Self-employment income tax evasion and inequality

Martina Bazzoli
Irvapp-FBK
Paolo Di Caro
Italian Ministry of Economy and Finance
Francesco Figari
University of Eastern Piedmont
Carlo Fiorio
carlo.fiorio@unimi.it

University of Milan
Marco Manzo
Italian Ministry of Economy and Finance

Research Article

Keywords: Income tax evasion, self-employment, inequality

Posted Date: June 13th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-6811499/v1

Additional Declarations: No competing interests reported.

Self-employment income tax evasion and inequality *

Martina Bazzoli[1], Paolo Di Caro[2], Francesco Figari[3,5], Carlo Fiorio[4,1*], Marco Manzo[2]

[1]Irvapp, FBK, Trento, Italy.
[2]Department of Finance, Italian Ministry of Economy and Finance, Italy.
[3]Department of Economics, Università degli Studi del Piemonte Orientale, Italy.
[4]Department of Economics, Management and Quantitative Methods, Università degli Studi di Milano, Italy.
[5]CeRP, Collegio Carlo Alberto, Italy.

*Corresponding author(s). E-mail(s): carlo.fiorio@unimi.it;

Abstract

This paper provides new evidence on the extent and impact of income tax evasion by the self-employed, using a novel dataset that links a household income survey with tax administration records in Italy. We find that the share of unreported income among self-employed households is approximately one-third of total income, similar to what has been documented for other countries with lower aggregate tax evasion rates. We show that income tax evasion primarily originates from liberal professionals, who underreport their income by about twice as much as small entrepreneurs. The results suggest that the high level of aggregate tax evasion in Italy is a consequence of the economic structure rather than differences in individual tax evasion propensities or social norms. Our distributional analysis reveals the negative impact of evasion on vertical and horizontal equity, as well as the increased unequal distribution of the tax burden across households.

Keywords: Income tax evasion, self-employment, inequality

JEL Classification: H26 , H31 , D31

---

*We wish to thank Ron Davies, Dominika Langenmayr, and two anonymous reviewers for helpful comments. We used EUROMOD version I2.0, and we acknowledge financial support from the European Union Easi programme (2014-2020). This paper is part of a research project developed at the Department of Finance of the Italian Ministry of Economy and Finance.
The views expressed in this paper do not reflect those of the institutions of affiliation.

1 Introduction

Tax evasion challenges the sustainability of public finances, causes market inefficiency, and distorts the distribution of the tax burden (Andreoni et al. 1998; Slemrod 2019). It also modifies the taxation-inequality nexus (Alstadsæter et al. 2019, 2022), raising concerns about both vertical and horizontal equity. A significant proportion of tax evasion is attributed to under-reporting of taxable income among the self-employed (Bergolo et al. 2023): their income is self-reported to a greater extent than income from wages, salaries and pensions, which are often reported by a third party, leaving less scope for manipulation (e.g. see Kleven et al. 2011; Carrillo et al. 2017; Slemrod et al. 2017). In advanced economies, self-employment tax evasion accounts for approximately one-third of total tax evasion.[1] In Italy, self-employment income tax evasion accounts for more than €31 billion, approximately 37% of total tax evasion (MEF 2024), one of the highest values among all advanced economies. In this country, moreover, self-employment is a key component of the total workforce: the share of self-employed individuals over the total working population is approximately 20%, which is seven percentage points higher than the average observed in European Union (EU) countries.

When data from random audit programs are available, such as in Denmark, the United Kingdom and the United States, they are used to estimate income tax evasion with a close approximation (Feinstein 1991; Kleven et al. 2011). However, estimates based solely on individual random audit data often understate evasion among those with very high incomes, as they fail to capture sophisticated evasion via offshore accounts and pass-through businesses (Guyton et al. 2021).

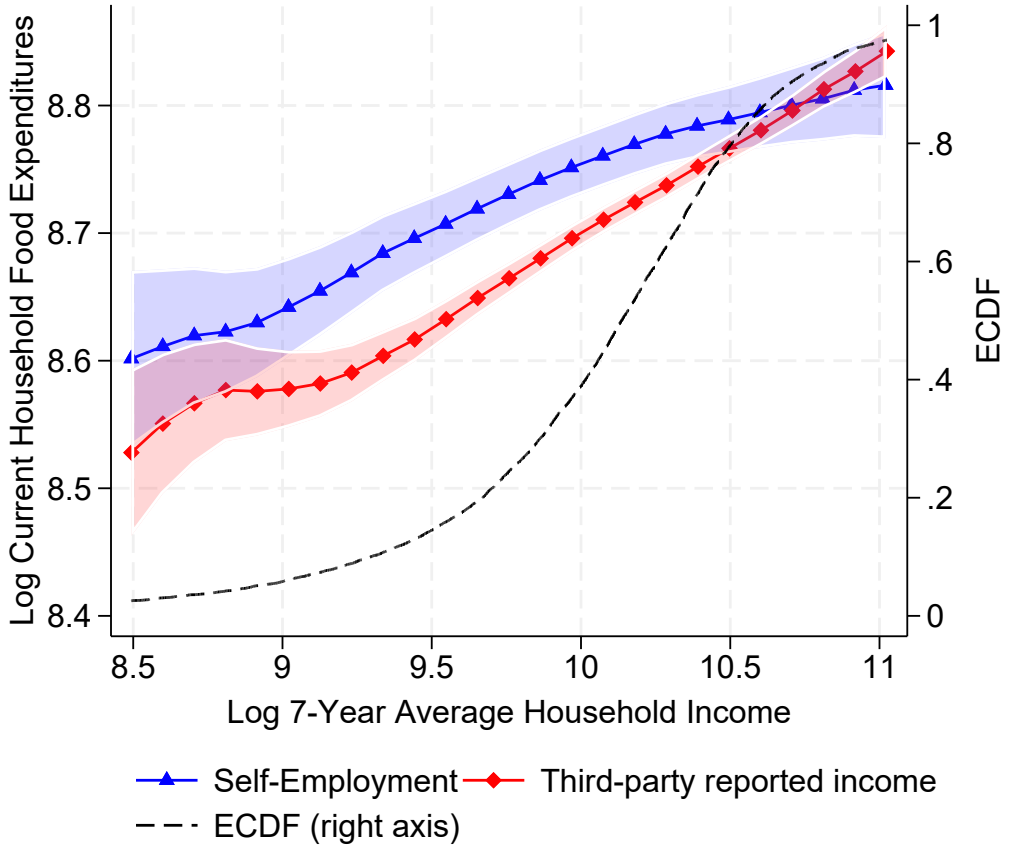
Measurement of self-employed tax evasion is difficult when data from random audits are not available, as in the case of Italy. In this case, researchers have employed alternative estimation methods based on comparing microdata from surveys and tax returns to detect traces of undeclared taxable income (Alm 2012; Nygård and Thoresen 2023; Fiorio and D'Amuri 2005). In the existing literature, moreover, self-employment is usually treated as a homogeneous unit of analysis, by estimating average evasion rates (Engström and Holmlund 2009; Cabral et al. 2021; Kukk et al. 2020), which is reductive for a working category with a high degree of heterogeneity. Indeed, the illicit traits and evasion decisions of shop vendors and craftsmen are likely to differ from those of lawyers and doctors (Levine and Rubinstein 2017).

Leveraging new and unique micro-data that exactly combine survey information on consumption and household characteristics from the Household Budget Survey (HBS) for the year 2013 with detailed administrative data on income, properties, taxes and social security contributions from tax returns over the period 2010-2016, this paper has a twofold objective. First, we provide novel evidence on self-employment income tax evasion in Italy, by applying the expenditure-income estimation approach (Pissarides and Weber 1989; Feldman and Slemrod 2007; Kukk et al. 2020).[2] The idea of this approach is to quantify evasion rates from the comparison of food expenditure-income

[1] We refer to the following countries for which information is available: Australia, Canada, the United Kingdom, and the United States (ATO 2022; CRA 2023; HMRC 2024; IRS 2022).

[2] This approach was firstly proposed for estimating self-employment income tax evasion in Great Britain (Dilnot and Morris 1981; Pissarides and Weber 1989) to overcome the lack of data from random audits. It is now widely applied in several countries (e.g. Hurst et al. 2014; Engström and Hagen 2017; Cabral et al. 2019, 2021; Nygård et al. 2018), and in cross-country settings (Kukk et al. 2020).

Figure 1 Nonparametric estimates of food-expenditure curves, by household type with 95% confidence bands.



Notes: The dashed black line shows the empirical cumulative distribution function (ECDF) of the total population of households.

curves estimated for self-employed, who self-report their income, and wage and salary workers, whose income is reported by a third party, typically the employer, with no incentive for underreporting. From Figure 1, it can be noted that, in Italy, the food expenditure curve estimated for self-employed consistently lies above that for third-party reporting income earners for almost all of the income support, similarly to what is reported for the United States (Hurst et al. 2014).[3]

[3]The opposite seems true only at the top of the income distribution, although the confidence intervals are large, and it involves less than 10% of all income earners (see the dashed line representing the empirical cumulative distribution function of the total population of households). It can also be observed that the food expenditure curve of self-employed households shows larger variation across different , suggesting the presence of heterogeneity in their underreporting behaviour.

Benefiting from the richness of our data, we limit the estimation bias of this approach by constructing a measure of permanent income (Solon 1992; Engström and Hagen 2017), which we obtain as the average income over the years 2010-2016. Then, we study heterogeneity in self-employment income tax evasion, by measuring the evasion rates of specific categories of self-employed, namely small entrepreneurs (i.e., shop vendors, craftsmen, etc.) and liberal professionals (i.e., lawyers, doctors, etc.).[4] The rise of liberal professionals, faster than all other occupations, is worth noting in recent years and is common in Italy, the EU and the US.[5]

Second, we investigate the impact of self-employment income tax evasion on tax revenues and inequality (Alstadsæter et al. 2019), using the evasion rates estimated in the first step of our analysis as input data for a microsimulation model analysis (Figari et al. 2015; Ceriani et al. 2021). We quantify the effects of self-employment income tax evasion on the existing tax-benefit system by explicitly looking at variations in standard measures of the distribution of the tax burden and vertical equity (Nygård et al. 2018). In our empirical analysis, we focus on households for two main reasons. First, distributional analyses are typically conducted at the household level, including when examining the impact of taxation and tax evasion (Alstadsæter et al. 2019). Second, the decision to evade taxes is often not an individual one, but rather a collective one, taken within households (Hashimzade et al. 2021). Self-employed workers, for instance, may have incentives to hire other household members, thereby reducing the tax base and potentially over-reporting employment costs. Alternatively, they may decide to mitigate the risk of self-employment by having other household members take on wage and salary occupations.

The main results of the paper can be summarised as follows. We find that in Italy, the share of unreported income among the self-employed is approximately one-third of their total income. These estimates align with what has been documented for other countries that are generally perceived as having low evasion rates compared to Italy[6], by applying the same methodology. We also show that self-employment income tax evasion is primarily due to liberal professionals, who underreport approximately twice the income of small entrepreneurs, namely 42% versus 23% of unreported income share. For both categories of self-employed workers, we find no clear evidence of heterogeneity in terms of property wealth or income levels. This is in contrast with evidence found using random audit programs, that normally find decreasing evasion with income, but consistent with evidence provided complementing random audit with off-shore, leaked and tax amnesties data: when accounting for sophisticated evasion is possible, tax evasion among the highest-income taxpayers is not lower than that of the rest of the population (Guyton et al. 2021; Alstadsæter et al. 2019).

[4] Specifically, our definition of entrepreneurs is similar to that of unincorporated self-employed (Levine and Rubinstein 2017), who are not separate legal entities as in the case of incorporated self-employed and are subject to personal income tax. In Italy, during the observation period, six out of ten self-employed were unincorporated businesses, which is similar to the situation in the US, according to data from the Bureau of Labor Statistics.

[5] From 2009 to 2019, the number of liberal professionals increased by about 18% and 27% in Italy and the EU, respectively. In the US, lawyers and accountants are projected to grow about 8% from 2021 to 2031, faster than the average for all occupations, according to data from the Bureau of Labor Statistics.

[6] For instance, Cabral et al. (2019) for the United Kingdom, Hurst et al. (2014) for the United States, Cabral et al. (2021) for New Zealand, Engström and Hagen (2017) for Sweden.

To explain the differences in tax evasion between liberal professionals and small entrepreneurs, we use additional information on employees hired by the self-employed and document that, in Italy, professionals are mostly sole proprietors with fewer employees compared to small entrepreneurs.

Our budgetary and distributional results show that the impact of evasion is mainly concentrated at the top of the income distribution: households in the 10th income decile generate over 70% of income tax revenue and 30% of social security contributions losses. We also find positive upward transition probabilities towards higher income decile groups for liberal professionals, once income is adjusted to account for evasion. Our findings suggest that if tax evasion were eradicated, the overall disposable income inequality, as measured by the Gini index, would be approximately 1-2 percentage points lower than its current level.

This study builds on and contributes to the following areas of the empirical public finance literature. Previous research has pointed out that self-employment income underreporting varies across several dimensions, including age, education, gender, and place of residence (Engström and Holmlund 2009; Feldman and Slemrod 2007; Cabral et al. 2021). What is missing is an empirical investigation of self-employment income tax evasion depending on working status (Cullen et al. 2021) and, most importantly, the implications of such heterogeneity on income distribution (Advani 2022). We also integrate the fast-growing literature that assesses the implications of self-employment on income inequality (Auten and Splinter 2019; Kopczuk and Zwick 2020; Alstadsæter et al. 2019), by explicitly looking at the impact of tax evasion on distributional measures (Nygård et al. 2018).

The paper is structured as follows. Section 2 provides a description of the data and summary statistics. Section 3 details the methodology used, including a discussion of how we extend the expenditure-income methodology to account for heterogeneity, and presents the main results. Section 4 presents the distributional analysis. Section 5 concludes the paper. Robustness checks and additional descriptive statistics are in the Appendix.

2 Data and summary statistics

2.1 Dataset construction

For the first time in Italy, we created a unique dataset that links the Italian Household Budget Survey on household consumption with several administrative registers, providing information on incomes, taxes, and properties. Administrative data are the ideal source of information on incomes reported to the tax authorities. The exact matching of all different data was carried out using a unique personal identification code to overcome errors that would occur using statistical matching, improving the informational content of the budget survey (Blanchet et al. 2022).

The Household Budget Survey

The Household Budget Survey (HBS) is run yearly by the Italian National Institute of Statistics to describe and monitor the consumption of Italian households.[7] We use HBS data for the year 2013, which contains information on 278 elementary consumption and self-consumption items. This survey is conducted on a large sample of households (approximately 20,000 annually, comprising around 50,000 individuals) selected using standard sampling procedures to ensure the sample is representative of the entire population of Italian households (Benedetti et al. 2024).

In the estimation of self-employment income tax evasion, our main expenditure variable is the monetary value (in euros) of total food consumption expenditures, which are recorded in the HBS on a daily basis from a diary kept by a member of the household for two weeks. This expenditure category includes information on 55 food items consumed by any household member at and away from home, including self-produced and auto-consumed goods. Food consumption expenditures are useful when applying the expenditure-income estimation approach, as this expenditure category is not influenced by employment status compared to other expenditures (Pissarides and Weber 1989; Hurst et al. 2014; Eika et al. 2020). Other consumption variables have been used when food consumption is not recorded. For example, Feldman and Slemrod (2007) use charitable donations because the Internal Revenue Service tax return data they use do not include information on food consumption. Albarea et al. (2020); Kukk et al. (2020) use household utility expenditure, which is one of the few consumption variables recorded in the European Union Survey on Income and Living Conditions (EU-SILC) they use.

The HBS also contains detailed information on the key demographic and economic variables used in our empirical analysis that include, among other variables, household composition (e.g., single adults vs couples, number of children), the educational level of the household head and the partner (if present), age, the gender of the household head.

Tax returns and property records

Information on incomes and property cadastral values is provided by tax returns and cadastral records, available at the Italian Ministry of Economy and Finance, and linked for all individuals in the HBS sample for each year during the period 2010-2016, meaning from year $t-3$ to year $t+3$, where $t = 2013$. This allows us to measure current income at the year of the HBS survey, as well as to construct a proxy for household permanent income, calculated as the 7-year average of current annual income declared to the tax authorities.

In Italy, tax declarations are filed on an individual basis, and the household composition enters the tax function only through tax expenditures for dependent family members, that is, members of the household with a yearly taxable income below €2,840. We combine tax return data for each member of each household in the HBS sample, defining household taxable income as the sum of individual taxable incomes

[7] Since 1966, the HBS has been used for several purposes, including the construction of national accounts, the measurement of poverty, and the measurement of inflation by household expenditure classes (Rondinelli 2014).

for each member.[8] Moreover, tax data contain information on income categories, taxes, tax credits and allowances, and social security contributions, which we will use for the analysis of the impact of self-employment income tax evasion on the tax-benefit redistribution and inequality (Meyer and Mittag 2019).

Administrative tax records are also used to classify self-employed households, specifically those with a given share of their total taxable income derived from self-employment, and to contrast them with the remaining households, which we refer to as wage and salary households. We exclude from the analysis households with only rental, capital or retirement income, as they are likely to be a different group from households with labour income. The main difference between self-employed workers and wage and salary workers is that the former self-declare their income, while the latter have their income declared by a third party, typically their employer, who has no incentive to underreport it to the tax authorities.[9] This classification is possible because labour incomes are taxed through different tax forms depending on their type. The details provided by administrative tax records also allows us to investigate the heterogeneity of income tax evasion among self-employment income earners, as disentangling self-employment income of liberal professionals (that is, individuals who provide specialised intellectual services based on their professional qualifications and skills, such as doctors, engineers, lawyers, accountants) and of small entrepreneurs (that is, individuals who start and run their small businesses, such as craftsman, shop vendors, artisans).

From tax returns, we use information on the province of residence of each taxpayer to construct clustered standard errors. Additionally, using cadastral property values included in cadastral records, we develop a measure of household property wealth (Cooper 2013).

2.2 Descriptive statistics

In the tax data, we have 8.2% of households whose income comes entirely from self-employment and 20.7% that have at least some self-employment income. The literature using the consumption-income approach to estimate income tax evasion often defines self-employed households as those that receive at least 25% of their total income from self-employment. An alternative approach to classifying self-employed households is to use the self-declared occupational status of the householder, as reported by individuals in the HBS. In Figure 2, we plot the proportion of households using the self-declared status of the household head (blue line) as opposed to using different shares of total income (from 5% to 95%, in steps of 5%, the red line in the figure). The latter decreases from a maximum of 24% to a minimum of 12%, and it equals 21.9% when the share of income from self-employment is 25%. When self-employed households are defined as those whose share of self-employment income is 25%, the proportion of self-employed households is not statistically different from the share obtained by self-declaration, which encourages us to use this threshold in our main empirical analysis.

[8] The anonymous matching procedure was produced for this project by the Italian Ministry of Economy and Finance in full compliance with privacy protection regulations.

[9] Although wage and salary workers may receive undeclared income, as in the case of unrecorded after-hours work, the assumption of no evasion for this category of workers is common in the literature for identification purposes (Feldman and Slemrod 2007).

Moreover, defining self-employed households as those with 25% of total income from self-employment allows us to compare the results for Italy with those produced using the same methodology in other countries. Some additional descriptive statistics and robustness analysis of our main results using different shares to classify households as self-employed is, however, produced in Appendix A and B, respectively.

Figure 3 shows the distribution of household types by quintile of total households (left panel) and the distribution of the population by household types, using the 25% threshold to classify self-employed households (right panel). The former shows that in the first 20% of the income distribution over half of the households have no income from work, i.e. is mainly comprised by retired people; about 10% of third-party-reported income households are in the first quintile and their frequency increase up to nearly 30% at the largest quintile; finally, over 30% of self-employed households belong to the highest quintile of income, and 15% of them belong to the first three quintiles. The right panel shows that labour income is the main source of income for over 61% of all households. We excluded the households with no labour income, and the final linked sample comprises 12,166 households, corresponding to 35,189 individuals.

Tables 1 and 2 show summary statistics for the main variables used in the empirical analysis for the whole sample of labour income households, distinguishing between the self-employed and third-party reporting households. Food consumption is, on average, higher for self-employed individuals than for households reporting third-party income (Panel A). In contrast, self-employed households declare on average less income than the rest of the sample (Panel B). The latter slightly increases when the 7-year average income measure is used instead. We also show descriptive statistics for other household characteristics used in the empirical analysis. Consumption of sin goods (e.g., tobacco and spirits) is used as a proxy to measure risk-taking behaviour (Barsky et al. 1997). Self-employed households are, on average, larger in size, more likely to be in a couple, and reside in the North of the country. They are better educated and are less likely to have a female head.

When examining the income distributions, some differences emerge. The income distribution for self-employed households is less concentrated, with a larger share of low-income households (who declare less than €20,000), and a longer right tail, with a share above €150,000 that is three times larger than that of wage and salary workers (Figure 4). Among self-employed households, 55.7% have income only from self-employment, while the remaining 44.3% also report earnings from wages and salaries.

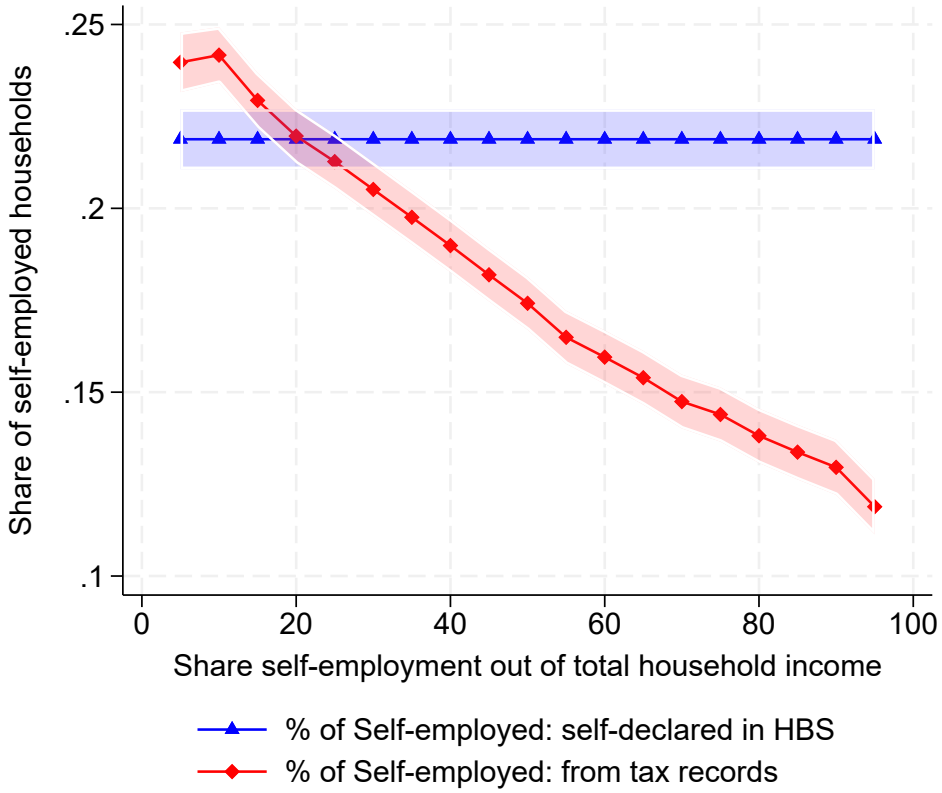
3 Income underreporting estimation

3.1 The homogeneous case

Our starting point is a food expenditure equation, like in most of the previous studies (e.g. Pissarides and Weber 1989; Hurst et al. 2014; Kukk et al. 2020), which describes the relation between log-income and log-consumption for a generic household i, namely:

$$lnc_i = \alpha + \beta lny_i + X_i^{'}\theta + \epsilon_i, \tag{1}$$

Figure 2 Households self-employment status



where lnc_i is the log-consumption of household i, lny_i is its log-permanent true income, and the vector X_i denotes a set of controls. The coefficient α is a constant term, β is the income elasticity of consumption, and ϵ_i is a composite error term with mean zero and variance equal to σ_ϵ^2.

In this framework, the first assumption is that the consumption of selected expenditures does not depend on the type of employment of workers (i.e., wage and salary workers versus self-employed individuals), after controlling for observable characteristics (Pissarides and Weber 1989). Not all consumption items are suitable for fulfilling this assumption. The self-employed may consume more fuel as they are more likely to visit clients by car, or spend more on elegant clothes and on durable goods, such as cars, to impress them. The best consumption item to satisfy this first assumption is food, which is typically recorded in household budget surveys over a set of consecutive days. However, as HBS also includes information on home utilities expenditure, such as home bills for heating, electricity, telephone, water, condominium expenses,

Figure 3 Distribution of household types by income quintiles (left panel) and over the whole sample (right panel)

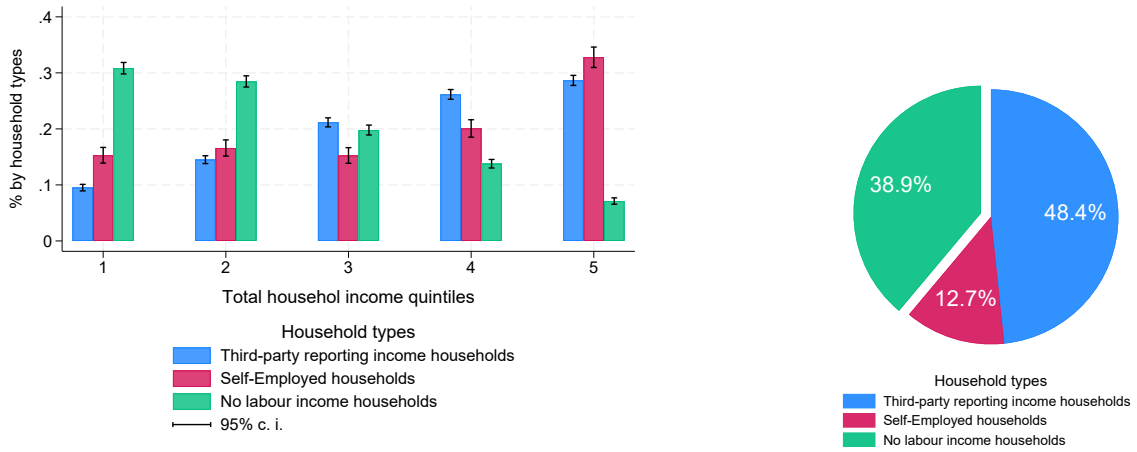


Table 1 Descriptive statistics, main variables.

	Whole sample	**Self-employed**	**Third-party reporting income**
	A. Log average expenditure measures		
Food expenditures	8.722	8.760	8.712
	(0.570)	(0.570)	(0.569)
Utilities expenditures	7.864	7.923	7.849
	(0.684)	(0.673)	(0.686)
	B. Log average income measures		
Pre-tax Total Household Income:			
current	10.197	10.051	10.235
	(0.974)	(1.205)	(0.901)
7-year average	10.228	10.106	10.259
	(0.867)	(1.059)	(0.808)
Sample size	12166	2465	9701

Notes: our calculation on the selected sample, standard errors in parentheses. Self-employed households are identified as those with self-employment income equal to or larger than 25% of total household income.

and ordinary and extraordinary maintenance expenses, we also test whether results would change if home utilities consumption were used instead of food consumption.[10]

The second key assumption is that self-employed workers can evade labour income tax, whereas wage and salary workers cannot: typically, self-employed workers self-declare their income, whereas wage and salary workers have their income reported by a third party, usually their employer. Self-employed workers cannot simply declare their income however they like. They must agree with the client not to register a transaction and be paid in cash to avoid income tax, thereby allowing the customer to evade value-added tax. However, once the transaction is registered, the received income must be reported. Wage and salary workers may receive additional income for unrecorded after-hours work; however, all income paid by the employer, as stipulated in the contract, is recorded and taxed accordingly. This assumption is widely accepted in the economic literature (e.g., see Feldman and Slemrod 2007; Kleven et al. 2011). It also implies that we can partition the population of households into two types, $j = W, S$, where W refers to wages and salaries (or third-party reported income) households, and S refers to self-employed households. We can denote income as y_{ij}^t, where the superscript, type t, can be R for reported, T for true, U for undeclared, so that $y_{ij}^T = y_{ij}^R + y_{ij}^U$. Hence, our second assumption implies that reported income for type-W households equals true income ($y_{iW}^R = y_{iW}^T$); conversely, reported income for type-S households can be lower than true income as they can decide to underdeclare part of it ($y_{iS}^R < y_{iS}^T$, if $y_{iS}^U > 0$). Let the average share of the true income undeclared by type-S households (y_S^U / y_S^T) be $0 \leq s \leq 1$, following Hurst et al. (2014), we can rewrite (1) as:

[10]Home utilities expenditures have been used when food consumption is not available, as in the European Survey on Income and Living Conditions (EU-SILC), Paulus (e.g. 2015); Albarea et al. (e.g. 2020).

Table 2 Descriptive statistics, other control variables.

	Whole sample	**Self-employed**	**Third-party reporting income**
	C. Control variables (continuous)		
Average household size	2.833	2.919	2.810
	(1.277)	(1.301)	(1.270)
Sin goods: Tobacco Products	2.116	2.123	2.115
	(3.117)	(3.127)	(3.114)
Sin goods: Lottery Tickets	0.755	0.687	0.773
	(1.930)	(1.862)	(1.947)
Sin goods: Spirits	0.608	0.594	0.612
	(1.790)	(1.776)	(1.793)
Number of rooms at home	4.171	4.438	4.103
	(1.417)	(1.545)	(1.374)
	D. Control variables (categorical)		
% households with kids	0.364	0.348	0.368
% female-headed households	0.255	0.198	0.270
% households in couple	0.659	0.700	0.649
Head age: % 36-50	0.429	0.461	0.421
Head age: % 51-65	0.351	0.352	0.351
Head age: % 66 and over	0.109	0.104	0.111
Head education: % with secondary	0.571	0.591	0.566
Head education: % with college	0.274	0.270	0.274
Spouse education: % with secondary	0.386	0.429	0.374
Spouse education: % with college	0.585	0.613	0.577
% resident in the Center	0.204	0.182	0.209
% resident in the South	0.296	0.276	0.301
Dwelling: % high-end apartment	0.097	0.116	0.092
Dwelling: % low-end apartment	0.787	0.730	0.802
Dwelling: % rural units	0.028	0.032	0.027
Building: % in condominium	0.705	0.678	0.712

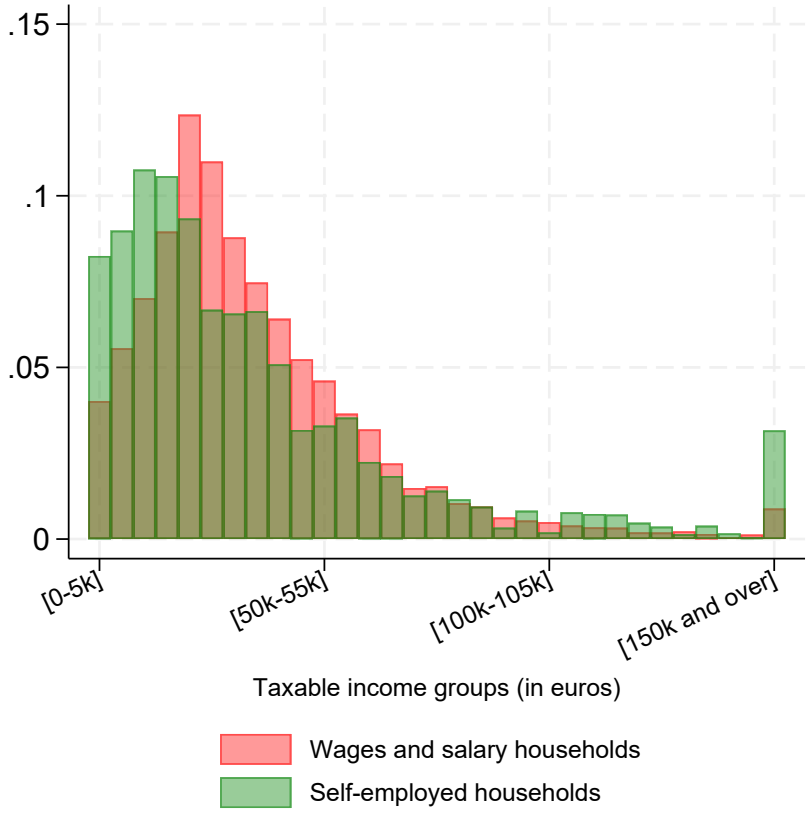
Notes: as in Table 1.
Omitted variables are: % male-headed households, % households without kids, Dwelling: % detached, Building: % single-family house, Head age: 35 and below, Head education: % with primary, % resident in the North, % households not in couple, Spouse education: % with primary

$$lnc_i = \alpha + \beta lny_i^R + \gamma I_j + X_i^{'}\theta + \epsilon_i, \tag{2}$$

where I_j is a dichotomous variable that is equal to one if $j = S$ and zero otherwise, so that γ measures the difference of log consumption between type-S and type-W households. As $\gamma = [-\beta ln(1-s)]$, the estimation of equation (2) allows for the identification of the average share of the true income reported by type-S households, $\kappa := (1-s) = [(y_S^T - y_S^U)/y_S^T] = (y_S^R/y_S^T)$. The estimate $\widehat{\kappa} = exp(-\widehat{\gamma}/\widehat{\beta})$ is then used to estimate $\widehat{s} = 1 - \widehat{\kappa}$, the average share of the true income that is unreported by the self-employed.

The primary empirical challenge in estimating consumption-income relations is accurately measuring household income to minimise transitory fluctuations and mitigate the attenuation bias caused by current income (Solon 1992). Previous research

Figure 4 Distribution of taxable income for self-employed and wage-and-salary households, in taxable income groups by steps of €5,000.



has employed an IV approach or approximated permanent income by averaging current income over several years to mitigate the bias caused by temporary fluctuations in current income. However, both strategies have flaws. On the one hand, an average income over a long period is a better proxy for permanent income, the longer the period considered; however, it may still incorporate some temporary fluctuations. On the other hand, finding an IV that convincingly satisfies the exclusion restriction is not easy, as the proposed instruments often correlate with income and consumption. For instance, Hurst et al. (2014) uses education as an instrument for current income, although there is plenty of evidence that education is also correlated with consumption; Engström and Hagen (2017) uses capital income as an instrument for household current income in consumption-income estimates, which is a better alternative than education, although, dividends and interests are often paid into the capital owner's current account and could be easily mistaken by the recipients as labour income.

In this paper, we do not use an IV approach because we lack a suitable instrument, we do not have information on households' capital income, and the education of the

household head cannot reasonably be considered to fulfil the exclusion restriction. However, by exploiting the panel dimension of our tax data, we use the average income to reduce temporary fluctuations and measurement errors over 7 years, which is a rather long period, corresponding to nearly 20% of an average working life.

In Table 3, we report the OLS estimates of equation (2), using as a proxy of permanent income either the current pre-tax income or the 7-year average pre-tax income and self-employment income tax evasion rates. We present the estimation of model (2) using either household expenditure on food or home utilities as dependent variables. A complete set of controls was included, namely the household head's age and gender, an in-couple dummy interacted with the education attainment (primary, secondary, or tertiary) of the spouse, household size, a dummy for the presence of children, total consumption of sin goods, a complete set of macro area of residence dummies, household head education and housing characteristics. Standard errors were adjusted for 109 clusters in the province of the household's residence. Defining type-S households as those with at least 25% of their income from self-employment, we find that the rate of self-employment income underreporting using food expenditure is 33% using current income and 29% using 7-year average income, showing that reducing the effect of transitory fluctuations slightly reduces the estimate of income underreporting. We also show that estimating using the home utility expenditure as the dependent variable yields similar evasion rates. Notably, estimates using home utilities as the consumption variable are only slightly smaller, but not statistically different from estimates using food consumption, which supports its use when food consumption data are not available.

Interestingly, these estimates align with those obtained for other developed countries, with the same methodology and main assumptions. In Sweden, using mostly administrative data, tax evasion is estimated between 15% and 30%, using a 7-year average income as we do to address endogeneity of the current income (Engström and Hagen 2017; Engström and Holmlund 2009); in the US tax evasion is estimated between 25-30%, using different data sources (the Consumer Expenditure Survey and the Panel Study of Income Dynamics) (Hurst et al. 2014); in the UK the very large tax evasion of self-employment income, ranging between 50 and 60%, originally estimated by Pissarides and Weber (1989) using the 1982 Family Expenditure Survey, was estimated around 30% with the same data 10 years later Lyssiotou et al. (2004), averaging between blue and white collars, recently updated at around 20% by Cabral et al. (2021) using a different dataset (the 2010-12 Living Costs and Food Survey). Similar results are found also for Spain (Martinez-Lopez 2013), for Norway (Nygård et al. 2018), for Finland (Johansson 2005), for Canada (Schuetze 2002), for New Zealand (Cabral et al. 2021).

To test the robustness of our homogeneous tax evasion estimates, lacking a convincing instrument, we first examine whether our results would change if type-S households were defined using lower or higher shares of household income from self-employment than 25%. Figure 5 plots the estimation of equation (2) using both current income and the average over a 7-year period of current income to proxy permanent income; we identify type-S households as those with a share of income from self-employment ranging from a minimum of 5% to a maximum of 95%. Estimates using food and home

Table 3 Self-employment income tax evasion estimates.

	Food expenditures		**Home utilities expenditures**	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Current income	0.085***		0.096***	
	(0.009)		(0.010)	
Average income (7-year)		0.109***		0.127***
		(0.011)		(0.009)
Self-Employed	0.034**	0.038**	0.036**	0.041**
	(0.015)	(0.015)	(0.018)	(0.018)
R-squared	0.199	0.203	0.200	0.204
N. households	12166	12166	12166	12166
N. obs Self-Employed	2465	2465	2465	2465
Share Self-Empl. income	0.758	0.758	0.758	0.758
Evasion rates $(1-\kappa)$				
Self-Employed	0.329***	0.291***	0.313**	0.277***
	(0.110)	(0.090)	(0.126)	(0.100)

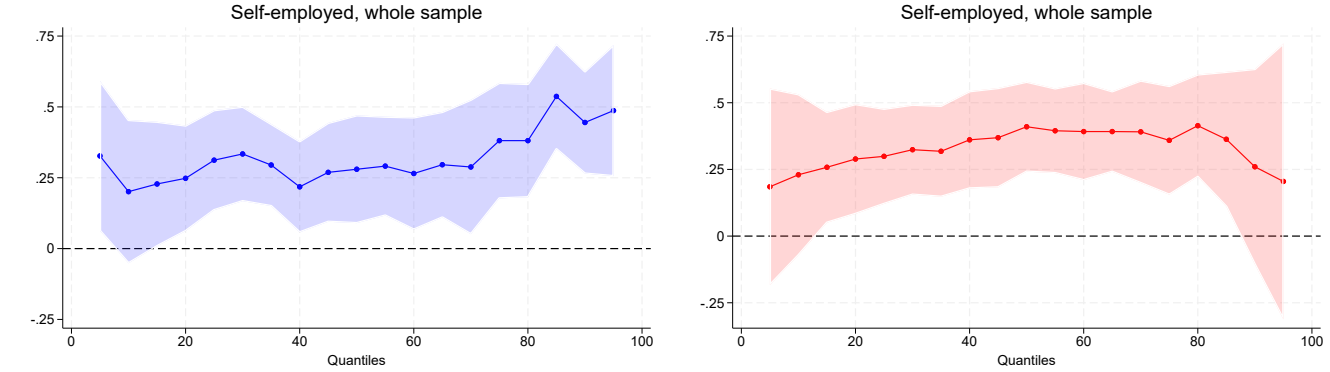
Notes: A self-employed household has at least 25% of its income from self-employment. Controls include the household head's age and gender, an in-couple dummy interacted with the spouse's education (primary, secondary, or tertiary), household size, a dummy for the presence of children, family consumption of sin goods, a full set of macro area-of-residence dummies, the household head's education, and housing characteristics. Standard errors are adjusted for 109 clusters in the province of family residence.

utilities consumption are provided on the left and right panels, respectively. It shows that using a 7-year average measure of permanent income provides a slightly lower tax evasion estimate with lower standard errors. The tax evasion rate slightly increases if a larger share of self-employment income is used to define households as self-employed, but the differences are not statistically significant. The black dashed line shows the share (on the right axis) of self-employed households within the total population for varying shares of self-employment income to classify households as self-employed. It shows that 25% of households have at least 5% of their income from self-employment, and this proportion decreases to a minimum of 8% when 95% of their income is from self-employment. Estimates are more precisely estimated and significantly different from zero when food consumption is used instead of home utilities.

Additionally, we also test the equation (2) using quantile regressions for a range of log-consumption quantiles that go from a minimum of 5 to a maximum of 95, holding the 25% share of self-employed household income to define type-S households. Figure 6 plots estimates using 7-year average income only as control, showing estimates for food consumption and home utilities on the left and right panels, respectively. The estimates of tax evasion for consumption quantiles up to the 70th are remarkably stable at around 25%, similarly to what we obtained previously for the average consumption. It largely increases only for households with high levels of expenditures, namely those in the top 30% of the consumption distribution.

Figure 5 Estimates of tax evasion and share of self-employed households using different percentages of household income that is from self-employment.

Figure 6 Quantile regressions, Food expenditures (left figure) and Home utilities expenditures (right figure).



Notes: With 95% confidence intervals.

3.2 Investigating the heterogeneity of tax evasion

Leveraging the richness of our administrative data, we investigate possible heterogeneities in tax evasion behaviour to shed some light on why tax evasion is so relevant in Italy. Here, we distinguish liberal professionals from all other self-employed workers based on the tax form they filed.

Self-employed individuals are those who work for themselves rather than being employed by a company or organisation. They run their businesses or offer their services directly to clients. Self-employment can span a wide range of professions and industries including freelancers (i.e., writers, graphic designers, and consultants who offer their services on a project-by-project basis), small business owners (e.g., individuals who own and operate their businesses, such as retail stores, restaurants, or service-based companies), independent contractors (e.g., professionals who provide services to other businesses or individuals under contract, such as construction workers, electricians, and plumbers) or liberal professionals (e.g., lawyers, doctors, architects, and accountants who run their practices, requires extensive education, specialized knowledge, and adherence to strict ethical standards).

We define two dummy variables I_p and I_e.[11] We then rewrite (2) as follows:

$$lnc_i = \alpha + \beta lny_i^R + \gamma_p I_p + \gamma_e I_e + X_i'\theta + \epsilon_i \tag{3}$$

and the estimated shares of true income now become $\widehat{\kappa_p} = exp\left(-\frac{\widehat{\gamma_p}}{\widehat{\beta}}\right)$ and $\widehat{\kappa_e} = exp\left(-\frac{\widehat{\gamma_e}}{\widehat{\beta}}\right)$ for professionals and entrepreneurs, respectively.

Within the two categories of self-employed, one can be interested in investigating further dimensions of heterogeneity based, for instance, on some observed characteristics I_l (e.g., a subset of the control set X_i), which led to ℓ mutually exclusive groups. We then rewrite the relation (3) as follows:

$$lnc_{ij} = \alpha + \beta lny_{ij}^R + \sum_{\ell=1}^{L} \gamma_{p\ell} I_p I_\ell + \sum_{\ell=1}^{L} \gamma_{e\ell} I_e I_l + X_i'\theta + \epsilon_{ij} \tag{4}$$

with the estimated shares of true income now being $\widehat{\kappa_{p\ell}} = exp\left(-\frac{\widehat{\gamma_{p\ell}}}{\widehat{\beta}}\right)$ and $\widehat{\kappa_{e\ell}} = exp\left(-\frac{\widehat{\gamma_{e\ell}}}{\widehat{\beta}}\right)$ for professionals and entrepreneurs, respectively.

Table 4 presents results of the heterogeneity analysis using the 25% share to identify self-employed households, as before, and focusing on food expenditure only as it is more precisely estimated.[12] The first two columns present the estimation of heterogeneity as in equation (3), and the last two columns show the analysis of heterogeneity by further decomposing the population by cadastral wealth, as in equation (4). Cadastral wealth is a large component (more than 50%) of total household wealth, and can be used as a proxy of total household wealth, which we do not measure.The heterogeneity by

[11] Wage and salary workers fill in tax form RC, liberal professionals tax forms RD, RE or L, and small entrepreneurs fill in tax form RG. Partnerships fill tax form RH. Liberal professional partnerships are a small proportion (5%) of all total partnerships (see data from MEF tax statistics, available at https://www1.finanze.gov.it/finanze/pagina_dichiarazioni/public/dichiarazioni.php) and we therefore include them in the entrepreneurs category.

[12] Some robustness analyses will be provided in the Appendix B.

type of self-employed occupation shows significant coefficients in the food-expenditure equation and a share of income tax evasion that is twice as large for liberal professionals as for entrepreneurs. In the last two columns, we found no clear evidence of significant heterogeneity by cadastral wealth. Some heterogeneity among entrepreneurs emerges only for high-wealth households.

Table 4 Self-employment income tax evasion estimates, heterogeneity analysis

	Food expenditures			
	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Current income	0.086***		0.085***	
	(0.009)		(0.009)	
Average income (7-year)		0.110***		0.109***
		(0.011)		(0.010)
Liberal professionals	0.062**	0.060**		
	(0.025)	(0.024)		
Entrepreneurs	0.024	0.029*		
	(0.017)	(0.017)		
Lib. prof. × Cadastral wealth below med.			0.061	0.060
			(0.038)	(0.037)
Lib. prof. × Cadastral wealth above med.			0.062	0.060
			(0.034)	(0.033)
Entrepreneurs × Cadastral wealth below med.			0.013	0.024
			(0.021)	(0.022)
Entrepreneurs × Cadastral wealth above med.			0.034	0.034
			(0.020)	(0.020)
R-squared	0.200	0.203	0.200	0.202
N. households	12166	12166	12166	12166
	$(1-\kappa)$			
Liberal professionals	0.513***	0.421***		
	(0.143)	(0.131)		
Entrepreneurs	0.242*	0.235**		
	(0.139)	(0.111)		
Lib. prof. × Cadastral wealth below med.			0.512**	0.423**
			(0.220)	(0.199)
Lib. prof. × Cadastral wealth above med.			0.518***	0.420**
			(0.194)	(0.176)
Entrepreneurs × Cadastral wealth below med.			0.140	0.199
			(0.207)	(0.152)
Entrepreneurs × Cadastral wealth above med.			0.331**	0.268**
			(0.146)	(0.127)

Notes: as in Table 3.
Cadastral wealth medians are computed using total cadastral wealth.

3.3 Interpretation of the results

Our main findings can be interpreted by looking at specific channels of influence of different evasion behaviours among taxpayer categories. Liberal professionals can conceal more income than small entrepreneurs, as they can engage in more collaborative evasion activities with clients (Doerr and Necker 2021). Moreover, liberal professionals (small entrepreneurs) are likely to be less (more) constrained by the need to deduct declared input costs, which are on average higher for small entrepreneurs. This implies that, all other things being equal, small entrepreneurs declare a larger share of their true income to the tax authority than liberal professionals (Buettner et al. 2023).

Another possible explanation for the high evasion rates can be related to the negative relation between firm size and tax evasion. Other things held constant, there is a trade-off between large-scale activities and tax evasion, that is, tax evasion decreases as firm size increases (Alm et al. 2019; Barth and Ognedal 2018).

Larger taxpayers may also opt for tax avoidance rather than tax evasion (Gumpert et al. 2016). Indeed, large taxpayers, who typically employ a large number of dependent employees, can reduce their tax burden through tax reliefs and tax planning strategies allowed by tax rules (Davies et al. 2018). In addition, small self-employed are more likely to use cash in their transitions than large taxpayers.

Tax audits are more effective when targeted at larger taxpayers, to whom tax administration addresses more resources (Bergolo et al. 2023) and design specific auditing strategies for them, including tailored compliance tools like cooperative compliance schemes, which are not available to small-scale taxpayers (Boeri et al. 2020). Finally, and importantly, small taxpayers are less able to evade tax controls (Basri et al. 2021; Almunia and Lopez-Rodriguez 2018; Carrillo et al. 2017)), as they cannot benefit from complex tax planning strategies as large taxpayers do.

Unfortunately, the data at our disposal do not include information on the number of employees hired by self-employed workers or additional details on firm size. To address this issue, we complement the data at our disposal by using four waves of the Survey on Household Income and Wealth (SHIW), conducted by the Bank of Italy, and restricting the analysis to the self-employed. This approach distinguishes liberal professionals from entrepreneurs, spanning the period from 2010 to 2016. These data indicate that in Italy, approximately 66% of self-employed individuals have no employees, and over 95% of them have fewer than five employees. We estimate a log-linear regression model for the number of employees employed by self-employed workers, using a dummy variable to distinguish between liberal professionals and entrepreneurs, year dummies, and controls for region of residence, a second-degree polynomial in age, and education. The dummy for liberal professionals is estimated with a coefficient of 0.18 (standard error of 0.04), suggesting that, on average, liberal professionals have 20% fewer employees than entrepreneurs. These results suggest a more complex structure at higher income levels among self-employed households.

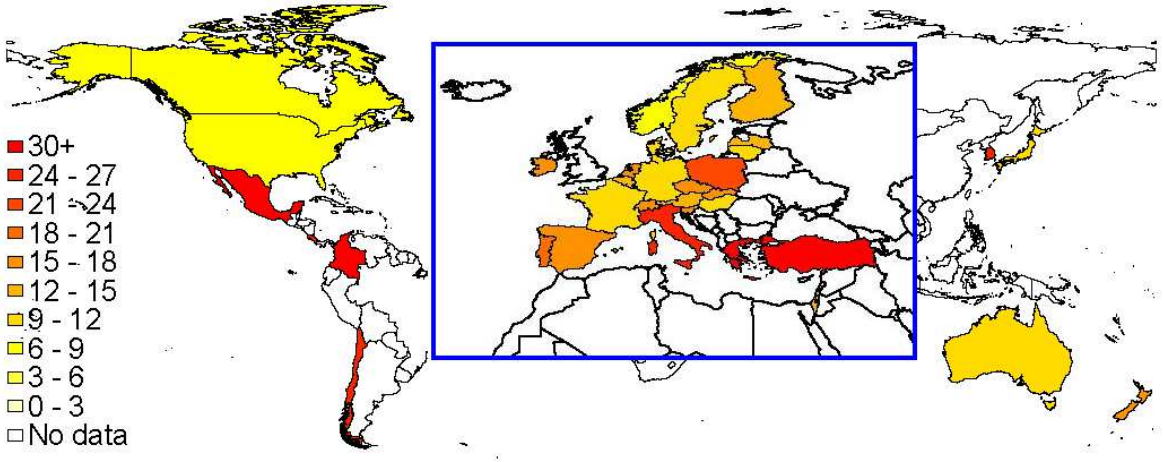
These results point to an explanation of tax evasion that goes beyond third-party reporting of taxable income (Kleven et al. 2011) or individual characteristics and social norms (Bergolo et al. 2022) and is consistent with traditional theoretical literature (Allingham and Sandmo 1972), suggesting that it is rational to evade taxes when

allowed to and suggests that income tax evasion is more pronounced among self-reporting taxpayers with simpler organizational structure. In other words, the smaller the organisation of the economic activity, the higher the chances of hiding some income, consistent with the theoretical model outlined by Kleven et al. (2016).

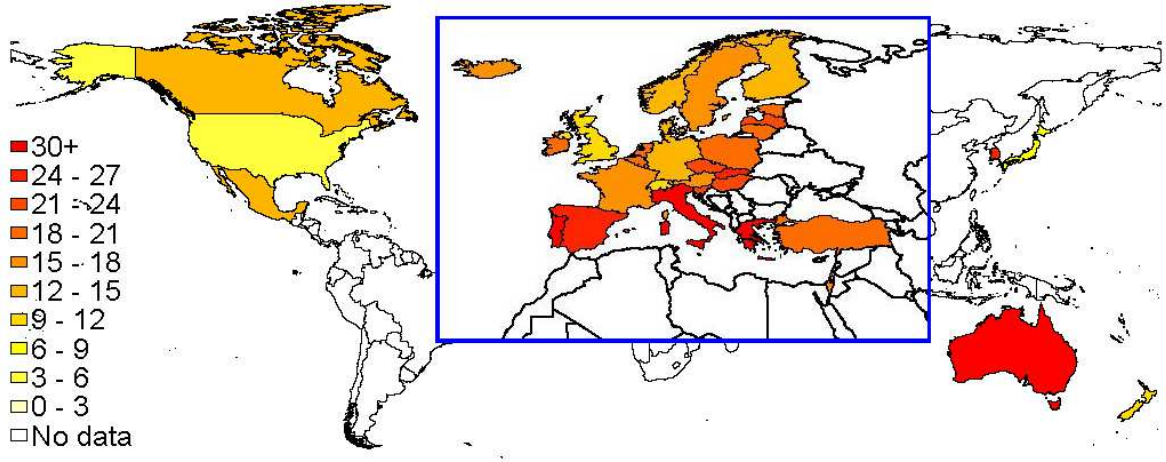
This explanation also suggests an understanding of the large size of tax evasion in Italy pointing at the structure of the economy, with a large share of self-employment as a percentage of total employment (Figure 7, panel (a)) and a large share of small enterprises (below ten employees) per 100 workers (Figure 7, panel (b)), where more room for informality is allowed. Compared to the EU, in particular, Italian self-employed is characterised by higher share of 'solo' self-employed workers without employees than in Europe (80% versus 69%).

Figure 7 Share of self-employed workers and share of small enterprises (less than 10 employees) per 100 employees.

Panel (a) Self-employment as % of employment



(b) People employed in enterprises (1-9 emp.) per 100 employees



Note: Our elaborations are based on OECD data. Average values for the years 2012/2016. The blue box was enlarged for ease of readability.

4 Budgetary and distributional analysis

Tax evasion has relevant budgetary and distributional effects (Nygård et al. 2018; Alstadsæter et al. 2019), affecting both personal income tax (IRPEF) and Social Security Contribution (SSC) revenues. The tax evasion rates estimated in the first part of the paper, taking into account the heterogeneity of the self-employed in terms of working status and cadastral assets, are used here with a twofold objective. First, to assess the budgetary impact of tax evasion by the self-employed by quantifying the loss of revenues associated with evaded taxes and social security contributions. Second, to quantify the effect of tax evasion by the self-employed on horizontal and vertical equity, highlighting the extent to which tax evasion affects average tax rates and income distribution, undermines the redistributive effects of the tax-benefit system, and disproportionately favours some taxpayers.

To achieve these objectives, we develop a set of counterfactual scenarios using EUROMOD, an established tax-benefit microsimulation model.[13] We use the static version of the model, meaning that simulations are performed assuming no changes in the labour supply behaviour of taxpayers across the different scenarios of tax compliance. We adjust income values for the years 2013 to 2018 by using validated updating indexes and policy rules that refer to 2018. Our starting point is a 'reported income scenario' where we consider only income reported to the tax authorities. We then construct a 'true income scenario' by adjusting household self-employment income using the evasion rates presented in the previous section[14], namely estimates from Table 4, last column.

In the third 'evasion-adjusted' scenario, we assume that self-employment income tax evasion is fully discovered and household income is computed under full tax compliance.

Tax evasion by self-employed persons leads to a loss of income tax and SSC revenue of about 40 billion euros in 2018, which corresponds to approximately 10% of total revenue. This is in line with the estimates of the Ministry of Economy and Finance MEF (2024) using a top-down approach based on the comparison of aggregate data from tax statistics and national accounts.

Ranking taxpayers using the OECD-adjusted equivalized-disposable income definition in the 'evasion-adjusted' scenario,[15] the revenue losses are unevenly distributed across taxpayers, as reported in Figure 8. The distribution of shares of income tax (blue bars) and SSCs (red bars) lost revenues by income decile is right-skewed, with over 70% of total tax revenue losses concentrated in the top income decile group. The distribution of losses from SSCs is less concentrated at the top decile due to the

[13] This model is widely used in academic and policy-related analyses to assess the distributional impact of tax-benefit measures in the EU (Sutherland and Figari 2013).

[14] As we define self-employed households based on the 25% share of total income, we can have some type-S households that also earn third-party reported income. This requires using a factor to inflate reported self-employment income for a household of type-S, without affecting their share of third-party reported income, which is assumed to be correctly reported. For type-W households that earn some self-employment income, we assume that they do not conceal taxable income and/or the amount concealed is negligible, as the median share of self-employment income for type-W households in our sample is zero, and the average share is 1.5%.

[15] Ranking taxpayers by reported income would lead to a biased estimates of the rate of tax evasion by income levels because the act of underreporting moves taxpayers down the distribution of reported income (Guyton et al. 2021).

combination of proportional SSC rates and maximum contribution rules. The concentration of lost revenues caused by income tax evasion at the top of the distribution is due to the uneven distribution of self-employment income across taxpayers (recall Figure 4), the upward distributional shift of taxpayers once evaded income is taken into account, and the overall progressivity of the tax system. The latter is responsible for the uneven gains of income concealment. This affects the horizontal equity of the tax system. Taxpayers with a third-party declared income above the median face an average tax rate that is 5 percentage points higher, with the difference reaching almost 20 percentage points in the top decile with respect to the reported income of the self-employed. In a full tax compliance scenario, the tax burden is more evenly distributed across taxpayer types, with low-income self-employed having a higher tax burden than other taxpayers due to social contributions (Table 5).

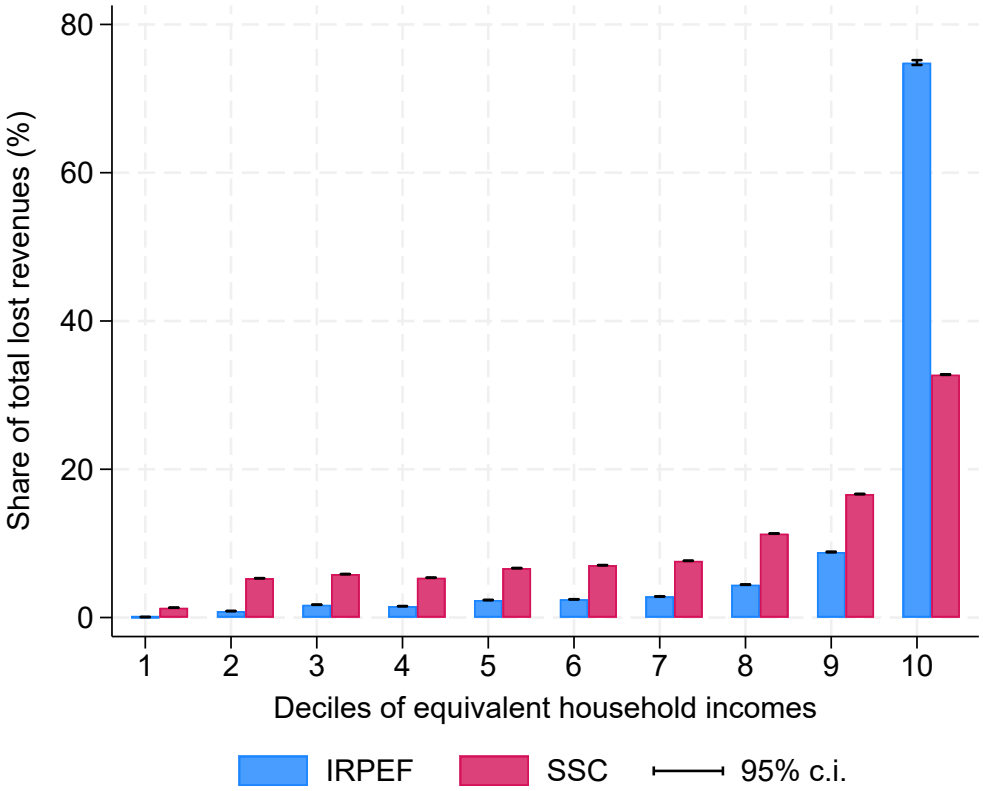
Table 5 Horizontal equity - Average tax rates under different scenarios, year 2018

	Wage and salaries		Self-employment income				
			Current income		7-year av. Income		
Deciles	Reported	%	Reported	Evasion-adjusted	Reported	Evasion-adjusted	%
1	0.149	10.56	0.225	0.240	0.184	0.233	9.54
2	0.195	10.54	0.265	0.285	0.225	0.285	9.44
3	0.212	10.39	0.270	0.291	0.218	0.296	9.66
4	0.240	10.37	0.286	0.321	0.236	0.321	9.63
5	0.264	10.38	0.296	0.343	0.249	0.352	9.27
6	0.339	10.04	0.305	0.375	0.264	0.374	8.74
7	0.386	10.22	0.320	0.392	0.287	0.393	7.44
8	0.419	10.05	0.325	0.415	0.291	0.407	7.13
9	0.457	9.81	0.315	0.436	0.299	0.436	9.32
10	0.509	7.63	0.308	0.474	0.317	0.475	19.84
Total	0.315	100.00	0.291	0.365	0.264	0.372	100.00

Notes: Averages of individual tax rates by equivalent income deciles, with relative frequencies of wage and salary workers and self-employed individuals in each decile. Decile groups based on taxable income in the evasion-adjusted scenario, taking into account evasion rates obtained from OLS estimates using 7-year-average income as in the last column of Table 4.
Source: authors' elaboration based on EUROMOD tax simulator.

Another issue hindering the horizontal equity of the system is the significant re-ranking of income distribution caused by tax evasion. In Figure 9, we show the joint probability distribution of household reported income (horizontal axis) and evasion-adjusted income (vertical axis) by self-employed working status and income level (i.e. below and above the median cadastral wealth), grouped by disposable income decile. The sum of frequencies of all the cells in each matrix adds up to 1, and in the case of no re-ranking, frequencies should lie only along the diagonal. Positive frequencies above the diagonal mean that evasion-adjusted income belongs to a higher decile than reported income, and vice versa. For the categories of self-employed workers for which we find positive and significant income tax evasion rates (panels (a) and

Figure 8 Distribution of the share of tax and SSCs revenues lost due to tax evasion, year 2018

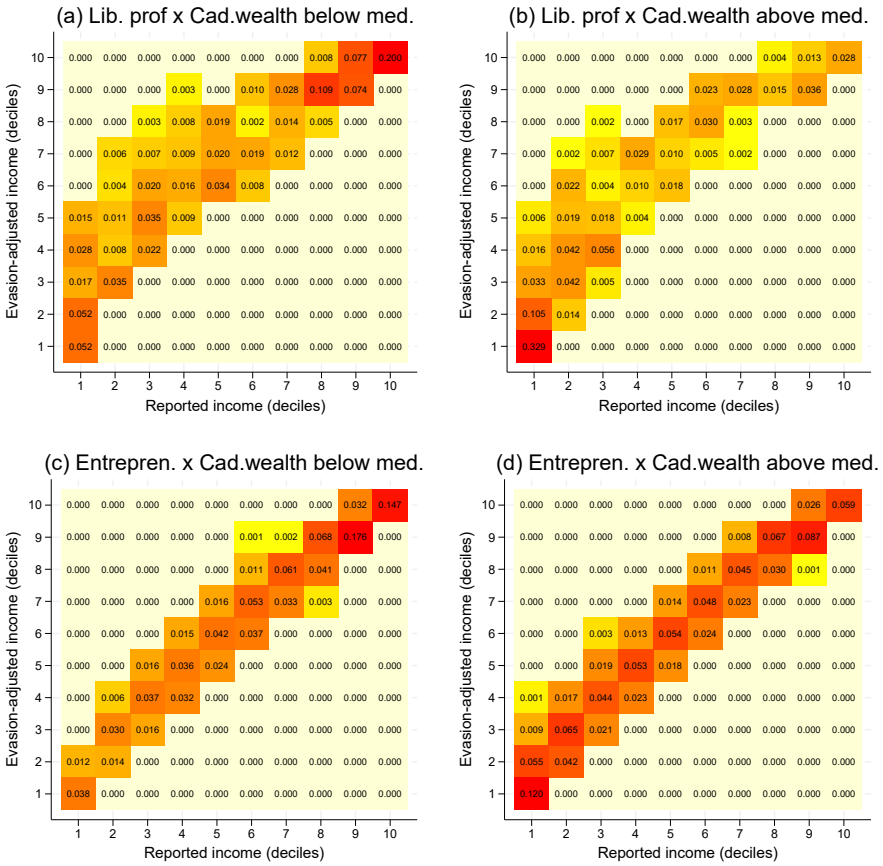


Notes: The graphs report the average share (%) of tax and SSCs revenues lost by decile groups, based on equivalized disposable income in the evasion-adjusted scenario. Income computed taking into account evasion rates obtained from OLS estimates using 7-year-average income as in the last column of Table 4; 95% confidence intervals used. Source: authors' elaboration based on EUROMOD tax simulator.

(b)), we observe positive upward mobility towards higher income decile groups once income is adjusted to account for evasion. Regarding liberal professionals, we observe positive upward transition probabilities for all decile groups after adjusting for income to account for tax evasion. Panels (c) show that tax evasion also impacts groups where no evasion is detected, as once the tax evasion correction is taken into account, their relative position worsens, causing some of them to move below the diagonal. These findings highlight the importance of re-ranking effects among the self-employed when tax evasion is considered. In the distributional analysis presented in this section, if we had used the reported income to rank individuals, we would have observed significant variations in disposable income, particularly at the bottom of the distribution, because the evasion-adjustment produces a substantial re-ranking, causing tax evaders to move

up the income ladder. Tax evaders who are found in the lower part of the reported income distribution are not necessarily poor, but may be wrongly considered so due to concealed income (Johns and Slemrod 2010).

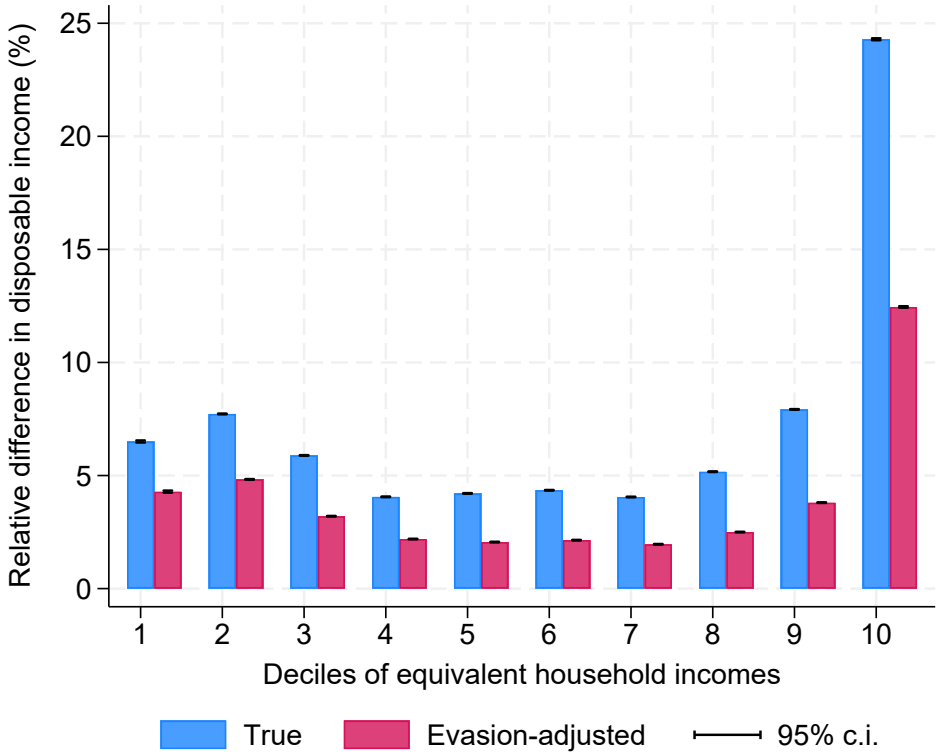
Figure 9 Transition matrices from reported to evasion-adjusted income



Notes: The graphs report the joint distribution of reported income (horizontal axis) and evasion-adjusted income (vertical axis) for professionals (panels a-b) and entrepreneurs, below and above the median cadastral wealth respectively (panels c-d), grouped by household income deciles. We use equivalized household disposable income and evasion rates obtained from OLS estimates from Table 4 (last column). Light-yellow-coloured areas denote zero-frequency cells; other coloured areas denote positive (from the minimum in yellow to the maximum in red), increasing frequencies. Numbers in cells estimate the relative frequency of each cell.
Source: authors' elaboration based on EUROMOD tax simulator.

In Figure 10, we show the distribution of changes in disposable income calculated under different scenarios by income decile groups. We find that the positive variations in disposable income under the true income scenario, where we observe evaded income that is not taxed, lie between +5% and +10% for all taxpayers and +25% for taxpayers in the top decile group. The positive changes in disposable income due to evasion are reduced in the evasion-adjusted scenario, where taxpayers pay taxes and contributions

Figure 10 Distribution of changes in disposable income in different tax compliance scenarios, year 2018



Notes: The graph reports the average individual percentage variation in disposable income in the 'true income scenario' (blue bars) and 'evasion-adjusted scenario' (red bars) with respect to the 'reported income scenario' by income decile groups, based on equivalized disposable income in the 'evasion-adjusted scenario'. Income computed taking into account evasion rates obtained from OLS estimates using 7-year-average income as in the last column of Table 4; 95% confidence intervals used.
Source: authors' elaboration based on EUROMOD tax simulator.

based on their overall income. Nevertheless, they are still relevant for all taxpayers and biased in favour of the high-income taxpayers.

The consequences of tax evasion in terms of income inequality are not negligible (Table 6). Italy is characterised by a relatively high level of disposable income inequality and a limited redistributive impact of the tax-benefit system. Estimates based on reported administrative data indicate a Gini index of approximately 0.56 for market income and 0.35 for disposable income, with a redistributive effect of the tax-benefit system (i.e., the difference between the Gini indices, known as the Reynolds-Smolenski index) of 0.21. However, considering the true income scenario, the inequality of market and disposable income would increase to around 0.57 and 0.37, respectively, and

Table 6 Gini inequality indices under different scenarios, year 2018

	Reported	Current income		Average income	
		True	Evasion-adjusted	True	Evasion-adjusted
Market income	0.556	0.583	0.583	0.574	0.574
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Taxable income	0.425	0.425	0.456	0.425	0.445
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)
Disposable income	0.347	0.390	0.367	0.375	0.359
	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)

Notes: The indices are calculated on income distributions equivalised using the OECD modified equivalence scale; Income computed taking into account evasion rates obtained from OLS estimates from the last two columns of Table 4; standard errors are in parenthesis.
Source: authors' elaboration based on EUROMOD tax simulator.

the redistribution would diminish to 0.20. In a scenario of full tax compliance (i.e. evasion-adjusted income), the Gini index of disposable income would be midway between the reported and the true scenario, suggesting that inequality measures based on reported income might be underestimated, as also discussed by (Hurst et al. 2014; Blanchet et al. 2022).

5 Concluding remarks

The increasing availability of microdata from administrative records provides researchers with new, detailed information that can be used to support policy-making more effectively (Pomeranz and Vila-Belda 2019). Thanks to the collaboration with the Italian tax administration, this work combined data on household consumption of a representative sample of the Italian population with data on their income and property values from administrative register records for the first time in Italy. The main findings of our study, which are robust to different sensitivity checks, suggest that self-employed workers in Italy underreport about 29-32% of their income relative to similar wage and salary workers, who are unable to evade income tax as their income is third-party reported. This aligns with the official estimates for Italy obtained by using National Accounts. It is also remarkably similar to what is found in several developed countries, as estimated by food-expenditure equations. Additionally, we find extensive heterogeneity of tax evasion among the self-employed, with tax evasion being more concentrated among liberal professionals, who underreport about twice the income of small entrepreneurs. These results point to an explanation of tax evasion that goes beyond third-party reporting of taxable income or individual characteristics and social norms and is consistent with traditional theoretical literature suggesting that it is rational to evade taxes when allowed to and suggests that income tax evasion is related to the structure of the Italian economy characterised by a large share of self-employment and of small enterprises. Moreover, these results support the notion that unproductive entrepreneurship also undermines productive entrepreneurship

through tax evasion (Baumol 1990). These findings provide potentially useful information for tax administrations to design effective tax auditing strategies (Battaglini et al. 2020). Policies to reduce tax evasion in Italy should also consider incentives for micro-firms to grow in size and complexity. Treating self-employment as a single analytical category can produce biased results and misleading policy prescriptions (Hurst and Pugsley 2011).

We have also documented the consequences of self-employed tax evasion on the public budget, income inequality and redistribution. Our distributional analysis reveals that households at the top of the income distribution reap the majority of the benefits from reduced tax and social security contributions. The effects of self-employment income tax evasion on public budgets are relevant, with a loss of about 10% of total actual revenues. Once evaded income is added to reported disposable income, inequality levels, as measured by the Gini index of equivalent disposable income, increase by 1-2 percentage points. This confirms the conclusions reached by Hurst et al. (2014) that neglecting self-employment underreporting introduces a noticeable bias into standard statistics and compromises the intended functioning of the tax-benefit system.

References

Advani, A. 2022. Who does and doesn't pay taxes? *Fiscal Studies 43*(1): 5–22 .

Albarea, A., M. Bernasconi, A. Marenzi, and D. Rizzi. 2020. Income Underreporting and Tax Evasion in Italy: Estimates and Distributional Effects. *Review of Income and Wealth 66*(4): 904–930. https://doi.org/10.1111/roiw.12444 .

Allingham, M.G. and A. Sandmo. 1972. Income tax evasion: a theoretical analysis. *Journal of Public Economics 1*(3-4): 323–338 .

Alm, J. 2012. Measuring, explaining, and controlling tax evasion: Lessons from theory, experiments, and field studies. *International Tax and Public Finance 19*(1): 54–77 .

Alm, J., Y. Liu, and K. Zhang. 2019. Financial constraints and firm tax evasion. *International Tax and Public Finance* 26: 71–102 .

Almunia, M. and D. Lopez-Rodriguez. 2018. Under the radar: The effects of monitoring firms on tax compliance. *American Economic Journal: Economic Policy 10*(1): 1–38 .

Alstadsæter, A., N. Johannesen, S.L.G. Herry, and G. Zucman. 2022. Tax evasion and tax avoidance. *Journal of Public Economics* 206: 104587 .

Alstadsæter, A., N. Johannesen, and G. Zucman. 2019. Tax evasion and inequality. *American Economic Review 109*(6): 2073–2103 .

Andreoni, J., B. Erard, and J. Feinstein. 1998. Tax Compliance. *Journal of Economic Literature 36*(2): 818–860 .

ATO. 2022. Commissioner of Taxation annual report 2020–21 – Australia Taxation Office, Australia. https://tinyurl.com/bddcxeau.

Auten, G. and D. Splinter 2019. Top 1 percent income shares: Comparing estimates using tax data. In *AEA Papers and Proceedings*, Volume 109, pp. 307–311. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Barsky, R.B., F.T. Juster, M.S. Kimball, and M.D. Shapiro. 1997. Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *The Quarterly Journal of Economics 112*(2): 537–579 .

Barth, E. and T. Ognedal. 2018. Tax evasion in firms. *Labour 32*(1): 23–44 .

Basri, M.C., M. Felix, R. Hanna, and B.A. Olken. 2021. Tax administration versus tax rates: Evidence from corporate taxation in indonesia. *American Economic Review 111*(12): 3827–3871 .

Battaglini, M., L. Guiso, C. Lacava, and E. Patacchini 2020. Tax professionals: Tax-evasion facilitators or information hubs? NBER Working Paper 25745, NBER.

Baumol, W.J. 1990. Entrepreneurship: Productive, unproductive, and destructive. *Journal of Political Economy 98*(5): 893–921 .

Benedetti, I., F. Crescenzi, T. Laureti, and L. Secondi. 2024. Adopting the fuzzy approach to analyze food poverty in Italy: A study on vulnerable households using household budget survey data. *Socio-Economic Planning Sciences* 95: 102006 .

Bergolo, M., G. Burdin, M. De Rosa, M. Giaccobasso, M. Leites, and H. Rueda. 2022. How do Top Earners Respond to Taxation? Evidence from a Tax Reform in Uruguay. https://ssrn.com/abstract=4007698.

Bergolo, M., R. Ceni, G. Cruces, M. Giaccobasso, and R. Perez-Truglia. 2023. Tax audits as scarecrows: Evidence from a large-scale field experiment. *American Economic Journal: Economic Policy 15*(1): 110–153 .

Blanchet, T., L. Chancel, and A. Gethin. 2022. Why is Europe more equal than the United States? *American Economic Journal: Applied Economics 14*(4): 480–518 .

Blanchet, T., I. Flores, and M. Morgan. 2022. The weight of the rich: improving surveys using tax data. *The Journal of Economic Inequality 20*(1): 119–150 .

Boeri, T., G. Giupponi, A.B. Krueger, and S. Machin. 2020. Solo self-employment and alternative work arrangements: A cross-country perspective on the changing composition of jobs. *Journal of Economic Perspectives 34*(1): 170–195 .

Buettner, T., B. Madzharova, and O. Zaddach. 2023. Income tax credits for consumer services: A tool for tackling vat evasion? *Journal of Public Economics* 220: 104836 .

Cabral, A.C.G., N. Gemmell, and N. Alinaghi. 2021. Are Survey-Based Self-Employment Income Under-Reporting Estimates Biased? New Evidence from Matched Register and Survey Data. *International Tax and Public Finance 5*(28) .

Cabral, A.C.G., C. Kotsogiannis, and G. Myles. 2019. Self-Employment Income Gap in Great Britain: How Much and Who? *CESifo Economic Studies 65*(1): 84–107 .

Carrillo, P., D. Pomeranz, and M. Singhal. 2017. Dodging the taxman: Firm misreporting and limits to tax enforcement. *American Economic Journal: Applied Economics 9*(2): 144–64 .

Ceriani, L., F. Figari, and C. Fiorio 2021. EUROMOD Report. Italy 2018-2021. Technical report, Institute for Social and Economic Research, University of Essex.

Cooper, D. 2013. House price fluctuations: the role of housing wealth as borrowing collateral. *Review of Economics and Statistics 95*(4): 1183–1197 .

CRA. 2023. overall federal tax gap report – Canada Revenue Agency. https://tinyurl.com/3e7psu8m.

Cullen, J.B., N. Turner, and E. Washington. 2021. Political alignment, attitudes toward government, and tax evasion. *American Economic Journal: Economic Policy 13*(3): 135–66 .

Davies, R.B., J. Martin, M. Parenti, and F. Toubal. 2018. Knocking on tax haven's door: Multinational firms and transfer pricing. *Review of Economics and Statistics 100*(1): 120–134 .

Dilnot, A. and C.N. Morris. 1981. What do we know about the black economy? *Fiscal Studies 2*(1): 58–73 .

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

Eika, L., M. Mogstad, and O.L. Vestad. 2020. What can we learn about household consumption expenditure from data on income and assets? *Journal of Public Economics* 189: 104163 .

Engström, P. and J. Hagen. 2017. Income underreporting among the self-employed: A permanent income approach. *European Economic Review 92*(C): 92–109 .

Engström, P. and B. Holmlund. 2009. Tax evasion and self-employment in a high-tax country: evidence from Sweden. *Applied Economics 41*(19): 2419–2430 .

Feinstein, J.S. 1991. An econometric analysis of income tax evasion and its detection. *The RAND Journal of Economics*: 14–35 .

Feldman, N. and J. Slemrod. 2007. Estimating tax noncompliance with evidence from unaudited tax returns. *The Economic Journal 117*(518): 327–352 .

Figari, F., A. Paulus, and H. Sutherland. 2015. Microsimulation and policy analysis, *Handbook of income distribution*, Volume 2, 2141–2221. Elsevier.

Fiorio, C.V. and F. D'Amuri. 2005. Workers' Tax Evasion in Italy. *Giornale degli economisti e Annali di economia*: 247–270 .

Gumpert, A., J.R. Hines Jr, and M. Schnitzer. 2016. Multinational firms and tax havens. *Review of Economics and Statistics 98*(4): 713–727 .

Guyton, J., P. Langetieg, D. Reck, M. Risch, and G. Zucman 2021. Tax evasion at the top of the income distribution: Theory and evidence. NBER Working Paper 28542, NBER.

Hashimzade, N., G.D. Myles, and H. Yousefi. 2021. Household tax evasion. *Journal of Public Economic Theory 23*(5): 985–1011 .

HMRC 2024. Measuring tax gaps 2024 edition - tax gap estimates for 2022 to 2023. Technical report, HMRC.

Hurst, E., G. Li, and B. Pugsley. 2014. Are Household Surveys Like Tax Forms? Evidence from Income Underreporting of the Self-Employed. *The Review of Economics and Statistics 96*(1): 19–33 .

Hurst, E. and B.W. Pugsley. 2011. What do small businesses do? *Brookings Papers on Economic Activity 42*(2 (Fall)): 73–142 .

IRS. 2022. Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2014-2016. Internal Revenue Service, US. https://www.irs.gov/pub/irs-pdf/p1415.pdf.

Johansson, E. 2005. An Estimate of Self-employment Income Underreporting in Finland. *Nordic Journal of Political Economy* 31: 99–109 .

Johns, A. and J. Slemrod. 2010. The Distribution of Income Tax Noncompliance. *National Tax Journal 63*(3): 397–418 .

Kleven, H.J., M.B. Knudsen, C.T. Kreiner, S. Pedersen, and E. Saez. 2011. Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark. *Econometrica 79*(3): 651–692 .

Kleven, H.J., C.T. Kreiner, and E. Saez. 2016. Why can modern governments tax so much? an agency model of firms as fiscal intermediaries. *Economica 83*(330): 219–246 .

Kopczuk, W. and E. Zwick. 2020. Business incomes at the top. *Journal of Economic Perspectives 34*(4): 27–51 .

Kukk, M., A. Paulus, and K. Staehr. 2020. Cheating in Europe: Underreporting of Self-employment Income in Comparative Perspective. *International Tax and Public Finance 27*(2): 363–390. https://doi.org/10.1007/s10797-019-09562-9 .

Levine, R. and Y. Rubinstein. 2017. Smart and illicit: Who becomes an entrepreneur and do they earn more? *The Quarterly Journal of Economics 132*(2): 963–1018 .

Lyssiotou, P., P. Pashardes, and T. Stengos. 2004. Estimates of the black economy based on consumer demand approaches. *Economic Journal 114*(497): 622–640 .

Martinez-Lopez, D. 2013. The underreporting of income by self-employed workers in Spain. *SERIEs* 4: 353–371. https://doi.org/10.1007/s13209-012-0093-8 .

MEF. 2024. Relazione sull'economia non osservata e sull'evasione fiscale e contributiva – anno 2024 (Report on the Shadow Economy and Tax and Contribution Evasion – Year 2024). Italian Ministry of the Economy and Finance. https://tinyurl.com/3eceye7d.

Meyer, B.D. and N. Mittag. 2019. Using linked survey and administrative data to better measure income: Implications for poverty, program effectiveness, and holes in the safety net. *American Economic Journal: Applied Economics 11*(2): 176–204 .

Nygård, O.E., J. Slemrod, and T.O. Thoresen. 2018. Distributional implications of joint tax evasion. *The Economic Journal 129*(620): 1894–1923 .

Nygård, O.E. and T.O. Thoresen. 2023. Controlling for fixed effects in studies of income underreporting. *European Journal of Law and Economics*: 1–22 .

Paulus, A. 2015. Tax evasion and measurement error: An econometric analysis of survey data linked with tax records. Technical report, ISER Working Paper Series.

Pissarides, C.A. and G. Weber. 1989. An Expenditure-Based Estimate of Britain's Black Economy. *Journal of Public Economics* 39: 17–32 .

Pomeranz, D. and J. Vila-Belda. 2019. Taking state-capacity research to the field: Insights from collaborations with tax authorities. *Annual Review of Economics* 11: 755–781 .

Rondinelli, C. 2014. On the structure of italian households: Consumption patterns during the recent crises. *Politica economica 30*(2-3): 235–260 .

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 .

Slemrod, J. 2019. Tax compliance and enforcement. *Journal of Economic Literature 57*(4): 904–54. https://doi.org/10.1257/jel.20181437 .

Slemrod, J., B. Collins, J.L. Hoopes, D. Reck, and M. Sebastiani. 2017. Does credit-card information reporting improve small-business tax compliance? *Journal of Public Economics* 149: 1–19 .

Solon, G. 1992. Intergenerational Income Mobility in the United States. *American Economic Review 82*(3): 393–408 .

Sutherland, H. and F. Figari. 2013. EUROMOD: The European Union Tax-benefit Microsimulation Model. *International Journal of Microsimulation 1*(6): 4–26 .

A Additional descriptive analysis

In this appendix, we present some descriptive statistics for the groups of liberal professional and entrepreneur households, which we used in the above empirical analysis. Liberal professionals and entrepreneurs share several similarities. The logarithm of the mean household income for both groups is quite similar, and food and utility expenditures are also nearly identical between the two groups (Table A.1). When considering other characteristics (Table A.2), entrepreneurs tend to have a slightly higher average household size, are more likely than liberal professionals to be in a couple, live in a household with children, and have higher levels of education. Liberal professionals, on the other hand, are more often in female-headed households.

Table A.1 Descriptive statistics for professional vs. entrepreneurs, main variables

	Professionals	**Entrepreneurs**
	A.Log average expenditure measures	
Food expenditures	8.774	8.755
	(0.581)	(0.566)
Utilities expenditures	7.964	7.907
	(0.673)	(0.672)
	B. Log average income measures	
B.Pre-tax Total Household Income:		
Current income	9.980	10.078
	(1.673)	(0.969)
7-year average income	10.122	10.100
	(1.387)	(0.904)
Sample size	680	1,785

Note: as in Table 1.

As shown in Figure A.1, the share of wage and salary income in total labour income varies across taxable income groups. At lower levels of taxable income, the average proportion of wage and salary income remains relatively low, indicating that households in these income groups primarily rely on self-employment income as a source of income. The highest average share of wage and salary income is found in the taxable income range of €40,000 to €60,000, indicating that households in this group tend to have a more balanced mix of self-employment and wage and salary earnings. Beyond this range, the share of wage and salary income becomes more dispersed, with notable fluctuations.

Table A.2 Descriptive statistics for professional versus entrepreneurs: other control variables

	Self-employed	
	Professionals	**Entrepreneurs**
	C. Control Variables (continuous)	
Average household size	2.692	3.006
	(1.339)	(1.276)
Sin goods: Tobacco Products	1.739	2.269
	(2.929)	(3.188)
Sin goods: Lottery Tickets	0.653	0.700
	(1.815)	(1.880)
Sin goods: Spirits	0.583	0.597
	(1.781)	(1.775)
Number of rooms at home	4.410	4.449
	(1.757)	(1.457)
	D. Control Variables (categorical)	
% household with kids	0.297	0.367
% female-headed households	0.278	0.168
% households in couple	0.622	0.729
Head age: % 36-50	0.443	0.468
Head age: % 51-65	0.357	0.350
Head age: 66 and over	0.113	0.101
Head education: % with secondary	0.830	0.500
Head education: % with college	0.146	0.318
Spouse education: % with secondary	0.476	0.411
Spouse education: % with college	0.765	0.564
% resident in the Center	0.189	0.180
% resident in the South	0.281	0.274
Dwelling: %high-end apartment	0.151	0.103
Dwelling: %low-end apartment	0.704	0.740
Dwelling: %rural units	0.034	0.031
Building: %in condominium	0.725	0.661

Note: As in Table 1.

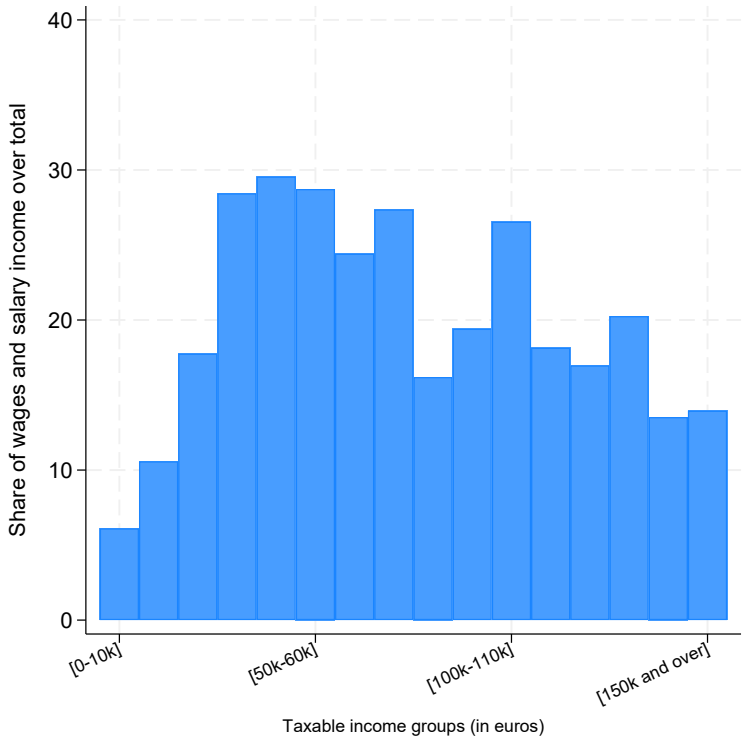


Figure A.1 Share of wage and salary income out of total labour income for self-employed households, in taxable income groups by steps of €10,000.

B Robustness analyses

In this appendix, we present some robustness tests of our main results. In Table B.1 and B.2, we test whether results are sensitive to the share of self-employment income over total income to identify self-employed households. Although evasion rates are slightly larger, the overall picture remains unaltered. In Table B.3, we test heterogeneity by household income, replacing it with heterogeneity by cadastral wealth, and obtain results that are very similar to those in Table B.2 when the 25% share of self-employment incomes is used.

In Figure B.1 and Figure B.2 and in Table B.4 we present main results of the distributive analysis using as tax evasion coefficients the estimates with heterogeneity by income levels instead of cadastral wealth, i.e. using the estimate of Table B.3, second column. Also, these results show that the analysis presented in Section 4 is robust.

Table B.1 Self-employment income tax evasion estimates using food expenditures, by self-employment working status, by different thresholds to identify self-employed households..

	Food expenditures					
	25%		**35%**		**50%**	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Current income	0.086***		0.085***		0.085***	
	(0.009)		(0.009)		(0.010)	
Average income (7-year)		0.109***		0.109***		0.110***
		(0.011)		(0.011)		(0.011)
Professionals	0.062*	0.060*	0.067*	0.065*	0.083*	0.081*
	(0.025)	(0.024)	(0.028)	(0.028)	(0.035)	(0.035)
Entrepreneurs	0.024	0.03	0.03	0.035	0.037*	0.043*
	(0.017)	(0.017)	(0.018)	(0.018)	(0.018)	(0.018)
R-squared	0.2	0.203	0.2	0.203	0.202	0.205
N. households	12166	12166	12077	12077	11924	11924
	$(1-\kappa)$		$(1-\kappa)$		$(1-\kappa)$	
Professionals	0.516***	0.424***	0.546***	0.449***	0.623***	0.521***
	(0.143)	(0.131)	(0.152)	(0.145)	(0.160)	(0.160)
Entrepreneurs	0.246*	0.238**	0.293**	0.277**	0.350***	0.327***
	(0.139)	(0.111)	(0.136)	(0.112)	(0.124)	(0.104)

Notes: Self-employed households are identified as having at least 25% of their income from self-employment (columns 1 and 2), or at least 35% (columns 3 and 4) or at least 50% (columns 5 and 6).

As in previous tables, controls include household head age and gender, in-couple dummy interacted with education (primary, secondary or tertiary) of the spouse, household size, a dummy for the presence of children, family consumption of sin goods, a full set of macro area-of-residence dummies, household head education, and housing characteristics. Standard errors are adjusted for 109 clusters in the province of family residence.

Table B.2 Self-employment income tax evasion estimates using food expenditures, by self-employment working status and cadastral wealth heterogeneity, by different thresholds to identify self-employed households.

Share of self-employment income	**Food expenditures**					
	25%		**35%**		**50%**	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Current income	0.085***		0.084***		0.085***	
	(0.009)		(0.009)		(0.010)	
Average income (7-year)		0.109***		0.108***		0.110***
		(0.010)		(0.011)		(0.011)
Professionals × Cadastral wealth below med.	0.061	0.060	0.053	0.051	0.081	0.079
	(0.038)	(0.037)	(0.041)	(0.040)	(0.044)	(0.044)
Professionals × Cadastral wealth above med.	0.062	0.060	0.081*	0.078*	0.085	0.080
	(0.034)	(0.033)	(0.036)	(0.035)	(0.047)	(0.047)
Entrepreneurs × Cadastral wealth below med.	0.013	0.024	0.012	0.024	0.027	0.040
	(0.021)	(0.022)	(0.022)	(0.023)	(0.023)	(0.023)
Entrepreneurs × Cadastral wealth above med.	0.034	0.034	0.047*	0.046*	0.047	0.045
	(0.020)	(0.020)	(0.022)	(0.022)	(0.024)	(0.024)
R-squared	0.200	0.202	0.200	0.203	0.201	0.204
N. households	12166	12166	12077	12077	11924	11924
	$(1-\kappa)$		$(1-\kappa)$		$(1-\kappa)$	
Professionals × Cadastral wealth below med.	0.512**	0.423**	0.465*	0.380	0.611***	0.515***
	(0.220)	(0.199)	(0.263)	(0.234)	(0.205)	(0.199)
Professionals × Cadastral wealth above med.	0.518***	0.420**	0.620***	0.514***	0.630***	0.519**
	(0.194)	(0.176)	(0.162)	(0.159)	(0.212)	(0.208)
Entrepreneurs × Cadastral wealth below med.	0.140	0.199	0.131	0.196	0.268	0.308**
	(0.207)	(0.152)	(0.224)	(0.164)	(0.186)	(0.139)
Entrepreneurs × Cadastral wealth above med.	0.331**	0.268**	0.425***	0.347***	0.422***	0.339**
	(0.146)	(0.127)	(0.135)	(0.122)	(0.151)	(0.135)

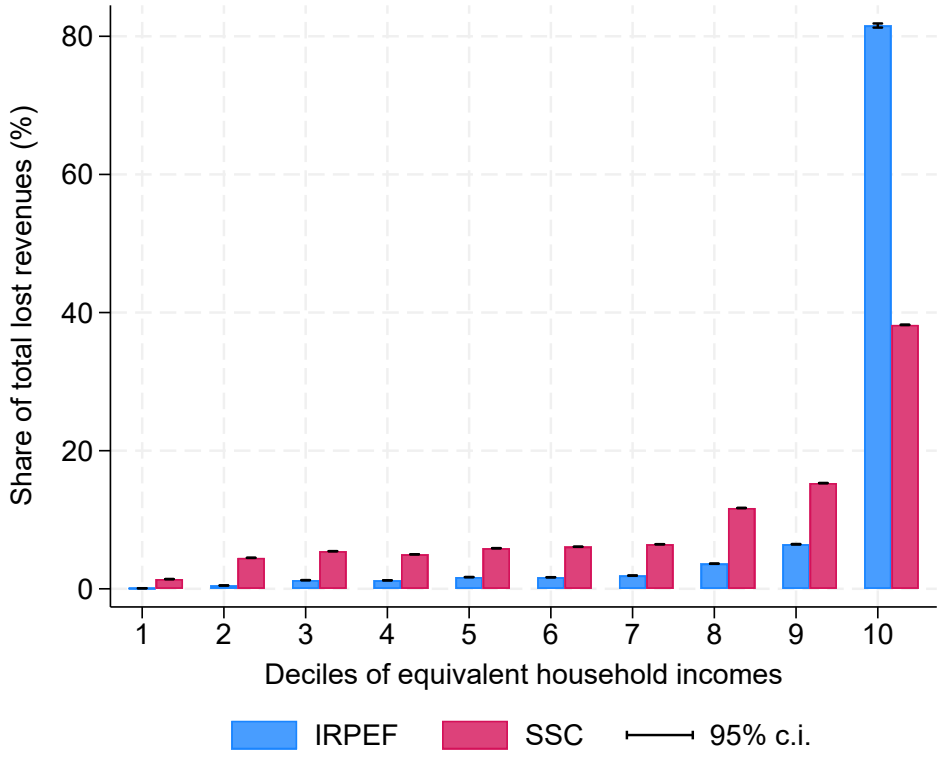
Notes: As in Table 4. Self-employed households are identified as having at least 25% of their income from self-employment (columns 1 and 2), or at least 35% (columns 3 and 4) or at least 50% (columns 5 and 6).

Table B.3 Self-employment income tax evasion estimates using food expenditures, by self-employment working status and income heterogeneity, by different thresholds to identify self-employed households.

Share of self-employment income	**Food expenditures**					
	25%		**35%**		**50%**	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Current income	0.085***		0.098***		0.087***	
	(0.009)		(0.009)		(0.010)	
Average income (7-year)		0.109***		0.128***		0.112***
		(0.010)		(0.010)		(0.011)
Professionals × Income below med.	0.056	0.055	0.092**	0.092**	0.100*	0.098*
	(0.039)	(0.038)	(0.043)	(0.043)	(0.052)	(0.053)
Professionals × Income above med.	0.066*	0.064	0.051	0.047	0.072	0.069
	(0.038)	(0.039)	(0.038)	(0.039)	(0.051)	(0.051)
Entrepreneurs × Income below med.	0.018	0.027	0.026	0.039	0.050*	0.062**
	(0.021)	(0.021)	(0.024)	(0.023)	(0.026)	(0.026)
Entrepreneurs × Income above med.	0.029	0.031	0.013	0.015	0.027	0.030
	(0.024)	(0.024)	(0.024)	(0.025)	(0.025)	(0.025)
R-squared	0.199	0.202	0.200	0.203	0.201	0.204
N. households	12166	12166	12077	12077	11924	11924
	$(1-\kappa)$		$(1-\kappa)$		$(1-\kappa)$	
Professionals × Income below med.	0.485**	0.398*	0.609***	0.514***	0.681***	0.581***
	(0.233)	(0.210)	(0.172)	(0.159)	(0.186)	(0.193)
Professionals × Income above med.	0.541**	0.443**	0.404*	0.308	0.559**	0.457*
	(0.213)	(0.202)	(0.232)	(0.215)	(0.267)	(0.258)
Entrepreneurs × Income below med.	0.192	0.223	0.231	0.261*	0.433***	0.424***
	(0.195)	(0.147)	(0.188)	(0.134)	(0.158)	(0.127)
Entrepreneurs × Income above med.	0.290	0.247	0.124	0.111	0.270	0.235
	(0.192)	(0.160)	(0.217)	(0.170)	(0.198)	(0.164)

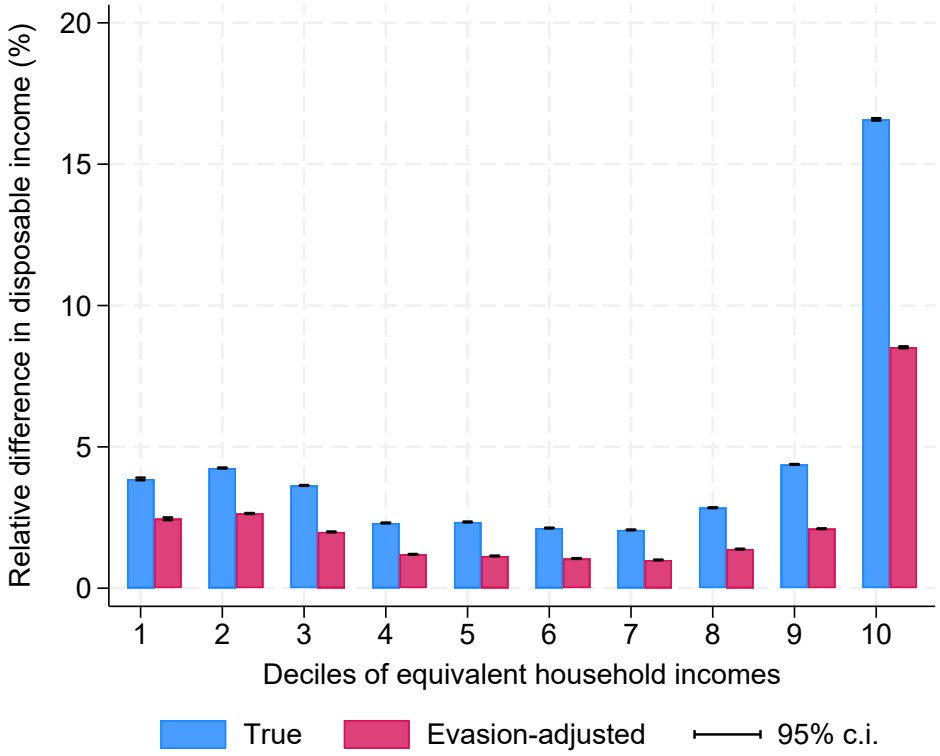
Notes: As in Table 4. Self-employed households are identified as having at least 25% of their income from self-employment (columns 1 and 2), or at least 35% (columns 3 and 4) or at least 50% (columns 5 and 6).

Figure B.1 Distribution of the share of tax and SSCs revenues lost comparing actual and evasion-adjusted scenario, year 2018



Notes: The graphs report the average share (%) of total tax and SSCs revenue losses by decile groups, as calculated in the evasion-adjusted scenario over equivalized household disposable income, with evasion rates obtained from OLS estimates using 7-year-average income as in column (2) of Table B.3; 95% confidence intervals used.
Source: authors' elaboration based on EUROMOD tax simulator.

Figure B.2 Distribution of percentage of true and evasion-corrected disposable income wrt observed



Notes: The graphs report the average share (%) of total tax and SSCs revenue losses by decile groups, as calculated in the evasion-adjusted scenario over equivalized household disposable income, with evasion rates obtained from OLS estimates using 7-year-average income as in column (2) of Table B.3; 95% confidence intervals used. Source: authors' elaboration based on EUROMOD tax simulator.

Table B.4 Gini inequality indices under different scenarios, year 2018

	Reported	Current income		Average income	
		True	Evasion-adjusted	True	Evasion-adjusted
Market income	0.556	0.577	0.577	0.570	0.570
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Taxable income	0.425	0.425	0.448	0.425	0.440
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)
Disposable income	0.347	0.381	0.363	0.368	0.357
	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)

Notes: The indices are calculated on income distributions equivalised using the OECD modified equivalence scale; Income computed taking into account evasion rates obtained from OLS estimates from the last two columns of Table B.3; standard errors are in parenthesis.
Source: authors' elaboration based on EUROMOD tax simulator.