.8% while the lower one 40%. On the contrary, the estimated range for the case of the HFCS is noticeably wide, spanning from a low 21.2% to a maximum of 53.6%. This is probably associated with the small sample which yields weaker results.

Despite the small sample issues, the HFCS has a wealth of information which is not available in the HBS database. In particular, it contains detailed information about households' real and financial assets and net wealth position. This information can be exploited to extend the PW framework by replacing food expenditure with a household's total assets. The rationale is that food, being a necessity good, might exhibit a concave relationship with income. Thus, as income grows, food expenditure increases at a slower pace until it eventually flattens. On the contrary, the accumulation

of wealth does not have an upper limit. Exploiting this relationship might better capture income under-reporting. Although not reported in detail here, some preliminary results indeed raise the figure of income under-reporting by about 20 percentage points, to about 60%, when using the log of total assets as a dependent variable instead of food expenditure.

4. Conclusions

This study has applied the method of Pissarides and Weber (1989) on two, household-level databases (the HBS and HFCS) and estimated income under-reporting in Cyprus be around 40% in 2009 (in both databases) before the domestic banking and economic crisis in 2012/13. Both datasets point to the same direction, though the results from the HFCS are weaker, exhibiting wide bounds, possibly due to small sample issues. The ranges of the estimates of the two databases differ due to the different sample sizes and information, yet they are still of interest. For the HBS the estimated range is quite narrow - the upper bound is 40.8% while the lower one 40%. On the contrary, the estimated range for the case of the HFCS is noticeably wide, spanning from a low 21.2% to a maximum of 53.6%, mostly due to the smaller sample size of self-employed households. The findings of the estimated 40% income under-reporting in 2009 suggest that for every 1 euro of income from self-employment received, the respective households report approximately 60 cents for taxation purposes. Our updated analysis for 2009 shows that income under-reporting has remained stable compared to that of 2002/3 when the study of Pashardes and Polycarpou (2008) estimated it to be in the range of 38% to 44%. Consequently, the level of the shadow economy has remained the same. The corresponding approximate size of the shadow economy was 8% of GDP in 2003 (Pashardes and Polycarpou, 2008).

The findings should be interpreted with caution. They relate to the specific sample under examination and should not be generalized over the whole population. However, they provide a rough estimate of income under-reporting and the size of the shadow economy. Another potential caveat is the assumption that salaried worker households accurately report both their income. Indeed, studies suggest that income from self-employment, financial assets, rents and secondary jobs is more prone to being misreported (Neri and Zizza, 2010). However, this does not guarantee that the reported income figures from salaried worker households are accurate.

Interesting future research could be the study of the evolution of income misreporting around the peak of the business cycle. Using survey waves pre- and post-2012/13, one could examine the impact of the Global Financial Crisis and its ramifications on

the shadow economy as captured by the respective surveys. Some preliminary results point to the direction of the reduction of income under-reporting in comparison to its pre-crisis levels. Yet, this topic deserves further investigation to uncover its magnitude and potential causes.

References

- Engström, P., & Holmlund, B. (2009). Tax evasion and self-employment in a high-tax country: evidence from Sweden. *Applied Economics*, 41(19), 2419-2430.
- Feldman, N. E., & Slemrod, J. (2007). Estimating tax noncompliance with evidence from unaudited tax returns. *The Economic Journal*, 117(518), 327-352.
- Fethi, M. D., Fethi, S., & Katircioglu, S. T. (2006). Estimating the size of the Cypriot underground economy: A comparison with European experience. *International Journal of Manpower*, 27(6), 515-534.
- Georgiou, G. M. (2007). Measuring the Size of the Informal Economy: A Critical Review. Working Papers 2007-1. Central Bank of Cyprus.
- Georgiou, G. M., & Syrichas, G. L. (1994). The underground economy: an overview and estimates for Cyprus. *Cyprus Journal of Economics*, 7(2), 125-43.
- Hurst, E., Li, G., & Pugsley, B. (2014). Are household surveys like tax forms? Evidence from income underreporting of the self-employed. *Review of economics and statistics*, 96(1), 19-33.
- IMF (2010) Executive Board Concludes 2010 Article IV Consultation with Cyprus. <https://www.imf.org/en/News/Articles/2015/09/28/04/53/pn10123>
- Johansson, E. (2005). An estimate of self-employment income underreporting in Finland. *Nordic Journal of Political Economy*, 31(1), 99-109.
- Kim, B., Gibson, J., & Chung, C. (2009). Using panel data to exactly estimate income under-reporting by the self-employed. KIEP Research Paper. Working paper, 09-02.
- Kukk, M., Paulus, A., & Staehr, K. (2018). Income underreporting by the self-employed in Europe: A cross-country comparative study. *Eesti Pank Working Paper 4*.
- Kukk, M., & Staehr, K. (2014). Income underreporting by households with business income: evidence from Estonia. *Post-Communist Economies*, 26(2), 257-276.

Lyssiottou, P., Pashardes, P., & Stengos, T. (2004). Estimates of the black economy based on consumer demand approaches. *The Economic Journal*, 114(497), 622-640.

Martinez-Lopez, D. (2013). The underreporting of income by self-employed workers in Spain. *SERIEs*, 4(4), 353-371.

Medina, L., & Schneider, F. (2018). *Shadow Economies Around the World: What Did We Learn Over the Last 20 Years?* (No. 18/17). International Monetary Fund.

Neri, A., & Zizza, R. (2010). Income reporting behaviour in sample surveys. *Bank of Italy Temi di Discussione (Working Paper) No, 777*.

Pashardes, P., & Polycarpou, A. (2008). Income tax evasion, inequality and poverty. *Cyprus Economic Policy Review*, 2(2), 37-49.

Pissarides, C. A., & Weber, G. (1989). An expenditure-based estimate of Britain's black economy. *Journal of public economics*, 39(1), 17-32.

Polycarpou, A. (2018). The Unified Database of Household Budget Surveys 1984/85, 1990/91, 1996/97, 2003, 2009 and 2015/16. *Economic Policy Papers*, 06-18.

Schneider, F. (2014). *The Shadow Economy and Shadow Labor Force: A Survey of Recent Developments* (No. 8278). IZA Discussion Papers.

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

Schuetze, H. J. (2002). Profiles of tax non-compliance among the self-employed in Canada: 1969 to 1992. *Canadian Public Policy/Analyse de Politiques*, 219-238.

Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20(4), 518-529.