Giornale degli Economisti e Annali di Economia
Volume 64 - N. 2/3 (Novembre 2005) pp. 247-270

WORKERS' TAX EVASION IN ITALY*

CARLO V. FIORIO**
University of Milan and Econpubblica, Bocconi University

FRANCESCO D'AMURI***
University of Pavia and Econpubblica, Bocconi University

Received: November 2004; accepted: May 2005

We apply a direct method to estimate tax evasion in Italy, assuming that tax evaders might consider declaring a closer-to-true income in an anonymous interview. The methodology is applied to employed and self-employed taxpayers, combining the Survey of Household Income and Wealth (SHIW) by the Bank of Italy and a large random sample of tax forms by SeCIT (Tax auditing office – Ministry of Finance), both referred to incomes received in 2000. Paying particular attention to the post-stratification of the data, we find that tax evasion is consistently higher for self-employment income than for employment income: the difference ranges from about 7% in lower deciles to 27% around the mode. This analysis shows that a relevant level of tax evasion arises also at low levels of employment income, although some under-sampling and misreporting problems need to be considered. An evaluation of the redistribution and incidence effects of tax evasion among workers is provided.

J.E.L. Classification: H26, D63, H24.
Keywords: Tax evasion, post-stratification, microsimulation, redistribution.

---

*Participants to the 2004 SIEP Conference and to the DEFAP Conference "La questione tributaria in Italia", 17-18 December 2004, and in particular Massimo Bordignon and Anna Marenzi, provided very helpful comments. An anonymous referee's suggestions are also gratefully acknowledged. Special thanks go to Alberto Zanardi for important suggestions, to SeCIT for the invaluable collaboration with the tax forms data and to Roberto Artoni for his support.

**Address: Dipartimento di Scienze Economiche, Aziendali e Statistiche, Università di Milano, Via Conservatorio, 7, 20122 Milano. email: carlo.fiorio@unimi.it.

***Address: Dipartimento di Economia Pubblica e Territoriale, Università di Pavia, Strada Nuova, 65, 27100 Pavia. email: francesco.damuri@unibocconi.it.

1. Introduction

According to recent estimates between 27% and 48% of official Italian GDP is hidden. Among major OECD countries Italy presents the highest levels of tax evasion (Schneider, 2000a). These figures raise concerns at a macro level, but they say little about who tax evaders are. Moreover, they do not provide much direction for policy. Although measuring tax evasion is a formidable task – Schneider (2000b, p. 1) describes tax evasion measurement "as a scientific passion for knowing the unknown" – we believe that a better understanding of tax evasion is necessary. Tax evasion alters the competitiveness of the market, introduces iniquities among equals and modifies the outcomes of public policies.

This paper attempts to provide an estimate of tax evasion by combining two rich microeconomic data sets that are representative of the Italian population. Such a route has already been walked by other authors and Section 2 summarizes the main results so far about the estimation of tax evasion in Italy, focussing mainly on microeconomic approaches. Here the main assumption is that taxpayers who decide to hide part of their income from tax authorities might consider declaring a more correct figure to an anonymous interviewer. Using this assumption, we estimate tax evasion in Italy in the year 2000 by comparing income data from the Survey of Household Income and Wealth (SHIW) produced by the Bank of Italy with the tax forms random sample developed by SeCIT. This methodology requires great care to verify the consistency of the two data sets. Section 3 deals with this problem. This section also makes clear why we focussed on tax evasion by active people only. Estimation results are then provided and discussed in Section 4. As tax evasion behavior changes significantly at different income levels, and results are particularly surprising for employees, an attempt to describe employment income distribution conditional on a set of individual and industry variables is undertaken using ordinary least squares and quantile regression tools. In Section 5 the effect of the estimated tax evasion on work income is analyzed using a tax-benefit microsimulation model and standard indices of redistribution, progressivity and tax incidence. Finally, Section 6 concludes and discusses some policy implications of the results obtained.

2. Available evidence about tax evasion in Italy

Tax evasion can be estimated using a direct or an indirect approach. Indirect methods estimate tax evasion considering it equal to the difference between aggregated macro indicators (e.g. the discrepancy between income and expenditure or the difference between the actual demand for money and the demand for money estimated in absence of taxes). Direct methods aim at estimating tax evasion through the use of sample survey micro-data based on voluntary participation or the results of the auditing activity of tax authorities. In contrast to indirect methods, direct methods are more suitable to analyze tax evasion at the micro level and they can highlight directions for policy.

Some of these methods have been applied to provide a measure of tax evasion in Italy. Among those who used indirect methods, Schneider (2000a) used the currency demand approach and Zizza (2002) also the factorial analysis. Zizza estimates the share of the underground economy (excluding illegal and criminal activities) on GDP for the years 1984-2000 between a maximum of 17.6% (1991) and a minimum of 14.3% (2000). According to Schneider (2000a), where estimates include also illegal and criminal activities, the share of the underground economy on the Italian GDP is very high and increasing (from 25.8% in 1994 to 27.8% in 1998), the highest rate among the OECD countries.

Cannari *et al.* (1995) Marenzi (1996), Bernasconi and Marenzi (1997), Bordignon and Zanardi (1997) and Calzaroni (2000) used direct methods. Calzaroni (2000) estimates labor supply and labor demand functions by sectors, using household and firm surveys respectively, and compares results at the national and the regional level. The difference between the two is considered to be the number of the irregular workers. This figure, multiplied for the average sectorial productivity estimated for regular workers, gives a first measure of the underground economy. The overall incidence of the underground economy is calculated by complementing this figure with coefficients correcting for the underestimation of the turnover and the balancing between aggregated input and output. The results indicate that, for 1998, the share of the underground economy on GDP in Italy is between 14.7 and 15.4%. Cannari *et al.* (1995), Marenzi (1996) and Bernasconi and Marenzi (1997) use a different approach: they assume that individuals report a more truthful income to an anonymous interviewer than to the tax authorities. Hence, when income data recorded in the Survey of Household Income and Wealth (SHIW) – produced by the Bank of Italy (BI) – is larger than that recorded in the analysis of tax forms – produced by the Ministry of Finance (MF) – the difference between SHIW and MF disposable income is considered as hidden income. Cannari *et al.* (1995) and Marenzi (1996) considered years 1989 and 1991 respectively. The estimate of tax evasion is performed for different groups of tax payers, identified by their main income (employment, self-employment, pension, etc.). Bernasconi and Marenzi (1997) find evidence of positive tax evasion in the two first deciles of employment income (27.2% in the first decile and 10.2% in the second) and a slight underreporting in the following ones. According to the same estimates, entrepreneurs show a much higher, but decreasing with the level of income, level of tax evasion. In particular, in the first four deciles of income tax evasion is higher than 50% (81% in the first decile), and remains above 40% across all income deciles. Also professionals show levels of tax evasion decreasing with the level of income, with an estimated evasion of 87.1% in the first decile, but lower than 20% in the last 5 deciles. In Cannari *et al.* (1995) tax evasion is on average zero for total employment income (for this kind of taxpayers an estimate by decile is not provided) and decreasing in the level of income for members of the art or professions and for entrepreneurs. Bernasconi and Marenzi (1997) provide also an estimate of redistributive effects using a tax-benefit microsimulation model. They show that had evasion not incurred, the vertical effect of taxation would have increased and the horizontal and re-ranking effect would have decreased by large proportions.

This methodology relies heavily on the hypothesis that SHIW data set is representative of the population and of its subgroups. As MF data refers to the population, a measure of tax evasion based on this methodology requires the SHIW data to be a good approximation of the population. This requirement must be verified carefully.

An estimate of tax evasion stemming from the comparison between SHIW and MF data also requires that income variables are defined consistently. Cannari *et al.* (1995), Marenzi (1996) and Bernasconi and Marenzi (1997) use total net income by group of taxpayers. However, SHIW data set is quite reliable for the measurement of labor income, but it is much less so for other types of income such as capital, estate and building income (Cannari and D'Alessio, 1992; Brandolini, 1999). This is due to two main reasons: first, these data are collected at the household level and they can only be imputed to the individual taxpayer; second, there is a tendency to misestimate the true value of these incomes, which is probably not voluntary and is however common also to other similar surveys.

Other direct methods to estimate tax evasion, although possible in other countries, have seldom been used in Italy. Bordignon and Zanardi (1997) look for determinants of personal income tax evasion in Italy by combining an institutional analysis with a quantitative one based on the results of the audits carried out by the tax administration on the self-employed in years 1987 and 1989. Although they attempt to correct for sample selection bias[1], their quantitative analysis is limited by the small number of variables that can be used to identify determinants of tax evasion.

[1] Audits are conducted non-randomly by the tax administration but selection methods are unknown to the authors.

Data limitation is also the main reason why a direct approach like that suggested by Pissarides and Weber (1989) has never been published for Italy. They use the British FES data and assume that income and expenditure are reported accurately by employees whose employer filed their income report. They estimate the expenditure function for these households as the true relation between income and expenditure given a set of individual and household characteristics. Assuming that expenditure is reported correctly also by other households, an estimate of their income is then estimated, and consequently of their tax evasion. However, the SHIW data are not suited to their approach, mainly because expenditure variables suffer large measurement error and no better data have been produced so far.

3. Applying a direct methodology

3.1 Consistency checks

In this paper a direct approach to tax evasion estimation is used, as previously used for Italy by Cannari *et al.* (1995) and Marenzi (1996). The basic assumption is that an income receiver who decides to evade tax payment will under-report her taxable income to tax authorities but declare the true income, or at least a closer approximation to the true income, to an interviewer who grants anonymity. As survey-based data tend to grant anonymity to increase the probability of participation in the survey and of truthful declarations, the comparison of income distribution using tax records and survey-based data sets should allow one to build up a picture of tax evasion behavior. However, before using a direct methodology to estimate tax evasion, some preliminary checks on the data available and their consistency are needed.

The data sets we will use are the 2000 Survey of Household Income and Wealth (henceforth, SHIW00), which reports data on incomes received during the year 2000, and the stratified random sample of individual tax forms referred to income received in year 2000 (henceforth, MF00).

SHIW00, published by the Bank of Italy, collects detailed microdata about 8,001 households and 22,268 individuals on disposable income, consumption, labor market, monetary and financial variables. The survey is collected by stratified random sampling from the population of households and is meant to be representative of the national household population. It also presents detailed information for each member of the household. The Bank of Italy, after some consistency checks, computes a set of sampling weights to gross-up estimates to population totals (for more details on SHIW00, see Banca d'Italia, 2002).

Although MF00 is not available to the authors, all requested elaborations were kindly performed at SeCIT (Tax auditing office – Ministry of Finance) on a very large data set (over 250,000 taxpayers). This data set is a non-proportional random sample of the population of tax forms[2] and it is representative of the population of tax payers for all levels of incomes (for more details on MF00, see Di Nicola and Monteduro, 2004).

The main differences between the two data sets are the following: (a) SHIW00 is a survey based on voluntary participation, MF00 is a random sample of tax forms that all Italian income receivers are obliged to submit; (b) the reference sampling unit is the household in SHIW00, the individual in MF00; (c) in SHIW00 there is information on income net of taxes and social contributions (which can only be simulated by using a tax-benefit microsimulation model), in MF00 gross and net income, as well as tax credit, tax deductions and compulsory social contribution paid are fully known; (d) under-sampling of high income households might be relevant in SHIW00 as the non-response issue increases with income (see references below), in MF00 this is not so, as participation is non voluntary; (e) a large range of information of individual and household characteristics is available in SHIW00, while mainly fiscal variables are available in MF00 and there is no way to recover household or family composition from these data, even using the information on fiscally-dependent relatives; (f) SHIW00 and MF00 are both micro-data sets, and the aggregation bias that would arise in comparing household micro-data sets with aggregated fiscal tables[3] can be avoided; (g) SHIW00 collects recall information (i.e. the interviewers are asked to recall the incomes they received in the previous year, without providing evidence) and can be affected by nonvoluntary approximation error and misreporting, MF00 presents no approximation error, as the report is based on official documentation and, at most, it should include only taxpayer mistakes in filling in the tax form and tax form data elaboration mistakes.

The reasons that make SHIW00 and MF00 particularly suited for a direct methodology of tax evasion estimation are that: (a) both present data on income received in the year 2000 and all bias due to approximate updating procedure can be avoided; (b) they are representative of the national population and analysis of tax compliance could be performed by regional areas, by age and the type of income received.

Although SHIW00 and MF00 are two good data sets for a direct approach to tax evasion, it is important that some *a priori* checks are performed. These include: the consistent definition of income variables to be

[2] Included tax forms are "Modello Unico", "Modello 730", "Modello 770".

[3] See, for instance, Fiorio and D'Amuri (2005a) the previous version of this paper.

compared; an analysis of the main hypothesis of the direct approach estimation, and first of all that there is a positive difference between the data reported to the anonymous interviewer and those reported to the tax authorities; and the analysis of how representative the data sets are of the total population, the post-stratification issue.

3.2 Variable definition and the main hypothesis for direct estimation

In this paper we only focus on the personal income tax (IRPEF) on income received by individual tax payers and by unincorporated firms (*società di persone*) in the year 2000. The aggregate amount of the IRPEF tax base in 2000 was divided as follows: 5% in income from farming activities and building and real estate property, 49% in employment income, 27% in pension income, 18% in self-employment, shareholding and unincorporated firm income[4], 2% in capital and other incomes (Ministero delle Finanze, 2004). We ruled out the possibility of tax evasion in pension income as an unlikely event and discrepancy of data could only be ascribed to data misreporting (which could be relevant as age increases). Only employment and self-employment income was considered, as it accounts for 92% of the total IRPEF tax base, excluding pension income. Moreover, the recording of other types of income in SHIW00 is highly unreliable (Cannari and D'Alessio, 1990, 1992; Brandolini and Cannari, 1994).

The definition of employment income in SHIW00 is relatively straightforward and definition is consistent with MF00. Self-employment income presents some challenges due to the difficulties of disentangling partnership income from SHIW00 consistent with the definition of partnership income in MF00. Hence, self-employment income is defined as the sum of members of the professions, sole proprietors, free lancers, family business and active shareholders/partners (of unincorporated companies – i.e. S.n.c. and S.a.s. – only).

Total employment income is on average the same in the two data sets, while self- employment income is over 20% larger in the SHIW00 than in MF00. Although this is the first positive evidence for using direct methods for estimating tax evasion, it is clearly not sufficient. Data declared to an anonymous interviewer might be affected by a number of flaws. For instance, SHIW00 data might suffer systematic under or over reporting. This could be due to the reasons mentioned in the Section 3.1, but also to a plausible psychological attitude that would induce low-income (high- income) in-

[4] This includes the sum of Sections RE, RF, RG and RH of the "Unico 2001/Redditi 2000" form.

dividuals to over-declare (under-declare) their income in the interview in order to feel closer to the social norm, that could be approximated by the median income. This would compress the declared income distribution introducing a bias into the results. Notwithstanding these data limitations, and possibly others that we do not see, we believe this is still a valuable exercise.

3.3 The post-stratification issue

It is well known that the non-response bias is larger, the larger the rate of non-response and the larger the differences between respondents and non-respondents (see among others Little and Rubin, 1987). Unfortunately the non-response rate in the SHIW00 is over 50% (Banca d'Italia, 2002) and some studies show that the non-response decision is not random. Cannari and D'Alessio (1992) analyzed the non-response bias in SHIW using the second wave of the panel sub-sample (the first wave was interviewed in 1987, the second in 1989). Knowing the characteristics of those who refused to respond again in 1989, they expanded the results to the whole sample. They found that non-response is more frequent in households who reside in urban areas and in the North. The participation rates decreases as income rises and household size decreases, while the relationship with the age of the head of the household is ambiguous. D'Alessio and Faiella (2002) also showed that non-response behavior is dependent on net financial wealth. They found this result using a supplementary sample of about 2,000 households, clients of a leading commercial bank. Although the sample can hardly be considered representative of the whole population and sub-samples size are not very large, individuals with financial wealth of more than Lit 1 billion (about € 0.5 million) have about half the response rates of other groups. D'Alessio and Faiella (2002) also used alternative methodologies, reaching the conclusion that non-response behavior is not random, and is more frequent among wealthier households.

For a direct estimation of tax evasion it is important that the characteristics of SHIW00 are as close as possible to those of the population. As MF00 is not affected by non-response bias as it is not a survey but a random sample from the population of tax forms, we used a post-stratification procedure for SHIW00 only. The post-stratification procedure computes a new weight that adjusts for differences between the sample data and the characteristics of the population to be modelled, reflecting differential non-responses between different groups in the sample. It involves stratifying the sample, after the data have been collected, by some relevant characteristics, and applying known proportions.

In more detail, the post-stratification procedure we used consists of assigning to each unit in a sample of dimension N a weight p_j, with $j = 1, \ldots, N$, such that some chosen statistics of interest calculated on the weighted sample coincide with the population statistics. The procedure is trivial if we need to reconcile the sample with the population using only one discrete statistic, s_k with $k = 1, \dots K$, such as family types or income ranges. In this case, we compute the probability of having the characteristic s_k in the sample, say $P(s_k)$, and make it equal to the probability of having the same characteristic in the population, say $p(s_k)$. If the dimension of the sample and of the population are N and n respectively, then the post-stratification weight is $p_j = np(s_k)/NP(s_k)$, i.e. the size of the cell with characteristic s_k in the population divided by the size of the cell with characteristic s_k in the sample. If more variables are included for the post-stratification procedure, then the interactions between the different variables (i.e. their joint distribution) should be considered. However, this conflicts with available information from external sources, which in general do not report the joint distribution of population variables but only the totals for each variable. For instance, it is generally possible to know the total number of single-parent families and the total number of self-employed in the population, but not how many single-parent families have self-employment income. Hence, the conditions imposed on the weights p_j are far less stringent than in the "full information" case we would have if the joint distribution were known, and in general there are many possible sets of weights p_j achieving the desired adjustment. To choose between them it is usually required that given a data set of dimension N, with original sampling weights p_j, $j = 1, 2, \dots N$, the set of post-stratification weights p_j have the least deviation from original weights, q_j (see, for instance Hollenbeck, 1976; Merz, 1983, 1985; Atkinson *et al.*, 1988). The main reason for this is that original weights present information – e.g. about the sampling procedure – that is valuable[5] . Both post-stratification and initial weights have to sum up to the population size: $\Sigma q_j = \Sigma p_j = n$. If original and post-stratification weights sum up to the sample dimension, they first have to be multiplied by $n=N$. It is then common practice to impose the condition that the new weights minimize the distance from initial weights. Merz (1983) suggests minimizing a measure of distance derived from information theory (see for instance Cowell, 1980):

$$d(p,q) = \sum p_j log\left(\frac{p_j}{q_j}\right) \tag{1}$$

It improves on other distance measures – for instance the quadratic dis-

[5] The choice of initial weights is however not crucial, as shown by replacing initial with uniform weights by Atkinson *et al.* (1988, pp. 15-16).

tance used in a number of studies including Hollenbeck (1976) – as it rules out the possibility that weights can be negative[6].

As for the optimal number of control totals to be included, no result is currently available. Although it is more common to face the problem of not having enough external sources than of having too many, Sutherland (1989, p. 15) warns against the risk of increasing the variance of weights, since the larger the number of control totals becomes, the smaller the number of observations in each "cell"(i.e. with each combination of characteristics being controlled for). Moreover, a particular set of post-stratification weights can closely reflect the characteristics of the population as for some variables but not for others.

SHIW00 is post-stratified using the variables sex, age group, area and dimension of the town of residence (p. 40, Banca d'Italia, 2002). However, it is not clearly stated what methodology and which control totals (e.g. which age groups) were used. Table 1 shows how much the weighted sample differs from population totals using the post-stratification weight provided in SHIW00. It can be seen that, using the BI weights, the differences between the post-stratified and actual figures are small (less than 1%) as far as sex and area of residence (North-West (NW), North-East (NE), Center (C) and South (S)) are considered, but they become worryingly large for age groups (especially by area of residence) and schooling. Moreover the post-stratified figures obtained using SHIW00 weights are rather different from those obtained using MF00. This is probably due to the fact that SHIW00 weights are calculated using the control totals provided by ISTAT. Since ISTAT control totals differ from those provided by MF, there could be a problem with post-stratified simulations. For instance, the effects of an hypothetical tax policy that affected mainly the self-employed would probably be overestimated as these groups are over-represented using the BI weights. All these issues are of particular relevance whenever an analysis of income by population sub-groups is performed. For these reasons a set of alternative post-stratification weights were estimated using the same methodology as Atkinson *et al.* (1988) using control totals found in Ministero delle Finanze (2004) and ISTAT (2004). In particular the new post-stratified weights are consistent with ISTAT for the main socio-demographic characteristics, while they refer to MF00 data for the number and the geographical distribution of employed and self-employed. In this way the weighted SHIW00 can be considered consistent with the Italian population of taxpayers and can be used for a direct estimation of tax evasion.

[6] Negative weights are clearly nonsensical, as an interviewed household should have a weight at least equal to one in the national population. For a detailed discussion on the grossing-up issue and some sensitivity analysis using the distance (1), see Atkinson *et al.* (1988).

TABLE 1 - *Discrepancies between population and weighted SHIW sample. Year 2000.*

Variable	External sources*	BI		Our Weight	
	Totals (a)	Totals (b)	Diff (b/a-1)	Totals (c)	Diff (c/a-1)
Total population	57,844,017	57,828,424	0.0%	57,844,127	0.0%
Males	27,796,000	28,068,065	1.0%	27,796,064	0.0%
Females	30,048,017	29,760,359	-1.0%	30,048,063	0.0%
Pop NW	15,153,050	15,151,408	0.0%	15,153,060	0.0%
Pop NE	10,681,233	10,645,864	-0.3%	10,681,278	0.0%
Pop C	11,159,583	11,108,163	-0.5%	11,159,577	0.0%
Pop S	20,850,151	20,922,989	0.3%	20,850,212	0.0%
Age<=19	11,349,415	11,495,187	1.3%	11,349,436	0.0%
19<Age<=65	35,938,667	35,962,509	0.1%	35,938,782	0.0%
Age>65	10,555,935	10,370,728	-1.8%	10,555,909	0.0%
Age<=19 NW	2,562,196	2,837,546	10.7%	2,562,209	0.0%
Age<=19 NE	1,814,818	1,949,195	7.4%	1,814,820	0.0%
Age<=19 C	1,983,300	2,150,925	8.5%	1,983,298	0.0%
Age<=19 S	4,989,101	4,557,521	-8.7%	4,989,109	0.0%
19<Age<=65 NW	9,654,836	9,644,351	-0.1%	9,654,847	0.0%
19<Age<=65 NE	6,752,727	6,805,086	0.8%	6,752,770	0.0%
19<Age<=65 C	6,969,449	7,061,154	1.3%	6,969,457	0.0%
19<Age<=65 S	12,561,655	12,451,918	-0.9%	12,561,708	0.0%
Age>65 NW	2,936,018	2,669,511	-9.1%	2,936,004	0.0%
Age>65 NE	2,113,688	1,891,583	-10.5%	2,113,688	0.0%
Age>65 C	2,206,834	1,896,084	-14.1%	2,206,822	0.0%
Age>65 S	3,299,395	3,913,550	18.6%	3,299,395	0.0%
Employed **	17,723,376	15,563,640	-12.2%	17,723,450	0.0%
Employed NW **	5,278,461	4,600,891	-12.8%	5,278,487	0.0%
Employed NE **	3,843,039	3,408,590	-11.3%	3,843,050	0.0%
Employed C **	3,398,935	3,272,614	-3.7%	3,398,947	0.0%
Employed S **	5,202,941	4,281,545	-17.7%	5,202,966	0.0%
Self-employed **	5,837,998	6,471,973	10.9%	5,838,009	0.0%
Self-employed NW **	1,827,274	1,763,441	-3.5%	1,827,278	0.0%
Self-employed NE **	1,296,121	1,442,827	11.3%	1,296,120	0.0%
Self-employed C **	1,233,728	1,396,112	13.2%	1,233,725	0.0%
Self-employed S **	1,480,875	1,869,593	26.2%	1,480,886	0.0%
Employed	15,131,000	15,051,997	-0.5%	17,077,358	12.9%
Employed NW	4,616,000	4,464,179	-3.3%	5,091,679	10.3%
Employed NE	3,247,000	3,302,600	1.7%	3,712,218	14.3%
Employed C	3,050,000	3,213,664	5.4%	3,338,018	9.4%
Employed S	4,218,000	4,071,554	-3.5%	4,935,443	17.0%
Self-employed	5,949,000	5,867,783	-1.4%	5,216,206	-12.3%
Self-employed NW	1,678,000	1,598,946	-4.7%	1,614,862	-3.8%
Self-employed NE	1,367,000	1,297,700	-5.1%	1,146,121	-16.2%
Self-employed C	1,205,000	1,302,585	8.1%	1,147,294	-4.8%
Self-employed S	1,699,000	1,668,552	-1.8%	1,307,929	-23.0%
Elementary schooling	19,766,000	19,628,313	-0.7%	19,237,834	-2.7%
Compulsory schooling	16,556,000	15,624,763	-5.6%	15,666,562	-5.4%
High school degree	14,291,000	15,436,445	8.0%	15,732,990	10.1%
Laurea	3,546,000	3,799,227	7.1%	3,944,788	11.2%
Agricolture	1,120,000	1,310,029	17.0%	1,369,843	22.3%
Industry	6,767,000	7,051,688	4.2%	7,576,260	12.0%
Services	13,193,000	12,362,652	-6.3%	13,163,325	-0.2%

Source: Our calculations on SHIW00

*External sources from ISTAT (2004) where not differently specified; ** External sources from MF00

4. ESTIMATING TAX EVASION: THE RESULTS

Using a direct approach to tax evasion estimation, the positive difference between disposable income from voluntary survey and tax forms data is considered concealed taxable income, which gives rise to tax evasion. The direct estimation methodology is applied to employment and self-employment income, as defined in Section 3.2, net of all taxes and compulsory social contributions.

We believe that this difference is a conservative measure of underground income (i.e. income from tax avoidance and evasion concerning legal activities) and of informal income (i.e. income from individual activity with a low level of organization, based on individual and familiar relationships, such as baby sitting, domestic cleaning, etc.). We are also convinced that criminal or illegal income (i.e. from tax evasion due to illegal activities such as drug trafficking or unauthorized medical practice) is not included, as those participating in such activities are unlikely to accept the interview or respond truthfully. For an analogous reason we believe there are no total tax evaders among the respondents to SHIW00.

The exercise we perform here is similar to that of Bernasconi and Marenzi (1997) and it can be seen as an update of that paper. However, our work differs from theirs in that we focus only on employment and self-employment[7] income at the individual level and because we carefully consider the issue of post-stratification to population totals.

Although the methodology was applied comparing SHIW00 and MF00 by centiles, results are shown in deciles for clarity reasons. The percentage of concealed income is estimated as the ratio between the mean net income within a SHIW00 centile and the mean after-tax income within the corresponding MF00 centile. Table 2 shows the estimates of concealed income for employee and self-employed in the whole sample by deciles. The percentage of concealed income is very high in both cases in the first three deciles of the distribution (although it is higher for the self-employed than for the employed). In the last six deciles the two distributions show instead a different pattern. Concealed employment income falls to zero, and becomes negative after the median; self-employment income shows a positive (though decreasing) pattern of concealed income. At any decile, a larger proportion of self- employment income is concealed compared to employment income. The difference is larger around the median – about 27% – and it is smaller at low levels of income – less than 10% – (see Figure 1).

This pattern of concealed income is robust with respect to geographical areas and age groups. The main relevant difference with respect to income concealment at the national level is that the self-employed in the South show a percentage of concealed income considerably higher (around 10% more across income deciles) than those in the North and in the Center. This difference does not clearly arise for employment income[8].

[7] Recall definitions provided in Section 3.2.

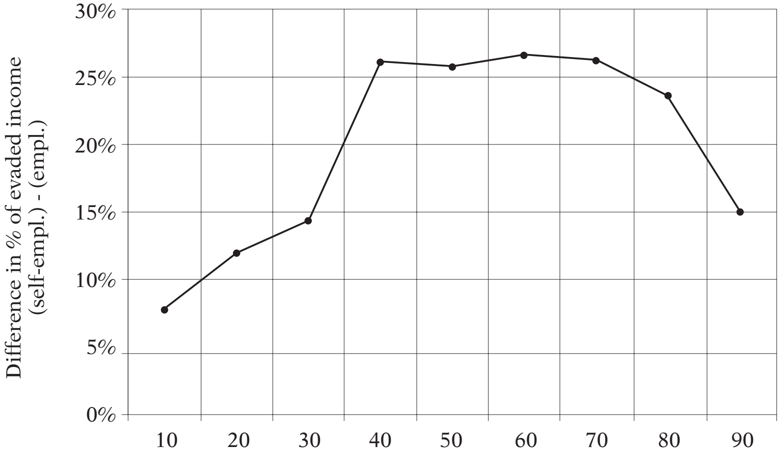
[8] For more details on estimation of income concealment by geographical areas and age groups, see the extended version of this paper (Fiorio and D'Amuri, 2005b).

TABLE 2 - *Tax evasion (in %) by type of labor income. Our calculations on SHIW00 and MF00, with post-stratified weight.*

Percentile	Employment income				Self employment income			
	BI employment income	MF employment income	Evasion (euro)	Evasion (%)	BI self employment income	MF self employment income	Evasion (euro)	Evasion (%)
10	3956	1470	2486	62,8%	2526	744	1783	70,6%
20	6592	3829	2763	41,9%	4453	2055	2398	53,9%
30	9506	7195	2311	24,3%	6401	3927	2474	38,6%
40	10853	9990	863	7,9%	8734	5755	2979	34,1%
50	12139	11907	233	1,9%	10532	7616	2916	27,7%
60	13066	13315	-249	-1,9%	12776	9618	3159	24,7%
70	14101	14732	-631	-4,5%	15377	12024	3352	21,8%
80	15453	16390	-937	-6,1%	18780	15483	3298	17,6%
90	17745	19016	-1271	-7,2%	24100	22220	1880	7,8%

Source: Our calculations on SHIW00 and MF00

FIGURE 1 - *Difference in % evasion between self-employment income and employment income. Our calculations on SHIW00 and MF00, with post-stratified weight).*



4.1 "Striking results! Any explanation?"

Let us discuss our main results[9]. Some readers might be surprised by the large proportion of taxable income that is concealed and some even by the mere existence of tax evasion among employees. Others by the fact that the share of hidden employment income becomes increasingly negative with income after the median.

[9] In a previous version of this paper (Fiorio and D'Amuri, 2005a) we applied the same methodology comparing the 1998 and the 2000 SHIW data sets with the tables aggregated by the Ministry of Finance (Ministero delle Finanze, 2004, 2002) from the tax forms received in each year and obtained similar results.

As we rule out the possibility that rich taxpayers declare to the tax authorities more income than they actually earn, we believe that a possible explanation of the negative percentage of concealed income over the median can be the fact that the response rate to SHIW00 by rich households decreases with income. Hence the under-sampling of these households reduces the thickness of the higher tail of the income distribution in SHIW00 and makes estimation more and more conservative as income becomes larger than the median. Our results are then consistent with those provided by Cannari and D'Alessio (1992) and D'Alessio and Faiella (2002) that propensity to accept an interview is decreasing with income (recall Section 3.3). However, our results seem to suggest that a certain level of under-sampling arises also among the poorest households. If the response rate was not decreasing with income but it was inverted-U-shaped, then also the bottom part of the income distribution would be less thick and evasion estimation biased upward at lower levels of income[10]. There could be possible reasons for a correlation of income to the response rate (e.g. value of time is increasing with income, lack of confidence or shame on own achievements could be large in poorer households), but further research on the response rate to SHIWs is beyond the scope of this paper[11]. A second plausible explanation of these results refers to the psychological attitude of individuals who are asked about their labor income. It might be that low-income individuals prefer to over-declare their true income and, conversely, high-income individuals prefer to under-declare it, as they might feel more comfortable with declaring a closer-to-median income. With this explanation we could see median income as a sort of social norm that individuals feel uneasy to break. This would imply that income declared in an interview is more concentrated around the median than it actually is.

However, if we assume that the rate of response is similar among the same percentiles of the employed and self-employed distribution, we should also accept the message of Figure 1: the self-employed hide between 7% and 27% more income from tax authorities than employees. It should moreover be noted that while self-employed workers are likely to under-declare their income to reduce their personal income tax and social contributions payment, employees' motivations might be different and not completely under their own control: badly-qualified employees might be induced to accept irregular jobs as the only alternative to remaining unemployed.

---

[10] For instance, the SHIW00 does not collect information on institutionalized individuals, who are very likely to have low incomes.

[11] Post-stratification on income using the distribution provided in MF00 would solve the under-sampling bias but would also cancel differences of incomes between SHIW00 and MF00, reducing direct estimation of tax evasion to zero for all income levels.

Although a non-response or misreporting bias could increase the estimation of tax evasion at low levels of income, some evidence provided by independent governmental agencies provides some support to our results. For instance, as for the high level of evasion among the first deciles of the employment income distribution, an inspection run in 2003 in 145,000 firms by INPS (National Institute of Social Protection) showed that about 63% of them were employing irregular workers or irregular payment methods (e.g. side-payments), mostly in order to avoid the payment of compulsory social contributions (INPS, 2004). Moreover, according to estimates of underground economy by ISTAT, irregularly employed workers[12] in 2000 were 3 millions (14% of total employed workforce), and this phenomenon was particularly concentrated in the agricultural sector (over 32%), in the building sector (over 20%) and in the trade and transportation services sectors (about 18%) (ISTAT, 2003).

How consistent are our results with the evidence found by INPS (2004) and ISTAT (2003)? Table 3 shows the distribution of employees in the first four deciles of the income distribution – with the highest employment tax evasion – by some sectors. It clearly emerges that low income employees are more likely to be employed in sectors such as agriculture, building, trade and transportation, and domestic services. For instance, the percentage of employees in agriculture is four times more concentrated in the first decile than in the whole distribution. Over 23% of the poorest employment income earners work in the agricultural sector.

There is a final striking point in our results: it emerges clearly that there is tax evasion not only among the self-employed – which was expected from available evidence – but also among employees, especially in lower deciles. Who are these employees who evade taxes?

Provided all the caveats concerning the reliability of SHIW00 data are kept in mind, the data allow us to draw a picture of the employees by income levels. By estimating a (log) wage equation conditional on individual characteristics using SHIW00 for employees only, we can have a clearer idea of the characteristics of employees at lower levels of income, where tax evasion is a more relevant issue. Here, we focus on employees only, as we want to characterize them better, rather then compare them to the self-employed.

OLS estimate the mean of log wage conditional on individual characteristics. They show that the average log wage is higher if the individual is the householder, it is increasing with age and experience and their

[12] It must be noted that, in contrast to our assumptions, irregular workers recorded by ISTAT include also workers who are unknown to the tax authorities.

TABLE 3 - *Distribution of employees among some activity sectors.*

Sector of economic activity	Income percentile					
	10	20	30	40	First four deciles	All the distribution
Agricolture and fishing	23.3%	11.1%	8.5%	2.9%	11.5%	5.6%
Industry	26.0%	37.1%	40.5%	48.3%	38.0%	36.4%
Costructions	9.4%	6.5%	6.8%	7.4%	7.5%	5.8%
Services	50.7%	51.8%	51.0%	48.8%	50.6%	58.0%
Commerce, car services, lodging and catering	13.4%	15.3%	16.8%	12.2%	14.4%	10.8%
Domestic services	9.3%	12.9%	6.6%	6.5%	8.8%	4.6%

Sector of economic activity	Income percentile					
	10	20	30	40	First 4 deciles	All the distribution
Agricolture and fishing	416	199	153	53	205	100
Industry	71	102	111	133	104	100
Constructions	164	113	118	128	131	100
Services	87	89	88	84	87	100
Commerce, car services, lodging and catering	124	142	155	113	133	100
Domestic services	204	282	145	142	193	100

Source: Our calculations on SHIW00, ISTAT 2003

returns are moderately decreasing[13], females earn 21% less than males, employees from the South earn 17% and 10% less than those from the North and the Center, respectively. Also the occupation variables are highly significant and present a reasonable sign and magnitude (e.g. managers' log wage is ceteris paribus 70% higher than blue-collar workers'). The sector-of-occupation variables show that *ceteris paribus*, employees in agriculture and domestic services to household earn the least. Finally, the average wage is 41% higher if the occupation is full time, it is 15% higher if it is tenured, it is 58% higher if it is occupied for a whole year. Quantile regression allows the study of the effect of a set of covariates on the whole distribution (Koenker and Bassett, 1978). Hence, using quantile regression it is also possible to document whether the picture provided by OLS is uniform across income levels or whether covariates have different effects depending on the income level considered. The statistical model used here specifies the τ-th quantile of the conditional distribution of (log) wages (y) given a set of covariates (x) as a linear function of the covariates:

$$Q_\tau(y_i|x) = \alpha(\tau) + x'\beta(\tau), \text{ with } \tau \in (0,1) \tag{2}$$

[13] The square of age and experience are highly significant but very small.

It shows that lower deciles are more likely to be occupied by young people, by residents in the Center, in the South and by people with part-time and temporary occupations. At lower quantiles the wage premium of being a manager is much smaller than the average (Figure 2)[14].

5. Effects of tax evasion on incidence and progressivity indices

Results presented in Section 4 raise a set of concerns about incidence and progressivity of taxation: they show that the proportion of income that is hidden from tax authorities is not evenly distributed across different income groups and across different types of income.

As in Bernasconi and Marenzi (1997) we assume that people belonging to the same income interval and with similar type of income have the same compliance behavior. In other words, we suppose that there is no re-ranking and horizontal inequity within the same income interval but only across them. "It is evident that the difference between the two data sources (BI and MF), computed by decile, is interpretable as evasion in the decile only if it is assumed that evasion did not move taxpayers from one decile to another before and after tax evasion" (Bernasconi and Marenzi, 1997, p. 24).

As SHIW00 records only disposable income after all taxes and social contributions, a microsimulation model is used to simulate before-tax income as well as to perform simulations. The TABEITA00 model used is a static microsimulation model that simulates personal taxation (IRPEF and "imposte sostitutive") using the SHIW00[15]. In order to avoid overestimating the amount of tax paid and of before-tax income, only income net of (imputed) tax evasion is used to simulate before-tax evasion. The imputation of tax evasion is taken from a table similar[16] to Table 2.

Results of the simulation are analyzed using some indices for the measurement of the effects of taxation and the Gini coefficient, before and after taxes. The Kakwani index is a very popular index of progressivity: it measures the departure from proportionality as the difference between the concentration coefficient of tax C_t and the Gini index of before-tax income, G_y:

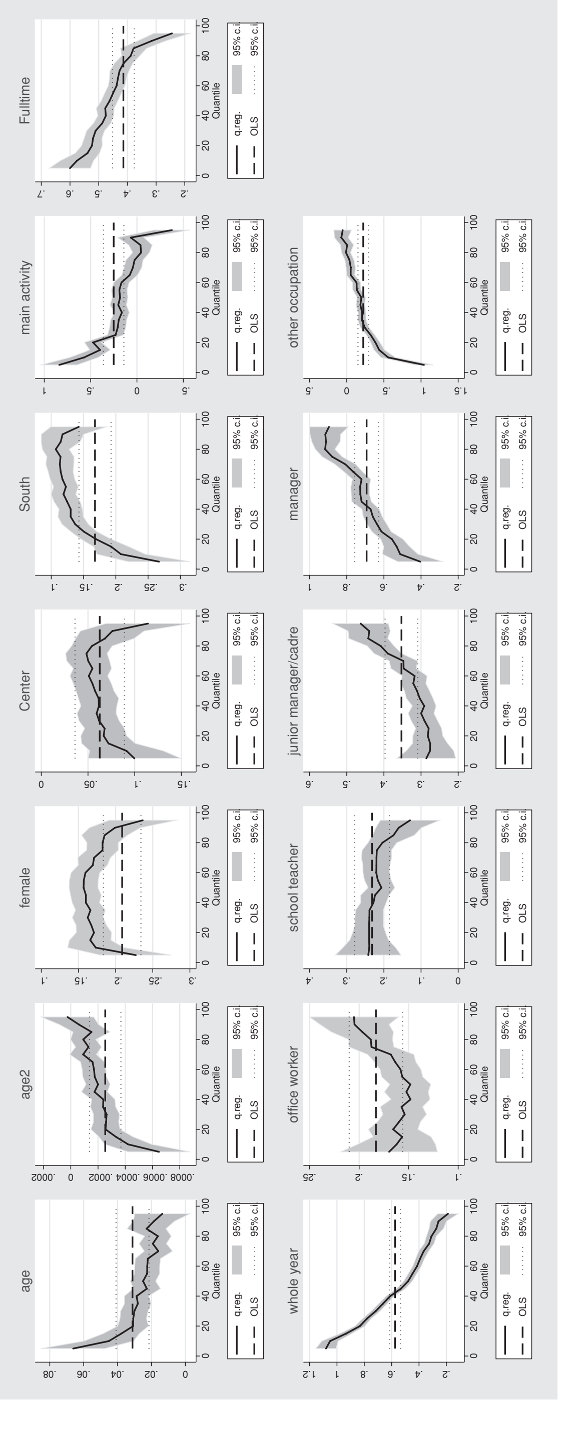
$$K_t = C_t - G_y \tag{3}$$

[14] For full details of the estimation results see Fiorio and D'Amuri (2005b).

[15] For details of the model, see Fiorio and D'Amuri (2004).

[16] The only difference is that Table 2 shows estimation by deciles while imputation is by centiles.

FIGURE 2 - *Selected quantile regression coefficients of a wage equation. Employees only*



For large samples the minimum value of the Kakwani index is – $(1 + G_y)$ and the maximum value is $1 + G_y$. The first case happens when the poorest person pays all the tax ($C_t = -1$), the second when all the tax is paid by the richest person, leading to maximal progressivity (Kakwani, 1977).

The redistributive effect looks at the shift from before-tax to after-tax income. With no re-ranking, the after-tax Lorenz curve coincides with the after-tax income concentration curve. The Reynolds-Smolensky index (*RS*) is equal to the difference between the Gini coefficient of before-tax income (G_y) and the concentration coefficient of after-tax income (C_{y-t}) (Reynolds and Smolensky, 1977). In the absence of reranking it is the reduction of the Gini coefficient achieved by the tax. It is also equal to the product of a progressivity index (e.g. K_t) and the average tax on net income (t/1–t):

$$RS = G_y - C_{y-t} = \frac{t}{1-t} K_t \tag{4}$$

Hence the redistributive effect is determined by disproportionality and tax incidence. However, as the re-ranking effects are likely to occur with the tax system, the Reynolds-Smolonsky index, which is an indicator of vertical equity, should be written as the sum of a redistributive effect (RE) and a re-ranking effect (RR) (Lambert, 1993, p. 185):

$$RS = RE + RR = (G_y - C_{y-t} - G_{y-t} + C_{y-t}) + (G_{y-t} - C_{y-t}) \tag{5}$$

These indices allow us to measure the importance of our estimates of labor tax evasion on the progressivity, redistribution and incidence of personal income taxation. Consistently with the rest of the paper, we focussed on labor income receivers only. Let us denote the evaded income – as estimated in Section 4 – with E, the before-tax income declared to fiscal authorities with BT, the net personal income tax applying the tax code on BT with PIT, the after-tax income with AT (AT = BT – PIT). Finally we denote the true BT (TBT = BT + E) with TBT, and the true AT (i.e. the disposable income after all PIT is paid on TBT) with TAT. Using the estimates of Section 4, two sets of incomes were then simulated: (a) the "*status quo*", where some conceal part of their labor income; (b) the "*ideal world*", where no income is concealed. The following exercise was performed:

The "*status quo*":

STEP 1 The net income as declared in SHIW and deducted by E is fed into TABEITA00 to obtain BT;

STEP 2 Using TABEITA00, PIT and AT are obtained as an output of the microsimulation model.
STEP 3 TBT and TAT are then computed adding E to both BT and AT.

The "*ideal world*":
STEP 1 as in "*status quo*";
STEP 2 TBT is computed as the sum of BT and E.
STEP 3 TABEITA00 is used on TBT to compute TAT.

Step 1 in the exercise above is necessary as, by definition, no tax is paid on E and SHIW00 data do not provide information about PIT paid. TBT and TAT are then used in (5) to analyze the effects of income concealment on personal income taxation.

Table 4 presents results. In the "*status quo*", the tax relative to net income is equal to 24.4% and it would be 5.1% higher if no income was concealed. In other words, the average tax rate would increase from 19.6% to 20.4%, increasing tax revenues (only from those receiving labor income) by 3.9%. The redistributive effect would however be slightly lower in the "*ideal world*" as the departure from proportionality (the K_t index) would be about 9% lower in the "*ideal world*". This result – together with a (slightly) higher Gini index in the "*ideal world*" after-tax income – might be puzzling to some readers, who would expect evasion to increase inequities. However they should recall that the direct methodology applied in Section 4 found much evasion in the lower deciles of employment income. Moreover these indices were computed using the subsample of those declaring some labor income only: labor tax evasion would certainly have a lower effect when the whole sample was considered.

6. POLICY INDICATIONS AND CONCLUDING REMARKS

This paper provides an estimate of labor tax evasion in Italy in the year 2000 using a direct methodology. The analysis was performed on labor income only, as other types of income are recorded with much greater ap-

TABLE 4 - *Indices of progressivity, redistribution and tax incidence.*

	RS	t/(1-t)	K	RR	Gini BT	Gini AT
ideal world: no income concealed	0.0511	0.2560	0.1996	0.0007	0.3983	0.3479
status quo: some income concealed	0.0540	0.2437	0.2215	0.0015	0.3983	0.3458

Source: our calculations on SHIW00 using TABEITA00 proximations in available survey data. After detailed analysis of the two data sets concerning the variable definition, the verification of the main hypothesis of the methodology and the post-stratification to population totals, results show the following: (a) employment income is hidden to tax authorities at low levels of income, but the evasion rate decreases constantly and becomes negligible above the median income; (b) self-employed taxpayers hide a consistently larger share of their income than employees in the corresponding income decile. The difference is over 7% for bottom deciles and increases to around 27% around the mode, where the estimates are less likely to be affected by under- sampling and misreporting bias. Although our estimates of tax evasion at bottom deciles could be biased upward because of SHIW00 under-sampling problem or the misreporting of income for psychological reasons, our story is consistent with that found by other independent institutions.

But who are these previously unnoticed tax evaders? What are their motivations to evade personal income tax? What public policies can be undertaken to reduce a phenomenon that is likely to cause inequalities among equals? Here we fully answer to the first question only, and we suggest some answers to the following ones. Using regression tools we show that those more likely to be in the bottom part of the income distribution are young, female, unmarried, blue-collar workers, working in the agricultural sector or providing domestic services. Moreover we also show that fixed-term employees and working part-time are more likely to receive lower-than-average wages.

A partial answer to the motivation of employees to evade personal income tax comes from an analysis of tax incidence and tax progressivity, which was performed using a tax-benefit microsimulation model. Focussing on the sub-sample of working taxpayers, if no tax evasion occurred at any level of employment and self-employment income, the average tax rate would be 4% higher, the tax incidence on net incomes would be 5.1% higher. The tax would however be less progressive, inducing an overall negative redistributive effect.

In other words, labor income tax evasion allows some workers, namely those in part-time, fixed-term, under-qualified occupations to supplement their meagre income. Shall we then turn a blind eye, or "smile indulgently" on tax evasion, as Cowell (1987, p. 195) was asking? Cowell suggests that considerations of horizontal equity might be an essential prop to the anti-evasion argument. We also add some final considerations. People who do not pay taxes generally also evade social contributions, thus reducing the chances of maintaining their income if or when in need. This might be the rational choice of a forward looking agent who gambles on his chances of living in good health and of receiving good returns from investing her hidden income in performing capital markets. Or, it might be the only option for someone who does not have a strong bargaining power on the job market. If the second possibility cannot be ruled out – as our knowledge of the Italian context suggests – there is room for discussion about what to do next. We believe that an indulgent smile on evasion should not be an option.

References

Atkinson, A. B., Gomulka, J., and Sutherland, H. (1988), Grossing-up FES data for Tax-Benefit Models, pp. 223-253, Atkinson and Sutherland (1988), Tax-Benefit Models, STICERD Occasional Paper, No. 10, London, London School of Economics, STICERD.

Banca d'Italia (2002), *I bilanci delle famiglie italiane nell'anno 2000*, volume XII of *Supplementi al Bollettino Statistico*. *Note metodologiche e informazioni statistiche*, Banca d'Italia, Roma.

Bernasconi, M. and Marenzi, A. (1997), Gli effetti redistributivi dell'evasione fiscale in Italia, *Ricerche quantitative per la politica economica*, pp. 1-38, Convegno CIDE-SADIBA, Banca d'Italia, novembre.

Bordignon, M. and Zanardi, A. (1997), Tax evasion in Italy, *Giornale degli Economisti e Annali di Economia*, 56(3-4):169-210.

Brandolini, A. (1999), The distribution of personal income in post-war Italy: Source description, data quality, and the time pattern of income inequality, *Giornale degli Economisti e Annali di Economia*, 58:183-239.

Brandolini, A. and Cannari, L. (1994), "Methodological appendix: The Bank of Italys Survey of Household Income and Wealth", in Ando, Guiso, and Visco, editors, *Saving and the Accumulation of Income and Wealth. Essays on Italian Households and Government Saving Behaviour*.

Calzaroni, M. (2000), L'occupazione come strumento per la stima esaustiva del PIL e la misura del sommerso, contributo al seminario dell'ISTAT, Roma, 12-13, gennaio.

Cannari, L., Ceriani, V., and D'Alessio, G. (1995), Il recupero degli imponibili sottratti a tassazione, *Ricerche quantitative per la politica economica*. Sadiba, Banca d'Italia, 3-5 novembre.

Cannari, L. and D'Alessio, G. (1990), Housing assets in the Bank of Italy's Survey of Income and Wealth, in Dagum and Zenga, editors, *Income and Wealth Distribution, Inequality and Poverty*, pp. 326-334, Berlin, Springer Verlag.

Cannari, L. and D'Alessio, G. (1992), Mancate interviste e distorsione degli stimatori, *Temi di discussione del Servizio Studi, Banca d'Italia* (172).

Cowell, F. A. (1980), On the structure of additive inequality measures, *Review of Economic Studies*, 47:521-31.

COWELL, F. A. (1987), The economic analysis of tax evasion, in Hey, J. D. and Lambert, P. J., editors, *Surveys in The Economics of Uncertainty*, Basil Blackwell, Oxford.

D'ALESSIO, G. and FAIELLA, I. (2002), Non-response behavior in the Bank of Italy's Survey of Household Income and Wealth, *Temi di discussione del Servizio Studi* (462), Banca d'Italia.

DI NICOLA, F. and MONTEDURO, M. T. (2004), Un modello di microsimulazione per l'imposta sul reddito personale, Technical report, Se.C.I.T., Rome. Downloadable from http://www.secit.finanze.it/approfondimenti.php.

FIORIO, C. V. and D'AMURI, F. (2004), TABEITA: a tax-benefit microsimulation model for Italian households. mimeo.

FIORIO, C. V. and D'AMURI, F. (2005a), Analysis of tax evasion effects in Italy using a tax-benefit microsimulation model, in Bosco, B. and Pisauro, G., editors, *Politiche pubbliche, Sviluppo e Crescita,* Franco Angeli, Milano. Downloadable from http://www.econpubblica.uni-bocconi.it/fiorio.

FIORIO, C. V. and D'AMURI, F. (2005b), Workers' tax evasion in Italy, *Econpubblica Working Papers, Università Bocconi* (104). Downloadable from: http:/papers.ssrn.comsol3papers.cfm?abstract id=702922.

HOLLENBECK, K. (1976), An algorithm for adjusting n-dimensional tabular data to conform to general linear constraints, *Proceedings of the American Statistical Association*, pp. 402-405.

INPS (2004), Rapporto annuale 2003. Technical report, INPS, Rome.

ISTAT (2003), La misura dell'occupazione non regolare nelle stime di contabilità a nazionale: un'analisi al livello nazionale e regionale, Rome.

ISTAT (2004), *Popolazione & statistiche demografiche*, ISTAT, Istituto nazionale di statistica, Rome. Data available at http://demo.istat.it/.

KAKWANI, N. (1977), Measurement of tax progressivity: an international comparison, *Economic Journal*, 87:71-80.

KOENKER, R. and BASSETT, G. (1978), Regression quantiles, *Econometrica*, 46:33-50.

LAMBERT, P. J. (1993), *The distribution and redistribution of income: a mathematical analysis*, Manchester University Press, Manchester, 2nd edition.

LITTLE, R. J. A. and RUBIN, D. B. (1987), *Statistical Analysis with Missing Data*. Wiley, New York.

MARENZI, A. (1996), Prime analisi sulla distribuzione dell'evasione dell'IRPEF per categorie di contribuenti e per livelli di reddito, pp. 305-341, in Rapporto CNEL, il Mulino, Bologna.

MERZ, J. (1983), Die konsistente hockrechnung von mikrodaten nach dem prinzip des minimalen informationsverlustes, *Allgemeines Statistisches Archiv*, 76:342-366.

MERZ, J. (1985), Microdata adjustment under minimum information loss with a special modified newton-raphson procedure, University of Frankfurt.

MINISTERO DELLE FINANZE (2002), *Analisi delle dichiarazioni dei redditi. Il modello Unico 1999/redditi 1998: le persone fisiche*, Ministero delle Finanze, Rome.

MINISTERO DELLE FINANZE (2004), *Analisi delle dichiarazioni dei redditi. Il modello Unico 2001/redditi 2000: le persone fisiche*, Ministero delle Finanze, Rome.

PISSARIDES, C. A. and WEBER, G. (1989), An expenditure-based estimate of britain's black economy, *Journal of Public Economics*, 39:17-32.

REYNOLDS, M. and SMOLENSKY, E. (1977), *Public expenditures, taxes and the distribution of income: the United States, 1950, 1961, 1970*, Academic Press, New York.

SCHNEIDER, F. (2000a), The increase of the size of the shadow economy of 18 OECD countries: some preliminary explanations, *IFO Working Paper*, (306).

SCHNEIDER, F. (2000b), The value added of underground activities: Size and measurement of the shadow economies and shadow economy labor force all over the world, *World Bank Working Paper,* Washington, D.C.

SUTHERLAND, H. (1989), Constructing a Tax-Benefit Model: What Advice Can One Give?, in *Taxation, incentive and the distribution of income*, number 141, STICERD, London School of Economics.

ZIZZA, R. (2002), Metodologie di stima dell'economia sommersa: un'applicazione al caso italiano, *Temi di discussione del Servizio Studi* (463), Banca d'Italia.