



TECHNICAL ASSISTANCE REPORT

SWEDEN

Estimating Income Underreporting Among
Self-Employed Taxpayers: A Time Series of
Basic Pissarides–Weber Estimates, 2003–2021

JUNE 2026

Prepared By
Per Engström and Johannes Hagen

Authoring Department

Fiscal Affairs Department

The contents of this report constitute technical advice provided by the staff of the International Monetary Fund (IMF) to the authorities of Sweden (the "TA recipient") in response to their request for technical assistance. This report (in whole or in part) or summaries thereof may be disclosed by the IMF to IMF Executive Directors and members of their staff, as well as to other agencies or instrumentalities of the TA recipient, and upon their request, to World Bank staff and other technical assistance providers and donors with legitimate interest, unless the TA recipient specifically objects to such disclosure. Disclosure of this report (in whole or in part) or summaries thereof to parties outside the IMF other than agencies or instrumentalities of the TA recipient, World Bank staff, other technical assistance providers and donors with legitimate interest shall require the explicit consent of the TA recipient and the IMF's Fiscal Affairs Department.

The results and opinions presented do not constitute official statements of the Swedish Tax Agency (STA).

[This technical assistance \(TA\) was provided with financial support from the European Union.](#)

Contents

Abbreviations and Acronyms	4
Preface	5
Executive Summary	6
Recommendations	7
I. Introduction	8
II. Methodology	9
A. The Expenditure-Based Approach	9
B. Limitations of the Current Implementation.....	9
C. Data	10
III. Key Findings.....	11
IV. Next Steps	14

Figures

Figure 1. Basic Pissarides–Weber Estimates of Self-Employed Income Underreporting	11
---	----

Tables

Table 1. OLS Estimates by HUT Vintage and Pooled across All Years	12
---	----

Annexes

Annex I. Technical Details of the Expenditure-Based Approach	15
Annex II. Materials Left with Skatteverket.....	17

Abbreviations and Acronyms

CD	Capacity Development
EC	European Commission
EU	European Union
FAD	IMF Fiscal Affairs Department
GDP	Gross Domestic Product
HUT	Swedish Household Budget Survey (Hushållens utgifter)
IMF	International Monetary Fund
IV	Instrumental Variable
OLS	Ordinary Least Squares
PIH	Permanent Income Hypothesis
PIT	Personal Income Tax
PW	Pissarides–Weber
SE	Self-Employed / Self-Employment
SEK	Swedish Krona
Skatteverket	Swedish Tax Agency
SME	Small and Medium-sized Enterprise
VAT	Value Added Tax

Preface

In response to a request from the Swedish Tax Agency (Skatteverket), a duty-station-based capacity development (CD) mission team comprising Mr. Per Engström (Uppsala University, FAD short-term expert), and Mr. Johannes Hagen (Uppsala University, FAD short-term expert) provided technical assistance to Skatteverket to estimate the degree of income underreporting among self-employed (SE) taxpayers in Sweden. The mission was financed by the European Union under the DG REFORM Program.

The analysis applies the expenditure-based methodology originally developed by Pissarides and Weber (PW) (1989)¹ — a workhorse approach in the empirical literature on self-employed tax evasion — to all available vintages of the Swedish Household Budget Survey (HUT). Estimates are produced separately for each vintage as well as pooled across all years, providing a time series of PW estimates of self-employed income underreporting in Sweden over the period 2003–2021.

The IMF team expresses its sincere appreciation to Skatteverket for the excellent cooperation and support provided before and during this mission. The team particularly acknowledges the support provided by Mr. Damián Migueles Chazarreta, PhD, Skatteverket, who prepared and analyzed the data underlying this report and was instrumental in the success of the mission.

This report represents the draft submitted to Mr. Damián Migueles Chazarreta on April 28, 2026. It consists of an Executive Summary and the following sections: (I) Introduction; (II) Methodology; (III) Key Findings; and (IV) Next Steps.

¹ Pissarides, C.A. and Weber, G. (1989). An expenditure-based estimate of Britain's black economy. *Journal of Public Economics*, 39(1), 17–32.

Executive Summary

This mission produced a time series of basic Pissarides–Weber (PW) estimates of income underreporting among self-employed taxpayers in Sweden, covering all available vintages of the Household Budget Survey (HUT) over the period 2003–2021. The PW method exploits the fact that self-employed taxpayers, if they underreport income to the tax authorities, will tend to spend more on food than wage earners at the same level of reported income. By comparing how much self-employed taxpayers spend on food relative to wage earners at the same reported income level, and how food spending generally changes with income, the method estimates the share of income that is actually reported (the compliance rate) — and, by extension, the share that is hidden.²

These basic estimates should be interpreted with caution and should not be taken as point estimates of the true level of self-employed underreporting in any given year. The estimates are based on current-year disposable income rather than a permanent income measure, and they include only a limited set of controls (gender of household head, age, age squared, and home-ownership). As discussed in Engström and Hagen (2017),³ using current income in place of permanent income affects the estimated income elasticity and, in turn, the implied underreporting rate. Using current income typically overstates true underreporting and exhibit considerable year-to-year volatility driven in part by transitory income fluctuations rather than genuine changes in self-employed compliance.

The pooled estimate across all HUT vintages indicates that self-employed households spend approximately 9 percent more on food than wage earners at the same level of reported disposable income, implying a basic underreporting rate of roughly 38 percent. Year-by-year estimates of the underreporting rate ($1 - \kappa$) naturally fluctuates a lot, with confidence intervals that are often wide. Despite the volatility, the basic estimates suggest a higher level of self-employed underreporting in the later vintages (2007 onward) than in the earliest vintages (2003–2006).

Although the level estimates are imprecise, the relative time pattern may carry useful information about the evolution of self-employed underreporting over time. To the extent that the bias from omitting permanent income operates in a broadly similar manner across vintages, year-on-year changes in the basic estimates may track underlying changes in compliance even when the absolute levels do not. This is an important, however cautious, basis for treating the time series as a useful diagnostic instrument for Skatteverket.

As the next step, it is recommended that Skatteverket updates these basic PW estimates in connection with each new HUT vintage. Maintaining a consistent time series of basic estimates would provide a low-cost monitoring tool for self-employed compliance trends. It is further recommended that work continue toward incorporating a permanent income correction along the lines of Engström and Hagen (2017), which would substantially strengthen the interpretation of the level estimates. It is also recommended to explore the possibilities of register-based proxies of consumption and other traces of real income. No single method for measuring income underreporting is free of limitations, and a portfolio of methods with distinct identifying assumptions will give a more reliable overall picture than any single method alone.

² In technical terms, the estimated intercept shift in the food Engel curve, combined with the estimated food income elasticity, yields an implied compliance rate κ and a corresponding underreporting rate $1 - \kappa$.

³ Engström, P. and Hagen, J. (2017). Income underreporting among the self-employed: A permanent income approach. *European Economic Review*, 92, 92–109.

Recommendations

No.	Recommendation	Due Date
1	Update the simple PW estimates of self-employed income underreporting in connection with each new HUT survey vintage, and compile the resulting time series. Although individual estimates should not be interpreted as point estimates of hidden income, the relative time pattern may be informative about the evolution of self-employed underreporting in Sweden.	Each HUT vintage
2	Treat the present estimates as one input among several. The PW method has known limitations, and a portfolio of methods with distinct identifying assumptions will give a more reliable overall picture than any single method alone.	Ongoing
3	Continue work on incorporating a permanent income correction along the lines of Engström and Hagen (2017). This is a methodological refinement that can substantially improve the credibility of the level estimates and is a natural next step once the necessary register linkages have been established.	2026–2027
4	Invest in training for Skatteverket data analysts to enable independent replication and extension of the methodology, including its application to future survey vintages.	Ongoing

I. Introduction

1. Income underreporting among self-employed taxpayers is a major source of revenue loss for Sweden's tax administration. Unlike wage earners, whose income is automatically reported to Skatteverket by employers under Sweden's third-party reporting system, self-employed taxpayers retain wide latitude in determining how much income to declare. Estimates from the Danish audit experiment of Kleven et al and from Skatteverket's own tax gap analysis indicate that a disproportionate share of the total tax gap is attributable to self-employment income.⁴⁵

2. The expenditure-based (Pissarides–Weber) method is one of the most established approaches to estimating income underreporting among the self-employed. The intuition is simple: if self-employed individuals systematically underreport their income to tax authorities while truthfully reporting their food expenditures in household surveys, they will appear to consume more food than wage earners at the same level of reported income. By comparing the food Engel curves of self-employed and wage earner households, it is possible to back out an implied underreporting factor.⁶ The method has been applied in many countries and remains a useful diagnostic instrument, even though the level estimates it produces are sensitive to a range of methodological choices.

3. This mission applies a simple version of the PW method to all available vintages of the Swedish Household Budget Survey (HUT), covering the years 2003–2009, 2012, and 2021. The estimates are based on current-year disposable income and a parsimonious set of controls (gender of household head, age, age squared, and home-ownership). Estimates are produced separately for each HUT vintage and pooled across all years. The result is a time series of basic PW estimates of self-employed underreporting in Sweden over almost two decades.

4. The level estimates produced by this exercise should be interpreted with caution. As emphasized by Engström and Hagen (2017), failure to account for the distinction between current and permanent income mechanically affects the estimated food income elasticity and, in turn, the implied underreporting rate (see section The literature has typically addressed this issue using instrumental variable (IV) techniques, but the validity of available instruments is difficult to verify without access to longitudinal income data. In the present exercise, neither a permanent income correction nor IV methods have been applied. The level estimates are therefore best understood as basic in nature, i.e., informative about orders of magnitude and about relative time patterns, but not as precise point estimates of hidden income in any given year.

⁴ Kleven, H.J., Knudsen, M.B., Kreiner, C.T., Pedersen, S., and Saez, E. (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, 79(3), 651–692.

⁵ Skatteverket (2022). *Skattefelsrapporten 2022*. Swedish Tax Agency, Stockholm.

⁶ The food Engel curve describes the relationship between a household's income and its spending on food. It builds on "Engel's law" — the empirical observation that as income rises, the share of income spent on food tends to fall.

II. Methodology

A. The Expenditure-Based Approach

5. **The PW method identifies income underreporting by comparing the food consumption Engel curves that describe the relationship between a household's income and its spending on a particular good or service, holding prices constant of self-employed and wage earner households.** Let c denote log food consumption and y log disposable income. The estimating equation is:

$$c_{it} = X_i\alpha + \beta y_{it} + \gamma SE_{it} + \varepsilon_{it} \quad (1)$$

6. **Here, c_{it} is consumption of household i at time t , y_{it} is household current income, X_i is a vector of household characteristics (in this implementation: gender of household head, age, age squared, and a home-ownership dummy), SE_{it} is a dummy variable equal to one for self-employed households, and ε_{it} is a random error term.** The parameter γ captures the upward shift in the food Engel curve for the self-employed: at a given reported income level, these households spend more on food. This pattern is interpreted as reflecting underreported income. The fraction of true income reported, i.e., the compliance rate, κ , is given by $\exp(-\gamma/\beta)$, so that estimated underreporting equals $1 - \kappa$.⁷ Standard errors for $1 - \kappa$ are obtained by the delta method, assuming β and γ as approximately independent.⁸

B. Limitations of the Current Implementation

7. **The implementation in this report is intentionally simple and has known limitations.** Three limitations are particularly relevant for the interpretation of the level estimates:

- a. **Current rather than permanent income.** Standard economic theory—particularly the Permanent Income Hypothesis—implies that food consumption depends on permanent rather than current income. Current income is volatile, reflecting transitory shocks that households largely smooth through consumption. As a result, it is only weakly correlated with actual consumption. Using current income as a proxy therefore introduces measurement error, which attenuates the estimated food income elasticity β toward zero... Since the estimated underreporting factor $1 - \kappa$ is a decreasing function of β , this attenuation typically results in an upward bias in the estimated level of underreporting. Engström and Hagen (2017) document, using equivalent Swedish data, that the food income elasticity rises by more than 40 percent when current income is replaced by a seven-year average measure, with a corresponding fall in the estimated underreporting rate of approximately one-third. The level estimates reported here are therefore likely to overstate true underreporting.
- b. **Limited set of controls.** The estimating equation includes only gender of household head, age, age squared, and home-ownership. A richer specification—including household composition, education, region, and other characteristics—would likely tighten the estimates and may shift the implied levels

⁷ The estimated amount of income underreporting is given by $1 - \kappa$, where $\kappa = \exp(-\gamma/\beta)$ is the fraction of true income reported by the self-employed, constructed from the estimated intercept shift γ and income elasticity β .

⁸ Standard errors for $1 - \kappa$ are computed as the square root of $(\partial g/\partial \beta)^2 \cdot \text{Var}(\beta) + (\partial g/\partial \gamma)^2 \cdot \text{Var}(\gamma)$, where $g(\beta, \gamma) = \exp(-\gamma/\beta)$. The independence assumption is a simplification: more precise estimates of the standard error would require bootstrap methods.

somewhat. The limited control set was chosen to keep the analysis tractable within the time available; refinements to the specification are a natural extension.

- c. **No correction for differential income volatility.** Self-employed households typically face more volatile incomes than wage earners. As discussed in Pissarides and Weber (1989), this differential volatility can affect the relationship between reported and true income for the two groups. The standard correction for this issue (described in detail in Engström and Hagen, 2017) has not been applied in the current implementation.

8. Despite these limitations, the time series of basic estimates may still carry useful information about the evolution of self-employed underreporting. If the bias from omitting permanent income operates in a broadly similar manner across HUT vintages, year-on-year changes in the basic estimates may track underlying changes in compliance even when the absolute levels do not. This logic underpins the recommendation, set out in Section IV, that Skatteverket update the basic PW estimates with each new HUT vintage and treat the resulting time series as a low-cost monitoring tool — while continuing work toward a permanent income correction that would strengthen the level interpretation.

C. Data

9. The data come from the Swedish Household Budget Survey (HUT), which is conducted by Statistics Sweden (SCB). The HUT provides detailed information on household consumption expenditures collected through a two-week expense diary, supplemented by information on household demographics and (in the more recent vintages) administrative income data. The mission uses all available HUT vintages: 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2012, and 2021. The pooled estimation sample comprises 20,529 households.⁹

10. Self-employed households are identified using a binary variable, an indicator equal to one if at least one household member has positive income from active self-employment (sole proprietorship, partnership, or active ownership of a closely held corporation) in the HUT survey year.¹⁰ The food consumption variable measures annualized household food expenditures derived from the two-week expense diary. The income variable is log household disposable income for the survey year. Descriptive statistics for the estimation sample are reported in Table A1.

⁹ The HUT survey is conducted by Statistics Sweden (SCB) on a periodic basis. The survey was conducted annually from 2003 to 2009 and was discontinued thereafter; it was reinstated for 2012 and most recently for 2021.

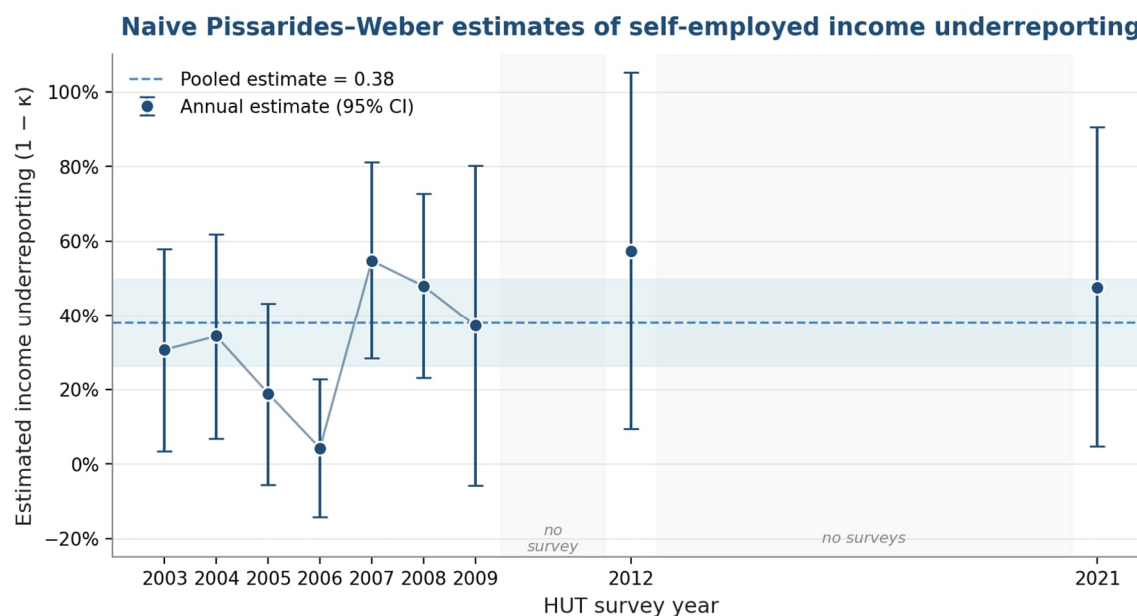
¹⁰ For a discussion of alternative definitions of self-employment in the PW context (including self-reported employment status, the share of self-employment income in total income, and restriction to households in which all adults are self-employed) see Engström and Hagen (2017), Section 5.

III. Key Findings

11. The pooled estimate across all HUT vintages indicates that self-employed households spend approximately 9 percent more on food than wage earners at the same level of reported disposable income. The estimated food income elasticity is 0.18, somewhat below the original Pissarides and Weber (1989) estimate of 0.27 for the United Kingdom but consistent with the lower elasticities typically obtained when control variables are added to the regression. Combining the two estimates yields a pooled basic underreporting rate of $1 - \kappa \approx 0.38$, with a delta-method standard error of 0.06.

12. Year-by-year estimates of the underreporting rate ($1 - \kappa$) range from approximately 4 percent in 2006 to 57 percent in 2012, with confidence intervals that are often wide. Figure 1 plots the annual estimates with 95 percent confidence intervals against the pooled estimate. Table 1 reports the underlying β and γ estimates by year together with their standard errors. The volatility across years partly reflects sampling noise (especially in the smaller vintages), partly the use of current rather than permanent income (which mechanically affects β in ways that vary with the realized income distribution), and possibly genuine year-to-year variation in self-employed compliance.

Figure 1. Basic Pissarides–Weber Estimates of Self-Employed Income Underreporting, by HUT Vintage



Source: IMF calculations based on Swedish Household Budget Survey (HUT) data provided by Skatteverket. Estimates of $1 - \kappa$ are computed as $1 - \exp(-\gamma/\beta)$ using the year-specific OLS estimates of β and γ from equation (1). Standard errors are obtained by the delta method, assuming independence of β and γ . Shaded grey bands indicate years in which no HUT survey was conducted.

Table 1. OLS Estimates by HUT Vintage and Pooled across All Years

Year	β (income elast.)	SE(β)	γ (SE intercept)	SE(γ)	Underreporting rate $1 - \kappa$ (SE)	Obs.
2003	0.304	(0.092)	0.111	(0.050)	0.31 (0.14)	2,267
2004	0.234	(0.071)	0.098	(0.040)	0.34 (0.14)	2,316
2005	0.318	(0.084)	0.067	(0.045)	0.19 (0.12)	2,063
2006	0.433	(0.050)	0.019	(0.043)	0.04 (0.09)	1,965
2007	0.173	(0.040)	0.137	(0.040)	0.55 (0.13)	2,240
2008	0.202	(0.045)	0.132	(0.039)	0.48 (0.13)	2,110
2009	0.129	(0.032)	0.060	(0.042)	0.37 (0.22)	2,045
2012	0.111	(0.029)	0.094	(0.059)	0.57 (0.24)	2,823
2021	0.145	(0.032)	0.094	(0.057)	0.48 (0.22)	2,700
Pooled	0.181	(0.016)	0.087	(0.016)	0.38 (0.06)	20,529

Source: IMF calculations based on HUT data provided by Skatteverket. Robust standard errors in parentheses. The estimating equation is $c = X\alpha + \beta y + \gamma \cdot SE + \varepsilon$, where c is log food consumption, y is log disposable income, and SE is a self-employment dummy. Controls (X) include gender of household head, age, age squared, and a home-ownership dummy. The pooled specification additionally includes year fixed effects.

13. Despite the volatility, the time series shows a moderate increase in basic underreporting between the earliest (2003–2006) and later (2007 onward) HUT vintages. The estimates for 2003–2006 cluster in a wide range from near zero to about 0.34, while those for 2007–2009 and 2012 are generally higher, in the range of 0.37–0.57. The 2021 estimate of 0.48 is broadly consistent with the post-2007 pattern, although the gap of nine years since the previous HUT vintage limits inference about the trajectory in the intervening period. The 2006 estimate (0.04) stands out at the low end and is statistically indistinguishable from zero; this reflects an unusually high estimate of β in that year (0.43, compared with a pooled value of 0.18) combined with a small γ , and is exactly the type of anomaly that a permanent income correction would be expected to dampen.

14. The relative time pattern is more robustly identified than the absolute level. The bias arising from omission of permanent income operates through the attenuation of β ; if this bias is broadly stable across vintages, it will affect the levels of the underreporting rate ($1 - \kappa$) in similar proportions across years and leave the relative ranking of vintages largely intact. This is a cautious interpretation — the

assumption of stable bias across vintages cannot be tested with the current data — but it provides a useful framework for treating the time series as informative about changes in compliance even when the absolute levels are not.

15. The estimates are based on a limited set of controls and an unrefined definition of self-employment. Robustness checks using richer specifications and alternative definitions of self-employment (for example, restricting attention to households where all adults are self-employed, or using the share of self-employment income in total income to define self-employment status) are a natural next step and are documented in Annex II. Based on the equivalent analysis in Engström and Hagen (2017), the qualitative time pattern is expected to be robust to such refinements, although the absolute levels may shift somewhat.

IV. Next Steps

16. It is recommended that Skatteverket updates these simple PW estimates in connection with each new HUT vintage and evaluates the resulting time series. Although individual year estimates should not be interpreted as point estimates of hidden income, the relative time pattern may be informative about the evolution of self-employed underreporting in Sweden. Updating the estimates with each new HUT vintage is a low-cost extension of the present exercise and provides a useful monitoring tool for tracking compliance trends over time. The estimating equation is straightforward and the data infrastructure is already in place.

17. It is recommended that work continue toward incorporating a permanent income correction along the lines of Engström and Hagen (2017). This refinement would substantially improve the credibility of the level estimates by mitigating the attenuation bias that arises from using current income in place of permanent income. The required data infrastructure — a longitudinal linkage of HUT households to administrative income registers, allowing income to be averaged over a multi-year window centered on the survey year — is in principle available in Sweden but has not yet been implemented in the current data setup at Skatteverket. This is a natural area for future technical assistance and methodological development. In the meantime, the basic estimates reported here should be interpreted as upper bounds on the true level of self-employed underreporting.

18. It is recommended that the present estimates be treated as one input among several. The PW method has known limitations, including the assumptions of equal food preferences across employment types and a constant underreporting factor, that may not hold in all circumstances. A portfolio of methods with distinct identifying assumptions is preferable to reliance on any single approach: if the flaws of different methods are sufficiently unrelated, they will tend to agree where their answers are most reliable, and disagree most where the underlying assumptions are most strained, which is itself a useful diagnostic. It is also recommended to explore the possibilities of register-based proxies of consumption and other traces of real income. Complementary approaches that Skatteverket may wish to pursue include audit-based estimates (along the lines of Kleven et al., 2011), bunching analyses around tax kink points, and discrepancies between different administrative data sources.

19. Finally, it is recommended that Skatteverket invests in building internal staff capacity to sustain and extend the analysis independently. Two or three data analysts trained in the econometric methodology and in working with the HUT data would suffice to maintain the time series and pursue the methodological refinements outlined above. Subject to funding the IMF stands ready to provide further capacity development support in the form of workshops or follow-up missions as needed.

Annex I. Technical Details of the Expenditure-Based Approach

The basic PW model. The PW approach estimates income underreporting by comparing the food Engel curves of self-employed and wage earner households. Let $c = \ln C_F$ denote log food consumption and $y = \ln Y_D$ log disposable income. The key estimating equation is:

$$c_{it} = X_{it}\alpha + \beta y_{it} + \gamma SE_{it} + \varepsilon_{it} \quad (1)$$

The parameter γ captures the upward intercept shift in the food Engel curve for the self-employed: conditional on the same level of reported disposable income, self-employed households spend γ log-points more on food. This excess consumption is attributed to income hidden from the tax authorities. Specifically, if self-employed households underreport income by a constant fraction, so that reported income equals a fraction κ of true income, then the intercept shift is $\gamma = -\beta \ln \kappa$. The compliance rate is therefore identified as $\kappa = \exp(-\gamma/\beta)$, and estimated underreporting is $1 - \kappa$. The control vector X includes gender of household head, age, age squared, and a home-ownership dummy. The pooled specification additionally includes year fixed effects.

Delta method standard errors for $1 - \kappa$. Standard errors for $1 - \kappa$ are obtained by the delta method, treating the estimates of β and γ as approximately independent. Let $g(\beta, \gamma) = \exp(-\gamma/\beta)$. The partial derivatives are $\partial g/\partial \beta = (\gamma/\beta^2) \cdot \exp(-\gamma/\beta)$ and $\partial g/\partial \gamma = -(1/\beta) \cdot \exp(-\gamma/\beta)$. Under the independence assumption, the variance of κ is approximated by $\text{Var}(\kappa) \approx (\partial g/\partial \beta)^2 \cdot \text{Var}(\beta) + (\partial g/\partial \gamma)^2 \cdot \text{Var}(\gamma)$, and the standard error of $1 - \kappa$ equals the square root of this expression.

Attenuation bias from using current income. Suppose that permanent income y^p (which is the relevant driver of food consumption under the Permanent Income Hypothesis) differs from current income y by a transitory component ω (i.e., $y = y^p + \omega$, where ω is mean zero and uncorrelated with permanent income). Substituting into equation (1) shows that the OLS estimate of β will be attenuated toward zero, because current income is a noisy proxy for permanent income. Since the estimated underreporting factor $1 - \kappa$ is a decreasing function of β , this attenuation typically produces an upward bias in the estimated level of underreporting. Engström and Hagen (2017) show, using equivalent Swedish data and a seven-year average income measure, that correcting for this attenuation reduces the estimated underreporting rate by approximately one-third. The basic estimates reported here can therefore be interpreted as approximate upper bounds on the true underreporting rate.

Key assumptions. The PW method requires the following identifying assumptions: (i) food preferences, conditional on true disposable income, are identical for self-employed and wage earner households; (ii) self-employed households underreport income by a constant factor κ across all income levels; and (iii) wage earners (WE) report their true incomes (or at least do not systematically underreport). The first assumption is intrinsically difficult to test. Engström et al (2023)¹¹ test this assumption by comparing consumption proxies among households defined as WE in year t but which transition to SE in the near future, to households that are stable WE. The presumption is that the future SE resemble the current SE in terms of consumption preferences but do not yet have the possibility to hide income. The second assumption is partially testable by examining whether the slopes of the two Engel curves are statistically equal; this equality is not rejected in the Engström and Hagen (2017) data. In Sweden, the third assumption is strongly

¹¹ Engström, P., Hagen, J., and Johansson, E. (2023). Estimating tax noncompliance among the self-employed—evidence from pleasure boat registers. *Small Business Economics*, 61, 1747–1771.

supported by the third-party reporting system covering wage income: Kleven et al. (2011) document that third-party reporting essentially eliminates underreporting among wage earners in the Nordic countries.

Table A1. Descriptive Statistics for the Estimation Sample, by Self-Employment Status (Pooled across HUT Vintages)

	Self-Employed (N = 1,783)		Wage Earners (N = 18,602)	
Variable	Mean	SD	Mean	SD
Log food consumption	11.323	0.690	11.146	0.735
Log disposable income	12.853	0.755	12.744	0.718
Share female household head*	0.274	0.446	0.330	0.470
Age of household head	48.93	12.48	49.25	15.29
Single-family house (= 1 if yes)	0.731	0.443	0.574	0.494

Source: IMF calculations based on HUT data provided by Skatteverket.

Annex II. Materials Left with Skatteverket

- Stata code for constructing the estimation sample from the HUT consumption survey, including variable definitions and sample restrictions.
- Stata code for estimating equation (1) by HUT vintage and in the pooled specification, and for computing $1 - \kappa$ and its delta-method standard error.
- Stata code for the recommended robustness checks: alternative definitions of self-employment (self-reported employment status; share of self-employment income in total income; restriction to households in which all adults are self-employed) and alternative control specifications.
- Detailed documentation describing the variable definitions, data sources, and sample construction.
- Stata code (provided in the mission) for constructing the 'stable households' sample required for the permanent income correction along the lines of Engström and Hagen (2017), to be applied once the necessary register linkages are in place.
- This Technical Assistance Report.