TAX EVASION AND SELF-EMPLOYMENT IN A HIGH-TAX COUNTRY: EVIDENCE FROM SWEDEN

PER ENGSTRÖM
BERTIL HOLMLUND

TAX EVASION AND SELF-EMPLOYMENT IN A HIGH-TAX COUNTRY: EVIDENCE FROM SWEDEN

Abstract

Self-employed individuals have arguably greater opportunities than wage earners to underreport their incomes. The incentives for underreporting should be especially strong in an economy with generally high taxes. This paper uses recent income and expenditure data to examine the extent of underreporting of income among self-employed individuals in Sweden. A key hypothesis is that underreporting of incomes among the self-employed would be visible in the data as "excess food consumption", for a given level of observed income. Our results confirm the underreporting hypothesis. In particular, we estimate that households with at least one self-employed member underreport their total incomes by around 30 percent. Underreporting appears to be twice as prevalent among self-employed people with unincorporated businesses as among those with incorporated businesses.

JEL Code: D12, H24, H25, H26.

Keywords: tax evasion, self-employment, Engel curves.

Per Engström
Department of Economics
Uppsala University
Box 513
751 20 Uppsala
Sweden
per.engstrom@nek.uu.se

Bertil Holmlund
Department of Economics
Uppsala University
Box 513
751 20 Uppsala
Sweden
bertil.holmlund@nek.uu.se

1. Introduction

It is widely believed that self-employed people can more easily underreport their taxable incomes than wage earners are able to do. Several researchers have tried to estimate the extent of underreporting by using information on consumption expenditure and reported income. The key idea is that expenditure on food is accurately reported by all people participating in an expenditure survey, or at least that there is no systematic misreporting that is related to employment status. If the self-employed underreport their incomes, this would be visible in the data as "excess food consumption" among self-employed people, for a given level of observed income. If one can get an estimate of the extent of underreporting, it is also possible to get an estimate of the size of the black economy relative to GDP that is related to self-employment activities.

The expenditure-based approach to estimation of tax evasion and the size of the black economy was pioneered by Pissarides and Weber (1989) and has been applied in a number of other studies. Pissarides and Weber examined British data from 1982; Apel (1994) provided evidence for Sweden using data from 1988; Johansson (2000) used Finnish data from the mid-1990s; Schuetze (2002) examined consumption and income in Canada over the period 1969-1992; Lyssiotou et al (2004) investigated consumption and income using British data from 1993. The studies by Apel, Johansson and Schuetze follow the Pissarides and Weber approach closely, whereas Lyssiotou et al develop a somewhat different methodology based on consumer demand systems. An alternative nonparametric methodology is developed by Tedd (2004).

The available empirical studies generally support the basic hypothesis: self-employed people do tend to underreport their incomes relative to wage earners. There is, however, considerable dispersion in the estimates of underreporting. Pissarides and Weber (1989) estimate the mean underreporting to around 55 percent, whereas Apel (1994) report a mid-point estimate of around 35 percent. Schuetze (2002) reports underreporting in the interval 11 to 23 percent, whereas Johansson (2000) estimates underreporting by some 27 percent on average. Lyssiotou et al (2004) estimate that true income for blue collar self-employed people is more than 100 percent greater than reported income, whereas true income for white-collar self-employed people exceeds reported income by 64 percent. The implications of these various estimates on the relative size of the black economy depend to the relative importance of self-employment in the economy. Pissarides and Weber conclude that black economy activities

related to self-employment amount to 5 percent of GDP in the United Kingdom; Apel gets an estimate for Sweden of only 1 percent, whereas Johansson gets 2 percent for Finland. The estimates in Lyssiotou et al imply that close to 11 percent of the UK economy involves black economy activities related to self-employment. The dispersion in estimates is thus considerable, and the reasons for the different estimates are not well understood.

This paper takes a new look at consumption expenditure and reported income among self-employed people and wage earners in Sweden. Sweden has the highest tax-to-GDP ratio in the world, and there is a presumption that the incentives for tax evasion would be particularly pertinent in such an economy.[1] We use two samples, the first for the years 1999-2001 and the second for 2003-2004. Our results confirm the underreporting hypothesis. Specifically, we estimate that the degree of income underreporting hovers around 30 percent of household income for a household with at least one self-employed member.

We also examine whether the degree of underreporting varies by the legal form of self-employment and find striking differences between incorporated and unincorporated businesses: self-employed in the latter category are much more prone to underreporting than those in the incorporated category. We interpret these differences as the outcome of differences in regulations that entail higher costs of tax evasion associated with incorporated business activities. These results based on expenditure patterns are consistent with observed earnings differences between paid employees and self-employed individuals. Controlling for human capital characteristics and industry affiliation, we find that self-employed individuals with unincorporated businesses report substantially lower incomes than employees as well as lower incomes than self-employed in the incorporated category.

The next section discusses briefly the Pissarides and Weber methodology. Section 3 of the paper presents the data, section 4 gives the results and section 5 concludes.

2. Estimating Tax Evasion from Expenditure Data

The Pissarides and Weber approach can be illustrated by means Figure 1. Let c denote log food consumption, $c = \ln C^F$, and y log disposable income, $y = \ln Y^D$. The figure shows two

---

[1] According to OECD data, the tax-to-GDP ratio in Sweden stood at 50.7 percent in 2004 which is the highest ratio among OECD countries; Denmark came second at 49.6 percent. Agell et al (1996) provide an overview and evaluation of the major 1991 reform of the Swedish tax system. Sandmo (2005) presents a recent survey of the theory of tax evasion.

log-linear consumption-income profiles (Engel curves), one for self-employed people (*SE*) and one for wage earners (*WE*). By assumption, the elasticity of consumption with respect to income is equal for the two groups; the Engel curves have thus the same slopes. However, we allow the intercepts to differ so as to capture the possibility of underreporting of income among self-employed people. Imagine two individuals, one self-employed and one wage earner, who both report consumption level $c*$. The wage earner reports income level y^w whereas the self-employed person reports income y^s. Suppose that consumption is truthfully reported by both types and that income corresponds to true income for the wage earner but not for the self-employed person. If preferences are the same, we can infer that the self-employed person's true income is y^w rather than reported y^s. The extent of underreporting is thus given by $y^w - y^s$.

Figure 1. Engel curves for wage earners (*WE*) and self-employed people (*SE*)
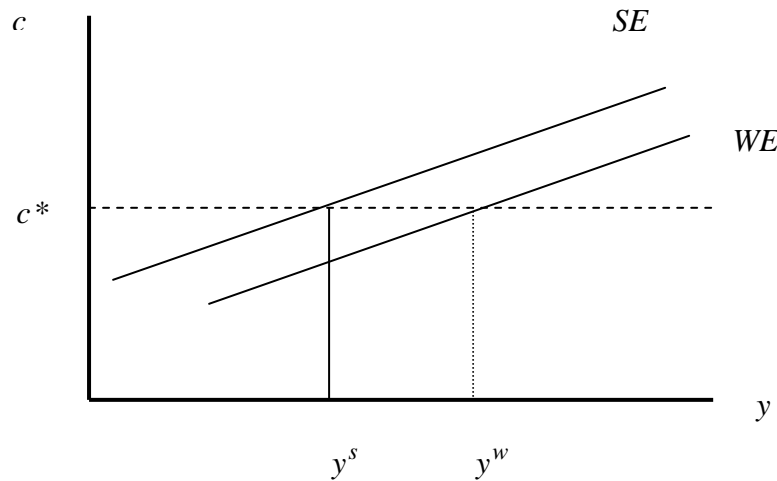
Now suppose that we pool the data for self-employed people and wage earners and estimate an equation of the form:

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

where subscript *i* denotes individual *i*, *X* is a vector of variables affecting consumption (in addition to income), *SE* is a dummy variable for self-employed persons and ε is a random

error term. The parameter γ captures the vertical distance between the two Engel curves; $\gamma > 0$ implies some underreporting of income among self-employed people. The degree of underreporting (in logs) is obtained as $y^w - y^s = \gamma / \beta$. For example, if $\gamma = .05$ and $\beta = 0.5$, we get $\gamma / \beta = 0.10$, which says that self-employed people underreport their incomes by around 10 percent. The percentage difference is generally obtained as $D \equiv 100 \times \left[\exp(\gamma / \beta) - 1 \right]$. We might alternatively express underreporting in terms of $k \equiv \exp(\gamma / \beta)$, which gives the number by which a self-employed person's disposable income has to be multiplied so as to get the true disposable income. With $\gamma = .05$ and $\beta = 0.5$, we get $k=1.105$.

The estimate of underreporting obtained from eq. (1) pertains to disposable income. We might alternatively express underreporting in terms of gross income. Let Y^* denote true gross income, U the amount of underreported income and define $\kappa \equiv Y^* /(Y^* - U)$ as the ratio between true and reported gross income. It follows that $\kappa = k(1-t)+t$, where t is the effective tax rate.[2] Clearly, we have $\kappa < k$.

Estimation of (1) raises several issues that are discussed in Pissarides and Weber (1989) and some of the subsequent literature. One issue is how income should be interpreted and treated in the estimation. The most relevant income concept is arguably permanent income rather than current income. However, existing data sets include only data on current incomes. Pissarides and Weber, as well as much of the subsequent literature, have attempted to resolve this problem by treating current income as endogenous and pursued instrumental variable estimation. A potential problem with this approach is that it relies on perhaps arbitrary exclusion restrictions for identification: one needs to find variables that affect disposable income without directly affecting also food consumption. We will present results from OLS as well as from IV estimations; in fact, the results are very similar.

---

[2] By effective tax rate we mean $t=1$ - (disposable income/total factor income). Transfers to households add to disposable income and thus reduces the effective (average) tax rate. Reported disposable income, Y^D, is given as $Y^D = Y^* - t(Y^* - U)$. We have $\kappa \equiv Y^* /(Y^* - U)$ and k is the ratio between true disposable income and reported disposable income. We thus get $\kappa = k(1-t)+t$.

3. The Data

We have used data from the Swedish Household Budget Survey (Hushållens utgifter, HUT) from 1999-2001, 2003 and 2004.[3] Data for 1999-2001 are based on surveys to households that include the second quarter of 1999 up to the first quarter of 2002. Around 9000 randomly selected households were approached over these years, but only slightly above 50 percent of those did actually participate in the surveys. In 2003 and 2004, 4000 households were approached each year and the response rate was 58 percent both years. The response rates are thus fairly low, although not exceptionally low for these kinds of time intensive surveys.

The 1999-2001 surveys are produced and presented by Statistics Sweden as one survey and we will refer to it as HUT 1999-2001. The other two surveys are produced and presented by Statistics Sweden as two separate surveys. We will concatenate the surveys for 2003 and 2004 and refer to the resulting linked data set as HUT 2003-2004. The HUT data contain no panel elements.

The participating households were asked to report their consumption expenditures during randomly selected two-week periods. In addition, supplementary questions were asked regarding expenditures over the past 12 months. Various other questions were asked so as to get information on household characteristics, including employment status and occupation. Information about disposable incomes is obtained from official income and tax registers and merged with the expenditure data.

The two key variables are annual food consumption and annual disposable income. We follow the literature by focusing on food consumption on the assumption that this is truthfully reported and that expenditure functions for food are similar for employees and self-employed people. Our measure of food consumption is reported in the data as "food purchases".[4] Disposable income is based on all types of (register-based) incomes, including transfers. Taxes are deducted from gross income so as to get household disposable income.

We focus on consumption and incomes among employees and self-employed people. Self-employment status can be defined in several ways. Most previous studies in this genre have

---

[3] The design and main results of the HUT-studies are presented in reports from Statistics Sweden (2003, 2004, 2005).

[4] We have obtained similar results for broader measures of food consumption, such as food purchases plus meals out.

classified households by using information on income shares attributed to paid employment and self-employment.[5] A problem with this approach, besides the difficulty of choosing the borderline, is that it may be sensitive to the legal form of the business. A person who runs an incorporated business may be formally employed by the company and receive the main part of the compensation in the form of wage income.

We prefer to make use of self-reported information on employment status of the individual household members. As a requirement for sample inclusion we require either that (i) at least one member of the household is employed or self-employed, or that (ii) two members are employed or self-employed. The self-employed category includes incorporated as well as unincorporated businesses and the data contain information about the legal form. We have excluded farmers, however, assuming that their food purchases exhibits a pattern relative to income that differs from other self-employed people. Following Pissarides and Weber (1989) and others in the literature, we also restrict the analysis to households with married or cohabiting couples. The bigger sample, fulfilling criterion (i), comprises around 3000 households for 1999-2001 as well as for 2003-2004. The smaller sample (ii) includes a little more than 2000 households for 1999-2001 as well as for 2003-2004. The share of self-employed varies between 13 and 16 percent according to the adopted definitions.[6] A first look at the data is given in Table 1.

The table displays average food consumption and disposable incomes among employees and self-employed in our sample. A household is here classified as self-employed if at least one member is self-employed. Average incomes are slightly higher among employees, whereas food consumption is higher among the self-employed. It is thus clear that food consumption relative to income is higher, on average, among self-employed people. The difference amounts to around 4 percentage points.[7] This pattern is what we would expect according to the hypothesis that self-employed people are more likely to underreport their true incomes. Of

---

[5] Pissarides and Weber (1989) define households as self-employed if income from self-employment accounts for at least 25 percent of total income.

[6] The fraction of self-employed in total employment is 10 percent according to the labor force surveys of 2000. Note that our sample is restricted to couples. Moreover, our measure of the rate of self-employment in households is not directly compared to the labor force survey data based on individuals.

[7] The food consumption ratios are substantially lower in 2003-2004 than in the previous surveys, a pattern also visible in the aggregate data published by Statistics Sweden. The likely main reason for the differences is that measurement techniques have changed. According to Statistics Sweden (2003), the surveys for 1999-2001 produced an upward bias in food consumption by inducing households to include some non-food items in their reports of food purchases.

course, this is just a first look at the data and we need to control for other variables that may affect consumption. This is the topic for the next section.

Table 1. Means of variables.

	1999 -- 2001				2003 -- 2004			
	Households with at least one member employed or self-employed		Households with at least two members employed or self-employed		Households with at least one member employed or self-employed		Households with at least two members employed or self-employed	
	Employees	Self-employed	Employees	Self-employed	Employees	Self-employed	Employees	Self-employed
$\ln Y^D$	12.736	12.6560	12.788	12.721	12.860	12.774	12.924	12.812
$\ln C^F$	10.783	10.852	10.818	10.865	10.646	10.682	10.689	10.708
C^F / Y^D	0.161	0.204	0.155	0.194	0.122	0.158	0.119	0.155
# obs.	2684	379	2052	314	2516	425	1874	363

Notes: Disposable income (Y^D) and food consumption (C^F) is in units of SEK, 2001 prices for 1999-2001 and current prices for 2003-2004. A household is classified as self-employed if at least one member is self-employed. A few households with income above two million SEK are excluded.

4. Estimation Results

We have estimated a number of models along the lines of eq. (1). That is, we regress log food consumption on log disposable income and a bunch of control variables. The latter include age and age squared, type of housing, number of children under age 20, average years of education among adult household members, and region of residence.[8] We use two alternative definitions of self-employment status. One variable (SE_1), with associated parameter γ_1, is a dummy for households where at least one member is self-employed. The second variable

---

[8] Age is the age of the "household head", which is the person with the highest income in the household; type of housing is a dummy for single family housing; regions are so called H-regions, which capture the degree of urbanization.

(*SETOT*), with associated parameter γ_2, is the number of persons who are self-employed in the household; this variable takes only three values in our sample, viz. 0, 1 or 2.

Preliminary analyses revealed very similar results for the 1999-2001 sample and the 2003-2004 sample. We have therefore decided to concatenate all data, thus obtaining a data set comprising five years. The estimation results for the basic specification are set out in Table 2.

Table 2. Estimation results for the pooled data. Dependent variable: $\ln C^F$

	Households with at least one member employed or self-employed				Households with at least two members employed or self-employed			
	OLS (1)	OLS (2)	IV (3)	IV (4)	OLS (5)	OLS (6)	IV (7)	IV (8)
$\ln Y^D$ (β)	.233 (10.75)	.234 (10.87)	.259 (3.52)	.258 (4.44)	.219 (8.36)	.222 (8.48)	.185 (2.26)	.184 (2.59)
SE_1 (γ_1)	.065 (3.55)		.070 (3.53)		.056 (2.92)		.054 (2.60)	
$SETOT$ (γ_2)		.064 (3.98)		.069 (4.35)		.055 (3.32)		.054 (3.14)
$k \equiv \exp(\gamma_j / \beta)$	1.32	1.31	1.31	1.31	1.29	1.28	1.34	1.34
R^2	.204	.204	.203	.204	.192	.193	.191	.191
# obs.	6004	6004	5965	5965	4603	4603	4574	4574

Notes: *SE*_1: at least one self-employed; *SETOT*: number of self-employed $\{0,1,2\}$. The other controls are age, age squared, number of children, single family house, average years of schooling in the household, six dummies for H-region and a dummy for 2003-2004. The instruments used in the IV estimations are income from capital and property taxes. Households with incomes above 2 million SEK are excluded. Robust standard errors, absolute *t*-values in parentheses.

In the IV estimations we use two variables to achieve identification, viz. income from capital and property taxes. Property taxes are closely related to the market value of owned housing (including housing used for recreational purposes). A comparison across columns shows that the coefficients pertaining to self-employment status are fairly robust as well as statistically significant at conventional levels. The interpretation is that self-employed households are characterized by "excess food consumption" by 5 to 7 percent. That is, they spend 5 to 7 percent more on food relative to wage earners with the same reported income.

The extent of underreporting, as measured by $k \equiv \exp(\gamma_j / \beta)$, varies in a fairly narrow band around 1.30. That is, the degree of underreporting amounts to 30 percent. Our estimates of underreporting are in the same (big) ballpark as those reported in most previous studies (with the exception of Lyssiotou et al, 2004).

These estimates of underreporting are expressed in terms of disposable income. Using the fact that the ratio between true and reported gross income is given by $\kappa = k(1-t) + t$, we obtain $\kappa \approx 1.23$ in our data. Effective tax rates are fairly low in our sample, reflecting the fact that taxes to a substantial degree are used to finance transfers to households.

Self-employment status according to our definitions does not exactly correspond to the share of household income that can be attributed to self-employment. Recall that $k=1.30$ means that households with self-employed members underreport their disposable incomes by 30 percent. However, self-employed households typically include also employees with wage incomes so 30 percent is not an estimate of how much underreporting of *self-employment income* there is. To arrive at such an estimate, we define self-employment income as including (i) entrepreneurial income[9] and (ii) wage income for households with incorporated businesses (on the assumption that owners of incorporated businesses are employed by their companies). Let μ denote the ratio between this (adjusted) measure of reported gross income from self-employment and total reported gross income among self-employed households. Under-reporting of income from self-employment can be derived as

$$(2) \qquad D^s = \frac{U}{U + Y^s} = \frac{\kappa - 1}{\mu + \kappa - 1}$$

---

[9] In Swedish, this concept is known as "inkomst av näringsverksamhet".

where Y^s is reported income from self-employment in self-employment households (assuming that underreporting is only relevant for self-employment income).

In our data for 2003-2004, we have $\mu \approx 0.43$ so we obtain $D^s \approx 0.35$ using $\kappa = 1.23$.[11] Thus, we can conclude that households with at least one self-employed member underreport their incomes from self-employment by roughly 35 percent. Note that this is a conservative estimate since μ is an upper bound of self-employment income relative to total income.

Heterogeneous Effects

Among the self-employed in our sample, around 50 percent runs an incorporated business. Those with incorporated businesses are presumably employed by their company and paid by the company. Indeed, our data reveal that wage incomes account for a much higher share of gross income among self-employed households with incorporated businesses than among those with unincorporated businesses.

Incorporated businesses are circumscribed by other and more elaborate legal rules than those relevant for unincorporated business activities. The transactions pertaining to the corporation and those pertaining to the owner should be kept apart. An annual financial report, which becomes public information, should be delivered for each financial year. Moreover, a certified auditor must be appointed and the auditor's report on the financial transactions of the company is public information. For unincorporated businesses, on the other hand, the borderline between the owner's and the company's financial transactions is fuzzy and the transactions are less easily scrutinized by outsiders. These differences in the legal form of self-employment are likely to influence incentives for tax evasion. There is a presumption that the more stringent legislation pertaining to incorporated businesses means higher costs of tax evasion.

We have examined whether the legal form of self-employment matters by including variables for incorporated and unincorporated businesses, respectively. The definitions are analogous to those already defined. For example, *SE*_1 (inc.) takes the value of one when at least one household member runs an incorporated business. The results are given in Table 3.

---

[11] It is more difficult to disentangle income sources in the 1999-2001 data so we focus on 2003-2004.

Table 3. Underreporting by the legal form of self-employment. Dependent variable: $\ln C^F$.

	Households with at least one member employed or self-employed				Households with at least two members employed or self-employed			
	OLS (1)	OLS (2)	IV (3)	IV (4)	OLS (5)	OLS (6)	IV (7)	IV (8)
$\ln Y^D$	.237 (10.92)	.237 (10.97)	.269 (.075)	.268 (3.56)	.224 (8.56)	.225 (8.61)	.195 (2.33)	.194 (2.32)
SE_1 (inc.)	.039 (1.58)		.040 (1.62)		.028 (1.09)		.030 (1.16)	
SE_1 (uninc.)	.089 (3.47)		.098 (3.37)		.084 (3.12)		.079 (2.55)	
$SETOT$ (inc.)		.041 (1.92)		.042 (1.96)		.031 (1.40)		.032 (1.45)
$SETOT$ (uninc.)		.086 (3.76)		.095 (3.65)		.080 (3.40)		.078 (2.82)
$k \equiv \exp(\gamma_j / \beta)$ (inc.)	1.18	1.19	1.16	1.17	1.13	1.15	1.17	1.18
$k \equiv \exp(\gamma_j / \beta)$ (uninc.)	1.46	1.44	1.44	1.43	1.45	1.43	1.50	1.49
R^2	.204	.204	.204	.204	.192	.193	.192	.192
# obs.	6004	6004	5965	5965	4603	4603	4603	4603

Notes: *SE*_1: at least one self-employed (incorporated, unincorporated); *SETOT*: number of self-employeed (incorporated, unincorporated). The controls are age, age squared, number of children, single family house, average years of schooling in the household, six dummies for H-region and a dummy for 2003-2004. The instruments used in the IV estimations are income from capital and property taxes. Households with incomes above 2 million SEK are excluded. Robust standard errors, absolute *t*-values in parentheses.

There is a clear pattern in the results suggesting that self-employment associated with incorporated businesses involves less tax evasion than self-employment in general. Self-employed with incorporated businesses underreport some 15 to 20 percent of their incomes, although these estimates are typically only marginally significant. Self-employed with

unincorporated businesses underreport at the rate of 40 to 50 percent. These results are consistent with the hypothesis that incorporated businesses face higher costs of tax evasion as a result of more detailed regulations.

We can also in this case ask how large proportion of self-employment income that is hidden from the authorities. Among households with unincorporated businesses, entrepreneurial income accounts on average for roughly 40 percent of total household income. To get an estimate of the amount of hidden entrepreneurial income, we use (2) and arrive at $D^s \approx 0.50$. That is, self-employed households with unincorporated businesses hide around 50 percent of their true incomes.

We have also checked for heterogeneous age effects by including interactions between the self-employment variables and a dummy for household heads over the median age (which is 44 in the data). However, we did not find any evidence that the propensity to underreport varies by age. Analogously, we have checked whether the effect varies by gender by interacting self-employment status and a gender dummy (equal to one if the self-employed person is a woman). We found no statistical significance for this interaction variable. We also checked whether skilled blue collar workers are more prone to tax evasion relative to other employees, a hypothesis motivated by anecdotal evidence as well as some previous empirical findings (see Pissarides and Weber, 1989, and Persson, 2006). However, we could not find any systematic evidence supporting this hypothesis.

Corroborating Evidence: Income and Employment Status
An arguably straightforward approach to examine tax evasion by employment status would be to look at differences in reported income between employees and self-employed individuals with similar measurable human capital and other characteristics. That is, we could estimate models of the form

(3) $\ln Y_i = Z_i \delta + \sigma SE_i + \eta_i$

where Y is gross income, i.e., income before taxes and transfers. Z is a vector of personal characteristics and SE a dummy for (individual) self-employment status. A negative sign of

σ would be consistent with underreporting of income among the self-employed.[12] Of course, a negative estimate could also reflect compensating income differences, which could be present to the extent that individuals place a positive value of self-employment status because it may confer non-pecuniary benefits such as personal independence ("being your own boss" etcetera).[13] On the other hand, the self-employed are known to work longer hours than the typical employees, a fact that would contribute to higher annual earnings.[14]

We have estimated a number of equations of the type given by (3) on data for 2004, focusing as before on married (or cohabiting) individuals aged 20 to 64. The specifications include, inter alia, some 50 industry dummies as well as six regional dummies. The results are displayed in Table 4. The estimates in the first three columns are based on a narrow income concept, whereas the last three columns are based on a broad concept that includes all incomes from labor and capital.

The estimates imply that self-employed individuals earn substantially lower incomes than wage earners. For given characteristics, the level of (reported) income is around 40-50 percent lower among the self-employed individuals with unincorporated businesses.[18] For individuals with incorporated businesses, the negative "effect" is smaller (in absolute value) and at best only marginally significant

The patterns set out in Table 4 are consistent with what we found in the analysis of consumption behavior, i.e., strong evidence of underreporting for households with unincorporated businesses but only weak evidence for households with incorporated businesses. The results in Table 4 might, however, reflect compensating income differences associated with employment status as well as underreporting of incomes. Although this seems likely, it is not clear why the legal form of a business ownership should matter much for job satisfaction.

---

[12] Persson (2005) compares earnings among employees and self-employed using Swedish data for 2002 and find substantially lower reported earnings among the self-employed even after having standardized for industry affiliation. The paper does not distinguish between incorporated and unincorporated businesses, however.

[13] See Hamilton (2000) for a study of the returns to self-employment in the United States. Hamilton finds a substantial "earnings penalty" associated with self-employment, a result that he interprets as evidence of nonpecuniary self-employment benefits. Blanchflower (2004) surveys the literature on self-employment and reports that job satisfaction is higher among self-employed than among wage earners.

[14] According to the Swedish labor force surveys for 2004, paid employees worked on average 35 hours per week whereas self-employed people worked 43 hours.

[18] The estimates in column (1) imply $\exp(-.701) - 1 \approx -0.5$, whereas the estimates in column (4) imply $\exp(-.500) - 1 \approx -0.4$.

Table 4. Income and self-employment status 2004. Dependent variable: $\ln Y$.

	Y: wages plus entrepreneurial income			*Y*: total factor income		
	Both sexes (1)	Men (2)	Women (3)	Both sexes (4)	Men (5)	Women (6)
Age	.118 (8.63)	.092 (5.58)	.142 (6.31)	.115 (8.63)	.096 (5.51)	.129 (6.30)
Age squared/100	-.123 (7.73)	-.097 (5.08)	-.148 (5.56)	-.119 (7.69)	-.099 (4.95)	-.132 (5.55)
Education (yrs)	.066 (9.93)	.077 (8.81)	.056 (5.47)	.069 (10.54)	.077 (8.45)	.063 (6.38)
Female	-.449 (13.13)			-.440 (3.10)		
# children	-.043 (2.92)	.008 (.045)	-.075 (3.05)	-.044 (3.10)	-.017 (.92)	-.068 (2.86)
SE – uninc.	-.701 (3.51)	-.642 (2.92)	-.766 (2.01)	-.500 (2.97)	-.392 (2.26)	-.751 (1.93)
SE – inc.	-.191 (1.79)	-.194 (1.64)	-.226 (.94)	-.060 (.55)	-.054 (.45)	-.144 (.55)
R^2	.256	.243	.178	.265	.227	.190
# obs.	2548	1306	1242	2551	1307	1444

Notes: All specifications include region dummies as well as a full set of (feasible) industry dummies (maximum 52). Total factor income include – in addition to wages and entrepreneurial income – dividends, interest payments and capital gains. Robust standard errors, absolute *t*-values in the parentheses.

5. Concluding Remarks

Our study of food expenditure and incomes among employees and self-employed has produced results that are broadly consistent with findings from previous studies. There is strong evidence of "excess food consumption" among the self-employed, for a given level of disposable income. A plausible interpretation is that the pattern reflects underreporting of

incomes among the self-employed. The magnitude of underreporting is non-trivial: our estimates suggest that the households with at least one member underreport their total incomes at the average rate of 30 percent and their incomes from self-employment by around 35 percent. Our results also imply that tax evasion is most pronounced among self-employed who run unincorporated businesses. We interpret this finding as indicating higher costs of tax evasion for owners of incorporated businesses since their transactions are more easily exposed to public scrutiny.

Our look at income differentials between employees and self-employed offers results consistent with the analysis of consumption data. Reported incomes among the self-employed, and in particular those with unincorporated businesses, are substantially lower than incomes among employees even after controlling for human capital characteristics and industry affiliation at a fairly detailed level. Although this pattern could reflect many different forces, including compensating earnings differences, it seems likely that tax evasion behavior is part of the story. In fact, longer work hours would tend to create an earnings advantage (rather than a penalty) for the self-employed.

There is no obvious policy prescription that follows from our results. However, the fact that tax evasion seems to depend on the legal form of self-employment is of potential policy relevance. The regulatory framework pertaining to small firms is subject to ongoing policy discussions, often emphasizing the need to simplify the regulations so as to facilitate entrepreneurship. One issue concerns whether independent auditing should be required for all incorporated businesses, including (very) small firms. The details of the regulatory framework matter for incentives for tax evasion, a consideration that needs to be acknowledged in the policy discussions.

References

Agell, J, P Englund and J Södersten (1996), Tax Reform of the Century – the Swedish Experiment, *National Tax Journal* 49, 643-664.

Apel, M (1994), An Expenditure-Based Estimate of Tax Evasion in Sweden, in M Apel, *Essays on Taxation and Economic Behavior*, Economic Studies 18, Department of Economics, Uppsala University.

Blanchflower, D (2004), Self-Employment: More May Not Be Better, *Swedish Economic Policy Review* 11, 15-73.

Hamilton, B (2000), Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment, *Journal of Political Economy* 108, 604-631.

Johansson, E (2000), An Expenditure-Based Estimation of Self-Employment Income Underreporting in Finland, Working Paper 433, Swedish School of Economics and Business Administration, Helsinki.

Lyssiotou, P, P Pashardes and T Stengos (2004), Estimates of the Black Economy Based on Consumer Demand Approaches, *Economic Journal* 114, 622-640.

Persson, A (2005), Om företagarnas faktiskt redovisade arbetsinkomster och andra indikatorer på deras ekonomiska standard (Reported incomes and other indicators of living standards among the self-employed), Skatteekonomiska meddelanden No 48, The Swedish National Tax Board.

Persson, A (2006), Svartarbetets omfattning och struktur (The extent and structure of black market activities), Skatteekonomiska meddelanden No 56, The Swedish National Tax Board.

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

Sandmo, A (2005), The Theory of Tax Evasion: A Retrospective View, *National Tax Journal* 58, 643-663.

Schuetze, H (2002), Profiles of Tax Noncompliance Among the Self-Employed in Canada: 1969-1992, *Canadian Public Policy* 28, 219-238.

Statistics Sweden (2003), *Hushållens utgifter* 1999-2001 (The Household Budget Survey 1999-2001), Stockholm.

Statistics Sweden (2004), *Hushållens utgifter* 2003 (The Household Budget Survey 2003), PR 35 SM 0402, Stockholm.

Statistics Sweden (2005), *Hushållens utgifter* 2004 (The Household Budget Survey 2004), HE 35 SM 0502, Stockholm.

Tedds, L (2004), Noparametric Expenditure-Based Estimation of Income-Underreporting and the Underground Economy, Working Paper 2004:17, McMaster University.

CESifo Working Paper Series

(for full list see www.cesifo-group.de)

_____