Are Household Surveys Like Tax Forms?
Evidence from Income Underreporting of the Self-Employed*

October 2012

Erik Hurst
University of Chicago

Geng Li
Board of Governors of the Federal Reserve System

Benjamin Pugsley
Federal Reserve Bank of New York

Abstract

There is a large literature showing that the self-employed underreport their income to tax authorities. In this paper, we quantify the extent to which the self-employed also systematically underreport their income in U.S. household surveys. To do so, we use the Engel curve describing the relationship between income and expenditures of wage and salary workers to infer the actual income, and thus the reporting gap, of the self-employed based on their reported expenditures. We find that, on average, the self-employed underreport their income by about 25 percent. We show that failing to account for such income underreporting leads to biased conclusions in a variety of settings.

1.   **Introduction**

Evidence shows that individuals systematically underreport income to tax authorities.[1] A separate literature finds that participants in experiments distort their behavior as a reaction to being studied.[2] Even administrative data such as the Vital Statistics data are not immune to problems of misreporting.[3] Collectively, previous research has shown that individuals tend to misreport their actual behavior to data collectors and administrative agencies when the incentives to do so are sufficiently large (e.g., to avoid tax payments) and/or the cost of doing so is small (e.g., changing their behavior in experimental settings). However, an implicit assumption made in the majority of empirical work using household survey data is that the data within the household surveys are immune to such systematic errors. In doing so, researchers are assuming that the problems that plague tax data, experimental data, and other types of administrative data do not plague household survey data. In this paper, we assess whether such an assumption is valid. Specifically, we address whether the self-employed, who have been shown to have misreported their income to tax authorities, have also misreported their income to household surveys.

A natural question is why interviewees would misreport their earnings to household surveys. On the one hand, it appears there is little to gain from misreporting because unlike reports to tax authorities, misreporting income to household surveys cannot decrease a household's tax liabilities. On the other hand, there is little to lose from misreporting income to household surveys. Additionally, the self-employed may perceive other indirect net benefits from underreporting income to surveys. For example, unlike wage and salary employees who receive W-2 forms, the self-employed have to expend efforts to accurately account for their true income. If the self-employed have already supplied (or have an intention to supply) a distorted income report to tax authorities, it may be easier to reuse this report for surveys

---

[1] See, for example, Clotfelter (1983), Slemrod (1985), Feinstein (1991), Andreoni et al. (1998), Slemrod (2007), and Feldman and Slemrod (2007).

[2] This phenomenon is sometimes referred to as the Hawthorne Effect, named for a series of studies at the Hawthorne Works factory where workers' productivity initially improved while under observation but declined soon after. There is a large literature within the social sciences on the Hawthorne Effect. See Levitt and List (2009) for a recent discussion.

[3] For example, Blank et al. (2009) compare administrative data from Vital Statistic records to data from the U.S. Census to show that in states where minimum age of marriage laws were binding, younger individuals appeared to have lied about their age to government officials when applying for their marriage license.

instead of computing a separate and more accurate measure of income. Households may also feel compelled to provide consistent measures between their income reported to tax authorities and household surveys if they believe that the information provided to surveys may not be completely confidential. Even a small probability of self-incrimination may lead to distorting their survey responses.

Our goals in this paper are twofold. First, we infer the extent to which the self-employed underreport their income within U.S. household surveys. As far as we know, this is the first paper that attempts to do so. Our empirical strategy is similar in spirit to the one set forth in Pissarides and Weber (1989). This procedure estimates the relationship between expenditures and income for wage and salary workers and uses the estimated coefficients from this relationship to predict the true income of the self-employed based on their reported level of expenditures. Using data from the Consumer Expenditure Survey and the Panel Study of Income Dynamics, we find that, on average, the self-employed underreport their income to household surveys by about 25 percent. The estimated magnitudes are nearly identical across both surveys for similar specifications.

Our estimating procedure makes a few key assumptions. First, it assumes no differential underreporting of expenditures by the self-employed relative to wage and salary workers. Second, it assumes that there are no definitional differences in household surveys between wage and salary workers' income and the self-employed's income. Finally, it assumes that the underlying relationship between income and expenditures, absent any underreporting, is similar between wage and salary workers and the self-employed. We test for the validity of all of these assumptions, as well as provide a variety of additional robustness analyses, in subsequent sections.

The second goal of the paper is to show several examples of how ignoring the underreporting of income by the self-employed can bias many different types of empirical analyses. For example, given that self-employment propensities differ over the lifecycle and across space, measures of lifecycle and spatial differences in earnings are sensitive to the extent of underreporting of income by the self-employed. For example, we find that roughly between 10 and 15 percent of the decline in earnings between the ages of 45 and 65 in both the Consumer Expenditure Survey and the Panel Study of Income Dynamics can be

2

explained by the underreporting of income by the self-employed given that self-employment propensities rise with age. We also show that the underreporting of income by the self-employed can alter standard estimates about (1) the importance of precautionary savings in explaining aggregate wealth holdings, (2) wealth differentials between the self-employed and wage salary workers, and (3) differences in earnings across cities.

Our work in this paper complements two different literatures. First, there is an important existing literature that looks at reporting errors within household surveys.[4] Our work adds to this literature by identifying another source of reporting errors. We find that a large group of individuals, the self-employed, *systematically* underreport their income to household surveys. Additionally, we show that the same misreporting issues that appear in non-survey settings may also exist in household survey responses. Second, our work also complements a large existing literature that estimates the extent of underreporting of income by the self-employed using tax records. Much of this work uses stratified random samples of tax returns subjected to a thorough audit. A separate strand of research uses data from actual tax returns (as opposed to random audits) to assess the amount of underreporting of income.[5] Our study differs from these in that we analyze the underreporting of income by the self-employed to household surveys as opposed to tax authorities.[6]

On balance, our work shows that it is naive for researchers to assume that individuals distorting the truth in some contexts would always provide accurate responses when participating in household surveys. While the benefits of providing distorted information to household surveys are small, so are the costs of providing inaccurate information. Such potential biases need to be accounted for when analyzing

---

[4] See, for example, Bound et al. (1994), Deaton (1997), Fitzgerald et al. (1998), Haider and Solon (2006), Meyer et al. (2009), and Aguiar and Bils (2011). Bound et al. (2001) survey nearly fifty years of research examining the extent to which household survey data are measured with error.

[5] For a summary of audit literature, see the recent surveys by Andreoni et al. (1998) and Slemrod (2007). Recent papers in the latter vein of the literature include Feldman and Slemrod (2007), LaLumia (2009), and Saez (2010).

[6] Several authors have examined the underreporting of income by the self-employed within household surveys from other countries. For example, Pissarides and Weber (1989) use their expenditure-based method to detect the underreporting of income by the self-employed within Britain's 1982 Family Expenditure Survey. They find that the self-employed underreported their income by roughly 35 percent within Britain in 1982. Using a similar methodology, Johansson (2005) finds that the self-employed underreport their income by 16-40 percent in Finland. Schuetze (2002) finds 17 percent underreporting in Canada, and Gibson, Kim, and Chung (2009) find roughly 40 and 50 percent underreporting in Korea and Russian, respectively. Engström and Holmund (2009) find that the self-employed underreport their income by about 30 percent within the Swedish Household Budget Survey.

data from household surveys.

2.   Data Description: CE and PSID

We use two nationally representative household surveys, the Consumer Expenditure Survey (CE) and the Panel Study of Income Dynamics (PSID), for the majority of our empirical analysis. In this section, we describe both surveys and discuss our sample selection criteria.

2A.   The CE Sample

We start with the pooled NBER CE extracts spanning all waves from 1980 to 2003.[7] We restrict the sample to households reporting expenditures in all four quarters surveyed and sum their quarterly responses for an annual expenditure measure. We further restrict the sample to include only households that have a male head between the ages of 25 and 55 (inclusive), who is working at least 30 hours in an average week, who has worked at least 40 weeks during the previous year.[8] We exclude any households where the male head is a wage and salary worker but the spouse (if present) was self-employed. We also exclude any household reporting farm income. Finally, we drop households with zero or negative reported household income or zero reported household expenditures. All workers in the CE data are asked to classify their primary employment type as either working for someone else in the private sector, working for the government, or being self-employed. We refer to households in the first two groups as being wage and salary workers while we classify the latter responses as being self-employed. Thus, our final sample includes 27,219 households, of which 2,508 are counted as self-employed.

We define three measures of expenditures and three measures of income. The expenditure measures are: total food expenditures, total nondurable expenditures, and total expenditures. Total food expenditures include expenditures on both food consumed at home and away from home. Following Aguiar and Hurst (2009), we define nondurable expenditures as spending on food, alcoholic beverages, tobacco, clothing and personal care, housing services, utilities, domestic care services, nondurable

---

[7] See http://www.nber.org/data/ces_cbo.html for details on the NBER CE extracts.
[8] Male heads are indentified as those male individuals who self-report themselves as being a head or, to be consistent with the PSID, those males who are married to someone who reports herself as being a head.

transportation, nondurable entertainment, and "other" nondurable expenditures.   Our total expenditure measure is the CE's measure of total household outlays including spending on nondurables, durables, education, health care, and other household outlays.

Throughout this paper we use three different measures of household income. Our first measure of household income sums all labor earnings from wage and salary employment plus all earnings generated by one's business.[9] Business earnings include both the returns to labor and the returns to capital for the male heads and their spouses (if a spouse was present). We refer to this income measure as "labor plus business income." Our second measure of household income is total family income, which includes all earnings, asset income, and transfer income received by the household. Our third measure of income is after-tax (disposable) total family income. The CEX records all taxes paid (net of any refunds) by the household during the prior 12 months. These taxes include federal income taxes, state income taxes, payroll taxes, and property taxes. To compute after-tax total family income, we simply subtract our measure of net taxes paid by the household from the measure of household total family income.

2B.   The PSID Sample

Compared with the CE, the PSID only collects limited information on household expenditures. Over the sample period, the only expenditure category measured consistently is food expenditures.   The PSID measures of household income are similar to our CE measures.[10]

We use the PSID data from 1980 to 1997 except for the 1988 and 1989 waves, during which food expenditures were not collected.   After 1997, the PSID started collecting data biennially, making us

---

[9] CE respondents are asked: "During the last 12 months, did you receive any money in wages or salary? Include all bonuses and overtime pay, commissions, tips, allowances, Armed Forces pay, severance pay, teaching fellowships, etc."   If the respondent answered in the affirmative to the above question, they were then asked: "During the last 12 months, how much did you receive in wages and salaries for all jobs before any deductions?"   This question was asked of all respondents regardless of their self-employment status. Respondents were then asked the following question about nonfarm business income: "During the last 12 months, did you have any income or loss from your nonfarm business, partnership, or professional practice?"   Respondents that answered in the affirmative to this question were asked "What was the amount of income or loss after expenses?" It is possible that the CE may understate the total income of the business owners if respondents do not report any retained earnings. We explore this concern in much greater depth in Section 5.

[10] The only difference is that PSID only includes reported measures of federal income taxes paid by the household up through survey year 1991. There are no measures of actual state income taxes or payroll taxes paid by the household in any survey year. For years up through survey year 1991, our measure of net federal taxes paid by the households is the actual reported taxes paid by the household. For survey years after 1991, our measure of net federal taxes paid by the household is our imputation of taxes paid by the household using NBER TAXSIM.[10]

unable to match income and expenditure occurred during the same year.[11]

With respect to income, business owners in the PSID were first asked: "Did the business show a profit, a loss, or did it break even in the prior calendar year?" For those reporting a profit, the question was followed up with "How much was your share of the total income from your business in [the prior calendar year]—that is, the amount you took out plus any profit left in?" A separate question was asked about the amount of business losses if the individual reported having business losses. Thus, the PSID respondents were specifically asked to consider the amount of income earned from the business that was reinvested back into the business.

Within the PSID, individuals are asked to report whether on their main job they "are self-employed" or "employed by someone else." Individuals can report being only employed by someone else, only self-employed, or both employed by someone else and self-employed. We define self-employed households as being a household where the male head reports being self-employed only. Wage and salary households are defined as ones where the male head reports only working for someone else. We exclude households who report being both self-employed and working for someone else from our analysis. Lastly, we exclude all households where the head is a wage and salary worker but the spouse (if present) was self-employed. None of our results are sensitive to whether these households are included or excluded from the analysis.

We form two samples using the PSID data. The selection criteria of the first sample, which exploits only the cross-sectional nature of the PSID data, are nearly identical to the CE sample discussed above. After applying these restrictions, our first PSID sample (PSID 1-Year sample) has 36,434 households, 4,446 of which are counted as self-employed. Our second PSID sample leverages the panel dimension of the data. We combine multiple waves of the PSID data to create a three-year average measure of income. This approach has been taken by many others in the literature to construct measures of permanent income within the PSID (see, for example, Solon 1992 and Gottschalk and Moffitt 1994).

---

[11] Households are asked to report income earned during the previous calendar year and to report current weekly or monthly food expenditures. We match the income report from survey year $t+1$ with food expenditures in survey year t, which is the standard approach in the literature.

Specifically, for each household in year *t*, we compute income measures that average income between *t-1* and *t+1*.[12] This sample imposes two additional restrictions above and beyond the restrictions to our one-year PSID sample. First, we impose that the household be in the survey for all consecutive years between *t-1* and *t+1*. Second, we impose that households who are classified as self-employed are self-employed in *t-1*, *t*, and *t+1* while households who are classified as wage and salary workers are wage and salary workers consistently in all three years. We exclude any household who transitioned from self-employment to wage and salary workers during the three-year period (or vice versa) from our sample. Additionally, when restricting the sample to include only those individuals with positive income, we make the restriction based on the three-year average of income.[13] These restrictions left us with 18,233 observations, 1,901 of which were self-employed.

2C.   Differences between Wage Workers and the Self-Employed

Table 1 shows descriptive statistics separately for wage and salary workers and the self-employed within the CE sample, one-year PSID sample, and three-year average PSID sample.[14] Across all surveys, the self-employed are slightly older, more likely to be married, much less likely to be black and are somewhat better educated. The two surveys are largely consistent in their demographic composition. It is not surprising that there are some moderate demographic differences between the PSID three-year average sample and the PSID one-year sample given that the former conditions on the household being in either full time employment or full time self-employment for three consecutive years. If lower educated households are more likely to transition between self-employment and wage work or if they are more likely to transition in and out of the labor force, then the three-year average PSID sample would appear more educated than the one-year PSID sample.

---

[12] We use this alternative measure of permanent income to evaluate our choice of education dummies as an instrument for permanent income in the one-year sample. As shown below, our estimates of underreporting using the two procedures to measure permanent income are very similar.

[13] The number of households we exclude was essentially zero when using the three-year average PSID sample. This is not surprising given that we are averaging income over three years for households with full time working heads.

[14] We convert all income and expenditure measures to year 2000 dollars using the CPI and express them in log levels. All estimates are weighted using the sampling weights of the respective surveys.

Before proceeding, we would like to highlight that, unconditionally, the self-employed report working roughly 10 to 12 percent more annual hours than wage and salary workers. The difference remains essentially unchanged even controlling for demographics. The difference in hours worked is notable and we will return to it when we address whether the expenditure differences we document conditional on income could be driven by varying tastes for leisure or difference in home production needs for the self-employed.

3.   Detecting the Underreporting of Income of the Self-Employed

Our central identifying assumption in understanding the relationship between income and expenditure is the log-linear Engel curve. Specifically, within each group, k, where k indicates either self-employed ($k = S$) or wage and salary ($k = W$), we assume that household i has a preference that generates, at least to a first order approximation, an Engel curve of the following form:

$$\ln c_{ikt} = \alpha_k + \beta_k \ln y_{ikt}^p + \Theta_k ' X_{ikt} + \mu_{kt} + \varepsilon_{ikt} \tag{1}$$

where β_k is the income elasticity, X_{ikt} is a vector of demographic controls, μ_{kt} is a time effect, and ε_{ikt} represents the cumulative effects of other unobserved determinants of the household's consumption. The vector of household controls includes: a series of five-year age dummies, a dummy if the household head is black, a dummy if the household head is married, and a series of family size dummies. We assume that households are able to borrow and lend in asset markets so that it is the permanent component of income, y_{ikt}^p, that determines the household's consumption, as is standard in modern theories of consumption.[15]

Estimating (1) directly is problematic because the household's annual income reported to the CE and PSID represents both y_{ikt}^p and a transitory component. Here we assume for the moment that all households accurately report their current income, y_{ikt}, so that:

$$\ln y_{ikt} = \ln y_{it}^p + \Omega_k ' X_{ikt} + \nu_{ikt} , \tag{2}$$

---

[15] We do not require that markets are complete, only that the household is able to save and borrow in some form sufficient to support a lifetime budget constraint in expected value. In section 5, we explore whether differential liquidity constraints between the two groups can explain the patterns we see in the data.

where $\Omega_k ' X_{ikt}$ and v_{ikt} represent respectively the predictable and unpredictable components of transitory income. The transitory fluctuations could represent both the true transitory variation in household income and classical measurement error in the report of income. We assume that the unpredictable component v_{ikt} is orthogonal to both the controls X_{ikt} and any unobserved determinants of the household's consumption ε_{ikt} (i.e., $E[X_{ikt}v_{ikt}]=0$ and $E[v_{ikt}\varepsilon_{ikt}]=0$). Rewriting (1) in terms of (2) implies:

$$\ln c_{ikt} = \alpha_k + \beta_k \ln y_{ikt} + \Psi_k ' X_{ikt} + \xi_{ikt} , \qquad (3)$$

where $\Psi_k = \Theta_k - \beta_k \Omega_k$ and $\xi_{ikt} = \varepsilon_{ikt} - \beta_k v_{ikt}$. Even if there was no measurement error in household income reports, transitory income fluctuations introduce attenuation bias in our estimation of β_k from (3) since, by construction, $E[\ln y_{ikt}\xi_{ikt}] \neq 0$.[16]

To mitigate the effects of both measurement error and transitory income fluctuations in our estimation of equation (3), we follow two separate approaches. First, when using the CE sample or the one-year PSID sample, we instrument for y_{ikt} using four educational attainment dummies (exactly high school, some college, exactly college, and more than college), assuming that educational attainments only affect consumption through affecting permanent income. Second, when using the PSID three-year average sample, we average the income measures over multiple years so that the transitory fluctuations are mitigated. Moreover, by comparing the two procedures for the same sample, we can examine the extent to which education dummies satisfy the exclusion restriction. As seen below, our estimates are very similar under both procedures.

We have assumed that the log-linear Engel curve is a good representation of household behavior as well as that the Engel curves between the two groups differ primarily due to differences in intercepts as opposed to differences in slopes. Figure 1 plots a nonparametric estimate of the income-expenditure Engel curves estimated separately for the two groups using the PSID three-year average sample where our

---

[16] Additionally, Andreoni (1992) shows that individuals can use income underreporting to the IRS as a way to smooth transitory fluctuations in income if they do not have access to other sources of credit. For this reason, underreporting may be correlated with transitory fluctuations in income. This suggests another reason to use measures of permanent income in our Engel curve estimation of income underreporting.

measure of income is log average total family income and our measure of expenditure is log average food spending. We first condition out our demographic controls by separately regressing log average total family income and log average food expenditure on our X vector of controls. We then use a nonparametric median spline procedure to estimate a nonlinear Engel curve from the residuals for each group. Figure 1 plots the estimated nonlinear Engel curve for each group.[17]

There are two interesting facts from Figure 1. First, the estimated Engel curve for the self-employed (dotted line) lies strictly above the estimated Engel curve for the wage and salary workers (solid line) over essentially the entire support of the data. In other words, for a given amount of income, the self-employed within the PSID spend more on food than the wage and salary workers. This is consistent with a potential story of the underreporting of income by the self-employed to household surveys. Second, even though Figure 1 was estimated nonparametrically, the slope of the Engle curve for the wage and salary workers was roughly constant as income increased. Moreover, the estimated slope of the Engel curve for the self-employed was roughly similar to the estimated slope for the wage and salary workers. When we estimate the log-linear Engel curve as specified in (3) separately for each group on this exact sample, the estimated slope of the log-linear Engel curve for the sample of the self-employed was 0.29 (standard error = 0.02) while the estimated slope of the Engel curve for the sample of the wage and salary workers was 0.34 (standard error = 0.01). Although the differences in the slopes are statistically significant (p-value = 0.02), the economic magnitude of the difference is quite small.

If income reports in the surveys are accurate, Figure 1 shows that the self-employed consume substantially more than wage and salary workers. We use a simple model that relaxes the assumption of accurate reporting of income and allows us to estimate the magnitude of underreporting by the self-employed. For households of each type, we now assume their income reports to the household surveys are determined by

---

[17] For expositional purposes, we re-centered the residuals at the unconditional log average income and log average expenditure for each group. Finally, when plotting the data, we restrict the range of the log average income residuals to be within the overlapping range for both groups. Within the overlapping range, we further restrict the support by truncating observations with log average income below the 1st percentile and above the 99th percentile.

$$\ln y_{iWt} = \ln y_{iWt}^p + \Omega_W 'X_{iWt} + v_{iWt} \tag{4}$$

$$\ln y_{iSt} = \ln \kappa_S + \ln y_{iSt}^p + \Omega_S 'X_{iSt} + v_{iSt} \tag{5}$$

For each group, we assume that the observed income in any time period, y_{ikt}, is a noisy proxy of

permanent income, where v_{ikt} includes both any classical measurement error in the income report and

the unpredictable component of transient variation around permanent income. As before, we assume that

v has mean zero and is orthogonal to the unobserved determinants of log expenditure, ε, in (1).

Equations (4) and (5) embed two additional assumptions. First, we assume that self-employed

households ($k = S$) systematically misreport their earnings by a factor, κ_S. Although we do not impose

that $\kappa_S \leq 1$, our empirical estimates below reveal that, in fact, $\kappa_S \leq 1$. We assume that κ_S is constant

across self-employed households. Second, we assume that the wage and salary workers provide a

systematically unbiased report of their income to household surveys, i.e., $\kappa_W = 1$. If wage and salary

households also systematically misreport their income to household surveys, κ_S would be an estimate of

the differential systematic underreporting by the self-employed.

The goal of this section is to provide a method to estimate κ_S. In order to do so, we also assume

that the parameters of the two Engel curves are constant between the self-employed and the wage and

salary workers. Lastly, we assume that household expenditures are not differentially misreported between

the two groups. This amounts to assuming that if the self-employed under or over report their

expenditures systematically, the wage and salary workers under or over report their expenditures

systematically to the same extent. We provide evidence for this assumption in Section 5.

Given these assumptions, we can estimate κ_S via the following expression formed from

combining (3) with (4) and (5):

$$\ln c_{ikt} = \alpha + \beta \ln y_{ikt} + \gamma D_{it} + \Psi X_{ikt} + \mu_t + \xi_{ikt} , \tag{6}$$

where D is a dummy variable indicating whether the household head is self-employed. As before

11

$\Psi = \Theta - \beta'\Omega$ and the unobserved determinants are expressed as $\xi_{ikt} = \varepsilon_{ikt} - \beta v_{ikt}$. The fraction of income reported by the self-employed, κ_S, is identified as $\exp(-\gamma/\beta)$, and we form the estimate $\hat{\kappa}_S$ as $\exp(-\hat{\gamma}/\hat{\beta})$ using the estimated coefficients from equation (6).[18] We will often express our results in terms of the amount that the self-employed underreport their income to household surveys, which can be defined as $1-\kappa_S$. When estimating (6) using the CE sample or the PSID one-year sample, we instrument for reported $\ln y_{ikt}$ using the set of education dummies. Estimating (6) with OLS would underestimate the income elasticity of consumption β and hence overestimate κ_S. When estimating (6) using the PSID three-year average sample, we use three-year average income when computing our measure of $\ln y_{ikt}$. Given this, we only report the OLS estimates for the PSID sample.

Before proceeding, it is worth discussing the potential of measurement error in our expenditure data. Given the fact that we are estimating (6), classical measurement error in our expenditure data will not bias our estimates of $1-\kappa_S$. There is evidence, however, that measurement error in the expenditure measures within the CE is large and that it may not be classical. For example, Bee et al. (2011) and Sabelhaus et al. (2011) show that CE respondents systematically underreport their expenditures to the survey and that the underreporting has grown over time. Additionally, there is evidence that the extent of the underreporting differs by expenditure category. Categories like food at home tend to be measured well within the survey while categories like clothing are not. Despite this, our estimates of $1-\kappa_S$ using the CE data will be consistent if (1) the systematic measurement error in expenditure reports does not vary between the self-employed and wage and salary workers and (2) the systematic measurement error in expenditure reports is uncorrelated with permanent income and other controls.

---

[18] In the online robustness appendix that accompanies this paper, we compare our estimation to Pissarides and Weber (1989). The main difference between the procedures is that Pissarides and Weber take a more parametric approach towards the income process that adjusts for the greater volatility of transient income fluctuations for the self-employed. Ignoring the differences in variance between groups could overestimate the extent of underreporting. In the appendix we show that adjusting for the differences in variance using the Pissarides and Weber model decreases the estimated underreporting only slightly. Moreover, their procedure only adjusts for differences in transient income volatility. If self-employed permanent income is also more volatile than for employees, this attenuates the bias. After also correcting for differences in permanent income volatility, we find estimates close to those in Table 3. The online robustness appendix can be found at http://faculty.chicagobooth.edu/erik.hurst/research/.

It is true that systematic biases in expenditure reports will affect the intercept of the estimated

Engel curves for both the self-employed and wage and salary workers separately. However, if the

systematic bias is similar between the groups, the difference in the Engel curves (our estimate of γ) will

not be altered. There is no obvious reason to believe that the systematic measurement error documented

in the CE differs between the self-employed and wage and salary workers. Moreover, in Section 5, we

test directly for differences in systematic reporting of expenditures between the self-employed and wage

and salary workers and reject any systematic difference. However, if the measurement error in

expenditure reports is correlated with permanent income, our estimates of β will be downward biased,

causing us to understate the amount of income underreporting by the self-employed. Aguiar and Bils

(2011) provide some evidence that the extent of income underreporting in the CE has increased to a

greater extent for high income households relative to low income households. Despite this, we believe

that our estimates are unbiased for two reasons. First, according to Aguiar and Bils (2011), the estimated

income elasticity for food in the CE (something akin to our estimate of β) has been relatively stable over

time, suggesting that the changing differential measurement error in total expenditures between rich and

poor households has not altered the estimates of the food income elasticity. This is consistent with food

being relatively better measured in the CE than other expenditure categories. Second, we are finding

similar estimates using PSID expenditure data which has not been shown to have the same systematic

measurement error problems.

4.   Estimating the Missing Income Using Expenditures

In this section, we show the baseline results from estimating (6) for our three samples—CE, PSID 1-year,

and PSID 3-year. Table 2 shows our estimates of both β and γ from equation (6) using different

measures of expenditures separately for the CE and PSID samples. For comparison, we show both the

OLS and IV results when using the CE and PSID one-year sample. Also, we separately show the results

where we use total family income and total after-tax family income.[19] As seen from Table 2, across all

---

[19] After-tax income, although subject to its own measurement issues, is our preferred measure of income. If the self-employed
pay less tax (whether through underreporting or other means), money that would have been paid in tax can instead be allocated to

expenditure measures and across both surveys, estimates of γ are positive and statistically different from zero (shown in panel A). For example, using the CE sample and using the nondurable expenditure measure, we estimate that the self-employed spend 18 percent more than wage and salary workers with similar reported income. The estimates of γ are roughly similar whether or not we instrument for annual income. However, the estimates of γ are slightly smaller when we use after-tax family income as our measure of income (e.g., 14 percent for nondurable expenditures in the CE). Finally, the estimates of γ for food expenditures are roughly similar between the PSID and CE samples.

The estimates of β from (6) are shown in Panel B of Table 2. Not surprisingly, given the existence of transitory variation around permanent income, in addition to measurement error in annual income, the estimates of β are sensitive to whether we estimate (6) via OLS or IV. The fact that current reported income is a noisy proxy of permanent income attenuates the estimated expenditure income elasticities. In all cases using the PSID one-year sample and the CE sample, the IV estimates of β are higher than the OLS estimates of β. The IV estimates of the income elasticities for food expenditures range from 0.31 to 0.42 across the various samples and specifications. The IV estimates of the income elasticities for total nondurables and total expenditures are much higher. These results are consistent with many empirical studies which find that food is a relative necessity compared to other consumption categories. It is also comforting that the food income elasticity estimated via OLS using the three-year average PSID sample is nearly identical to the food income elasticity estimated via IV using the one-year average PSID sample and after-tax income. This allows us to provide some estimates of underreporting using the PSID three-year average sample that does not rely on the validity of our instruments.

Table 3 presents the central results of the paper, quantifying the self-employed's missing income. In this table, we use the estimates of β and γ shown in Table 2 to construct estimates of $(1-\kappa_s)$ - the estimated underreporting of income by the self-employed relative to wage and salary workers. We

---

consumption. Even if income is accurately reported to household surveys, the self- employed's expenditures would be higher for a given level of before-tax income. Using after-tax income is robust to this source of excess expenditures.

estimate this reporting gap as $(1-\hat{\kappa}_S)=1-\exp\left(\dfrac{-\hat{\gamma}}{\hat{\beta}}\right)$ and compute standard errors using standard

asymptotic approximations. In this table, we only show the results from the IV regressions for the CE

and PSID one-year samples, which are the appropriate way to estimate (6) given the measurement error in

income. When reporting the results using the PSID three-year average sample, we compute the extent of

underreporting using the OLS results.

The results of Table 3 are striking. The estimated amount of underreporting is similar across the

two surveys and are very similar regardless of the measures of income and expenditures used in the

estimation. For example, the estimated underreporting using total family income and food expenditures

are 31.7 percent in the CE, 31.5 percent in the PSID one-year sample, and 29.4 in the PSID three-year

sample. Using after-tax total family money and food expenditures yields estimates of 25.2, 28.8, and 27.8

in the three samples. respectively. The estimated underreporting is slightly smaller using broader

measures of consumption. For example, the estimated underreporting is roughly 25 percent using both

nondurable expenditures and total expenditure when the income measure is total family income. With

after-tax total family income, the estimates fall to below 20 percent. All results are statistically

significant at the one percent level. In the next section, we show that the underreporting of vehicle

expenditures by the self-employed (which are included in both nondurable expenditures and total

expenditures) could be the reason that our income underreporting estimates are lower when we use more

aggregate expenditure measures as opposed to the parameters we estimated we use food expenditures.

As discussed above, within the PSID, respondents were asked only to recount their actual federal

taxes paid up through 1991. After that, we have imputed the federal taxes paid using the NBER Taxsim

data. Is the imputation potentially biasing our results with respect to the importance of conditioning on

after-tax income? It does not appear so. We infer this from the fact that if we compare the results

between using pre-tax and after-tax income during the pre-1991 sample, they are similar to the results for

the full sample. For example, using the PSID three-year sample and restricting the sample to pre-1991

data (where no imputation was necessary), the estimated underreporting using pre-tax total family income

and after-tax total family income was 28.1 percent and 27.5 percent, respectively.   However, the gap between the two estimates in the pre-1991 period is roughly similar to the gap between the pre- and after-tax estimates in the full sample.

The results in this section suggest that household surveys are like tax forms in that households with an incentive to misreport their income to the tax authorities also misreport their income to household surveys. The results also suggest more broadly that household surveys are not immune to the potential problems of mis-measurement that plague other types of data.   Our estimates suggest that the self-employed under-report their income by roughly 20-30 percent.

5.   Examining the Validity of Key Identification Assumptions

In this section, we assess the plausibility of our identifying assumptions, placing particular emphasis on underreporting as the first order source of differences in predicted expenditure. First, our model imposes, conditional on our controls, a log-linear Engel curve. Previous studies, including Pissarides and Weber (1989), have found this to be a reasonable approximation. Besides analytic convenience, we consider several alternative models incorporating higher order terms; in each of these models we fail to find consistent evidence that the higher order terms are either significant or quantitatively important.   Despite potential nonlinearities for specific expenditure categories such as alcohol or clothing, we find the log-linear form fits the data well for food, nondurable, and total expenditures. This is consistent with our findings in Figure 1 where we estimated the Engel curve for wage and salary workers nonparametrically. Second, our baseline results constrain the effect of controls and the income elasticity to be identical across groups. In all samples, we fail to reject equality of the covariate coefficient vector. We do find some weak evidence that the income elasticity of the self-employed is lower than wage and salary households. However, the difference is inconsistent across samples and small in magnitude.

Given our procedure, our estimate of underreporting remains the residual explanation for the difference in predicted expenditure across groups conditional on income. There are other potential explanations for this expenditure gap even if income were accurately reported. In the remainder of this

section we address several alternative explanations and show that our estimates holds when taking these alternative explanations into account.

5A.   Potential Systematic Differences in the Reporting of Expenditures

The fact that some business expenditures can substitute for personal expenditures may undermine the validity of the assumption of no differential misreporting of household expenditures between the two groups. For example, the self-employed sometimes purchase vehicles for their business. In doing so, they may treat gasoline and maintenance associated with the vehicle as business expenses. It is possible that some of the expenses they attribute to the business actually should instead be recorded as household expenditures. The story could also hold in reverse. The self-employed could conceivably report all outlays on vehicle expenses as being personal consumption even though some of the expenses are for their business. Given this, the self-employed could be underreporting or over-reporting their personal expenditures to household surveys.[20]

To test for the importance of this, we perform two separate exercises. First, using the CE sample, we re-estimate the amount of underreporting of income using a variety of expenditure subcategories in order to explore the robustness of the results across the subcategories. Then, we can verify whether expenditure subcategories that are not likely to be subject to the confounding of business and personal use yield different estimated underreporting gaps than categories where the potential substitution for business and personal use is larger. For example, while vehicle expenditures may be easily attributed to the business even if they were for personal use, it is less likely that an individual would attribute the spending on home utilities as being a business expense. To implement this procedure, we re-estimate (6) using the log of spending on different expenditure sub-categories as the dependent variable using the CE data. Focusing on utilities as a measure of expenditures and using after-tax family income as our measure of income yields an underreporting of income by the self-employed of roughly 30 percent, nearly identical to

---

[20] For some categories of expenditure, such as gasoline, the CE explicitly asks the share of that expenditure used for business. For some other expenditure categories, such as auto repairs and parking, the CE explicitly reminds the consumers not to include such expenditure for business use. However, it is still possible that some consumers counted personal expenditure as business expenditure or vice versa, leading to reporting inaccurate personal expenditure to the CE.

the food expenditure estimates.

Across all the subcategories we explored, nearly all of them yielded an estimated underreporting gap in the 20 to 30 percent range. One notable exception, however, is nondurable transportation expenditures (which includes spending on vehicle gasoline, car maintenance, parking fees, tolls, etc.). Using this subcategory, we estimate underreporting of income by the self-employed close to zero. This is consistent with anecdotes that the self-employed frequently classify a large amount of their vehicle expenditures associated with personal use as business expenses. Overall, however, we find no evidence that the results across any of the other expenditure categories differ in any substantive way from our main results even though the substitutability with business expenditures differs markedly across the categories. Given that nondurable vehicle spending is a component of nondurable spending and total outlays, it is not surprising that our estimates of underreporting are lower for nondurable spending and total spending than they are for all the other individual categories like food expenditures and utilities.

The results above, however, raise a different question. Is it possible that the self-employed systematically misreport their expenditures on all categories? To test for this hypothesis, we estimate the relationship between the *share of spending* on a given category and total expenditures based on Deaton and Muellbauer's (1982) Almost Ideal Demand System (AIDS) and allowing for systematic measurement error of expenditures. If underreporting were uniform across categories, then while total expenditures would be underreported, budget shares would be unaffected. Our procedure tests whether the budget shares of necessary goods and luxury goods, conditional on total expenditure, differ between the self-employed and wage and salary workers. For example, if the self-employed underreport expenditures, then their budget shares for luxury goods will be too high for their measured level of total expenditures.

To be concrete, suppose individuals report their expenditures on different consumption categories (*j*) such that:

$$\log c_{ikt}^{j} = \log c_{ikt}^{j*} + \log \omega_k + \phi_t \tag{7}$$

where c_{ikt}^{j} is the reported amount of spending on consumption category *j* by household *i* from group *k* in

18

year t and c_{ikt}^{j*} is the true amount of spending on consumption category j by household i from group k in

year t.[21] The actual report is subject to two types of multiplicative error: a group specific term that is

constant across all categories in all time periods (ω_k) that represents the fraction of true expenditures

reported, and a time specific term that is constant for all groups and for all categories (ϕ_t). We include a

time effect because it is well documented that measurement error in the CE has been growing over time.[22]

The key assumption in our procedure to estimate systematic misreporting of expenditures by the self-

employed is that there is a group specific measurement error term in reported expenditures that is constant

across both goods and time and that is additive in log expenditure.

Given this specification, reported total expenditure, c_{ikt}, is just the sum of reported expenditures

on each of the individual categories and actual total expenditures, c_{ikt}^*, is just the sum of the actual

expenditures on the individuals categories. With these assumptions, we can express total reported

expenditures as also being a function of the three components such that:

$$\log c_{ikt} = \log c_{ikt}^* + \log \omega_k + \phi_t . \qquad (8)$$

Under an AI-style demand system, individuals have expenditure shares for consumption category j that

can be expressed as:[23]

$$s_{ikt}^{j*} = \tilde{\alpha}^j + \tilde{\beta}^j \ln c_{ikt}^* + \tilde{\Theta}^j X_{ikt} + \tilde{\varepsilon}_{ikt}^j \qquad (9)$$

where $s_{ikt}^{j*} = \dfrac{c_{ikt}^{j*}}{c_{ikt}^*} = \dfrac{c_{ikt}^j}{c_{ikt}}$ and X_{ikt} is the same vector of demographic controls used above. Conditional on

demographics, the share of spending on some expenditure categories will be increasing in total

---

[21] We assume that within each year households face similar prices for the expenditure categories so income differences are driving changes in expenditure shares.

[22] See Aguiar and Bils (2011) We wish to note that we could enrich the structure of the measurement error even further by allowing for either a time-invariant category specific component or a time-varying category specific component. As long as we assume that these category specific terms do not vary with self-employment status, our estimating procedure to identify systematic misreporting of expenditures by the self-employed will not be altered.

[23] This expression omits a price term. Although there may be year-to-year price variation, since we will aggregate up to fairly broad categories where substitution effects are much smaller, omitting prices is reasonable. We also estimate the system for specific years where price variation should be minimal and find the same effects.

expenditures (i.e., luxuries), while the share of spending on other expenditure categories will be decreasing in total expenditures (i.e., necessities).   As above, we assume that the parameters of expenditure share elasticity function are constant between the self-employed and wage and salary workers.   Given the above specifications and the assumptions on measurement error, we estimate:

$$s_{ikt}^{j*} = \tilde{\alpha} + \tilde{\beta} \ln c_{ikt} + \tilde{\gamma} D_{it} + \tilde{\Theta} X_{ikt} + \tilde{\mu}_t + \tilde{\xi}_{ikt} , \qquad (10)$$

where D_{it} is an indicator variable taking the value of one if the household is self-employed in year t and $\tilde{\mu}_t$ is a vector of year dummies.   All other variables are defined above.

The coefficient $\tilde{\gamma}$ in (10) represents the extent to which the self-employed will under or over report their expenditures relative to wage and salary workers. In particular, $\tilde{\gamma}$ can be shown to be equal to $-\tilde{\beta}_j \log \omega_S$ where ω_S is the fraction of expenditures reported by the self-employed relative to wage and salary workers. If the self-employed underreport their expenditures by a constant amount on all expenditure categories relative to wage and salary workers, $\tilde{\gamma}$ will be positive for luxuries and $\tilde{\gamma}$ will be negative for necessities.   The patterns will be the opposite if the self-employed over-report their expenditures by a constant amount on all expenditure categories.   If $\tilde{\gamma}$ is equal to zero for luxuries and necessities, this says that the self-employed are not systematically under-reporting or over-reporting their expenditures relative to wage and salary workers.

To implement our estimation procedure, we segment our nondurable consumption categories into luxuries, necessities, and other.   We use the estimated consumption category expenditure elasticities from Aguiar and Bils (2011) to segment the categories.   We define luxuries to be those consumption categories with estimated expenditure elasticities greater than 1.3. These categories include food away from home, men's and women's clothing, and nondurable entertainment expenditures.   Necessities are defined to be those consumption categories with estimated expenditures less than 0.7. These categories

include food at home, children's clothing, and utilities.[24]   Nondurable expenditure categories with expenditure elasticities between 0.7 and 1.3 were considered neither luxuries nor necessities. We also exclude transportation which, given our results above, may be miscategorized as business spending by the self-employed. We estimate the demand system from (10) defined over the three categories of expenditures. Our estimate of $\tilde{\gamma}$ from the regression using the luxury share as the dependent variable is 0.005 (standard error = 0.003) while the estimate of $\tilde{\gamma}$ from the regression using the necessity share as the dependent variable is -0.003 (standard error = 0.003). An F test for $\tilde{\gamma} = 0$ for each category has a p-value of 0.25. This suggests that ω_s, the fraction of expenditures that is systematically reported by the self-employed relative to wage and salary workers, is very close to one.

Both of our additional tests suggest that, aside for the possibility of nondurable vehicle expenditures, there is no systematic evidence that the self-employed misreport their expenditure data to household surveys relative to wage and salary workers.

5B.   Potential Other Reasons That Could Cause Systematic Differences in the Income-Expenditure Relationships.

In this subsection, we assess whether differences in the income-expenditure relationships between wage and salary workers and the self-employed are driven by factors other than the systematic underreporting of income by the self-employed. In particular, we examine 1) differences in the potential conceptual measures of income between the two groups, 2) differences in desired expenditures due to differences in hours worked, and 3) differences in desired saving propensities between the two groups.

Different Measures of Income

We begin by assessing whether the above underreporting results could be driven by differences in the conceptual measures of income between the two types of individuals. As noted above, there are a few

---

[24] Aguiar and Bils (2011) estimate their expenditure elasticities from a regression of log expenditure on a given category on log total expenditure and controls as opposed to estimating the share of expenditure on a given category on log total expenditure and controls (as we do). Hence, their estimates range from 0.25 to 2.10.

potential issues with our income measures. First, the results in Tables 2 and 3 focus on pre-tax and after-tax family money. Do our results differ instead if we use before-tax business plus labor income? As seen from column (1) of Table 4, the IV results using pre-tax business plus labor income are very similar to the results shown in column (1) of Table 3. The regression we ran is identical to what we reported in Table 3 aside from changing the income measure to pre-tax business plus labor income.

Additionally, as noted above, the CE data do not allow us to measure the retained earnings of the self-employed. To the extent that some of the desired saving of the self-employed is done via retained earnings, the self-employed could have a higher propensity to consume out of the income they draw out of the business. This, however, is not a problem with the PSID data. To test for the importance of the conceptual difference in income as an explanation of our results in the CE, we examine the underreporting behavior of the self-employed with zero reported business wealth relative to the underreporting behavior of the self-employed with non-zero business wealth. The assumption is that if the self-employed are reinvesting a substantial amount of their income into the business, it should show up in the value of their business. Such an identification is made possible given the fact that a substantial fraction of the self-employed owns a business with little value (see Hurst and Lusardi 2004). This is consistent with the facts documented in Hurst and Pugsley (2011), who show that many of the self-employed are in industries where very little physical capital is needed (skilled craftsmen, real estate agents, lawyers, etc.).

To do this, we re-estimate (6) using the CE data including separate dummies for the business wealth of the self-employed. The results from this specification show that the estimated underreporting gap is still substantial even for those self-employed with zero business wealth. These findings suggest that the differences in the conceptual measures of income between the self-employed and the wage and salary workers is not the primary driver of the underreporting of income we documented above.

Difference in Expenditures Due to Differences in Hours Worked

As documented in Section 2, the self-employed do systematically work more hours than wage and salary workers, even within the restricted sample of individuals working full time. Theory suggests that

individuals who work more may engage in less home production compared to individuals who work less (Becker 1965 and Ghez and Becker 1975), or that increased work hours may raise the marginal utility of consumption. Empirically, Aguiar and Hurst (2005, 2007, and 2009) show that a given consumption bundle produced with more market inputs, as opposed to time inputs, is indeed more expensive. To the extent that the self-employed engage in less home production, because they are allocating more hours to market work, expenditures may be higher simply because they are spending more for an equivalent consumption bundle.

To control for this potential difference in the market cost of the consumption bundle, we follow a procedure similar to that of Blundell, Browning, and Meghir (1994) and further control for log work hours. We re-estimate (6) with a log work hours control using after-tax total family income as our income measure to compute new estimates of $1 - \hat{\kappa}$. These results are shown in column (2) of Table 4. For each expenditure measure, the results are very close to the original estimates shown in column (2) of Table 3.

Differences in Saving Propensities

A potentially more significant problem is the extent to which the self-employed have different saving incentives relative to wage and salary workers. If the self-employed face greater income risk, they may want to accumulate more wealth to insure themselves against potential shocks. Likewise, if the self-employed face more binding liquidity constraints, they may accumulate more wealth to overcome such constraints. Finally, if the self-employed receive fewer fringe benefits, they may have to save more for health shocks and for retirement. All of these factors will shift the income-expenditure relationship downward (as a result of the higher saving propensities), causing us to *under-estimate* the extent to which the self-employed underreport their income.

However, the self-employed may also have a higher tolerance for risk (Barsky et al., 1997). As a result, for a given amount of risk, the self-employed may have a lower desire to accumulate precautionary savings. The differences in risk preferences could result in less desired saving by the self-employed and, as a result, cause us to overstate the extent to which the self-employed underreport their income. A

23

similar result would occur if the self-employed systematically expect higher income in the future.

We can partially assess whether such possibilities affect the expenditure gap by segmenting the sample by age. High expected future income, binding liquidity constraints, and the existence of labor income risk are much more important for young households relative to older households (Gourinchas and Parker, 2002). In columns (3) and (4) of Table 4 we re-estimate (6) on a sample of young and older households, respectively. We define younger households as households between the ages of 25 and 40 and define older households as households between the ages of 40 and 55. Again, for comparison to the results above, we use after-tax total family income as our measure of income. Across expenditure measures, the estimated amount of underreporting is very similar between old and young households. These results suggest that differences in future income expectations, differences in binding liquidity constraints, and differences in precautionary motives are likely not the primary driver of our results.[25]

As an additional robustness exercise, we examined the extent of underreporting by the level of household wealth. To do this, we augmented (6) to include dummies for whether the household was in the second wealth quartile, the third wealth quartile or the top wealth quartile along with our entrepreneurship dummy interacted separately with all of the wealth quartiles. The goals of this specification are twofold. First, we ask the general question as to whether differences in household wealth (which could result from past savings behavior) could explain the estimated underreporting by the self-employed. Second, the specification allows us to assess whether the extent to which households underreport differs by wealth quartiles. The results from this specification shows that the underreporting estimates do not differ significantly across the wealth quartiles.

6.   The Importance of Income Underreporting by the Self-Employed

In this section, we show how not accounting for the systematic underreporting of income by the self-employed can lead to biased conclusions in a variety of settings. To do so, we re-examine several empirical examples where variation from the self-employed plays a prominent role. For the most part, we

---

[25] Our results are not sensitive to the age cutoff. Even looking at 45-55 year olds or 50-55 year olds, we see little change in our estimates of the amount of underreporting.

restrict the scope to results that can be explored using our existing PSID and CE samples. But, for some analyses, we must slightly change the sample criteria. Below, we only give a brief overview of the procedures. Collectively, the results in this section underscore our central point that taking a household survey at face value without thinking about the incentives the households have to report correctly could undermine a wide range of empirical analyses.

6A.   Differences in Earnings and Wealth between the Self-Employed and Wage and Salary Workers

The most straightforward implication of our findings pertains to papers that estimate earnings differentials between wage and salary workers and the self-employed. In a recent example, Hamilton (2000) estimates the returns to job tenure and experience in both self-employment and wage employment. Using data from the Survey of Income and Program Participation (SIPP), he constructs lifecycle wage profiles separately for wage and salary workers and for the self-employed under several alternative measures of self-employed income. He finds the median self-employed individual's wages, cumulating over 10 years, are 35 percent lower than a comparable wage and salary worker, and he assigns a prominent role to non-pecuniary returns to self-employment over alternative explanations.

Throughout his analysis, Hamilton did not adjust his estimates for the fact that the earnings data of the self-employed could be substantially underreported. In this subsection, we do not attempt to replicate the analysis of Hamilton using the SIPP data. The reason for this is that it is obvious that not accounting for the underreporting of income by the self-employed will severely bias these results. As we show above, the self-employed earnings are underreported by roughly 25 percent. As a result, it is likely that a substantial amount of the earnings gap estimated by Hamilton is simply due to the measurement error in household surveys with respect to self-employment income.[26]

6B.   The Empirical Importance of Precautionary Savings

---

[26] We want to stress that we are not saying that non-pecuniary benefits are not important for the self-employed. Moskowitz and Vissing-Jorgensen (2002) do account for the fact that the self-employed do underreport their income. Even with this adjustment, they still conclude that non-pecuniary benefits of self-employment could be very large. Additionally, in the NBER working paper version of the paper, we show how the underreporting of income by the self-employed can alter the estimates in Gentry and Hubbard (2004) showing that the self-employed are much wealthier than wage and salary workers conditional on income. Not surprisingly, if the self-employed underreport their income, their wealth conditional on income will be biased upwards.

The above example shows how accounting for the underreporting of income by the self-employed could lead to substantial changes in the results that directly compare the economic behavior of the self-employed to that of wage and salary workers. In the next several examples, we show that the underreporting of income by the self-employed can bias a variety of other types of analyses.

We start by focusing on the importance of precautionary savings. Carroll and Samwick (1998) use data from the PSID during the 1980s to explore the relationship between wealth holdings and income risk conditional on the level of income and other demographics. They measure income risk as the variance of actual household income as well as decompose the variance of household income into its permanent and transitory components. They instrument income risk using industry dummies. They conclude that up to 50 percent of household wealth accumulation for households under the age of 50 is due to precautionary motives.

Defining a sample that is analogous to Carroll and Samwick and using the same specification as theirs, we find that 47.5 percent of household saving can be attributed to precautionary motives. This estimate is nearly identical to the results reported in their work.[27] However, if we increase the income of the self-employed by 25 percent to adjust for the underreporting, the estimated importance of precautionary savings is 10.3 percent lower (from 47.5 percent to 42.6 percent). What drives this decline? Industries in which the self-employed are more prevalent are also the industries which are found to have more income risk. When regressing wealth on income risk controlling for household income, it appears that households in industries where there is more income risk are holding more wealth for a given amount of income. However, some of this relationship is simply due to the fact that we are mismeasuring the level of income for those households. Our results show that ignoring the income underreporting of the self-employed can introduce upward biases in the importance of precautionary savings in explaining total wealth accumulation of younger households.

---

[27] The details of the specification and sample can be found in the online robustness appendix.

6C.   Adjustments to Life Cycle Earnings Profiles

Self-employment probabilities change over the lifecycle. For example, within our PSID and CE samples, the fraction of individuals that are self-employed rises from 4.4 percent and 5.9 percent respectively when households are 25 to 18.3 percent and 11.2 percent respectively when households are 45. They rise further to 39.9 percent and 21.6 percent when households are 65. Given that the self-employment propensities are changing over the lifecycle, the extent to which individual earnings is measured with error is changing over the lifecycle.

How much of the decline in average earnings between the ages of 45 and 65 could be driven by the underreporting of income by the self-employed? Within the PSID and CE samples, respectively, average earnings fell by 24.4 percent and 21.5 percent between the ages of 45 and 65.[28] Simply adjusting for underreporting reduces these declines by about 3.9 and 2.8 percentage points. Taken together, between roughly 10 and 15 percent of the decline in lifecycle earnings between the ages of 45 and 65 in the PSID and the CE can be accounted for by the underreporting of income by the self-employed.

6D.   Spatial Differences in Average Earnings

Finally, we show that the underreporting of income by the self-employed could bias conclusions about the spatial differences in earnings. As shown by Glaeser (2007), self-employment propensities differ markedly across U.S. cities. Given this, comparing earnings across space will yield biased results if one does not account for differences in self-employment propensities across space.

Using the Public Use Microdata Sample (PUMS) from the 2000 Census, we illustrate this point. We construct a sample of male heads between the ages of 25 and 64, who are currently working, usually work more than 30 hours per week, and worked 40 weeks during the previous year. We further restrict this sample to only include individuals who live in a defined MSA. Within this sample, we find a tremendous amount of variation in self-employment rates across MSAs. Across the MSAs, the mean and median self-employment rates were 12.5 percent and 11.9 percent, respectively. The standard deviation

---

[28] A more formal analysis of this question would want to adjust for either individual or cohort fixed effects. This example is only meant to be illustrative.

in self-employment rates across the MSAs was 2.9 percent.

To illustrate the potential for the underreporting of income by the self-employed to yield biased results when computing earnings differentials across place, we look at a few pair wise comparisons. To do this, we examine conditional earnings differentials. Specifically, we regress log earnings on one-year age dummies and a series of education, usual hours worked per week, and weeks worked per year dummies. We then compare these residuals, with and without adjusting for the underreporting of income by the self-employed, across MSAs.

For example, not accounting for the underreporting of income and conditional on the demographic and work hour adjustments, individuals in our sample from Nassau County, New York earned 8.9 percent more than individuals in our sample from Detroit. 16.8 percent of the individuals from Nassau County were self-employed while only 10.2 percent of the individuals from Detroit were self-employed. Adjusting upward the income of all self-employed individuals, conditional on observables, by 25 percent increases the Nassau-Detroit earnings differential by nearly 15 percent (from 8.9 percent to 10.3 percent). The earnings differential between the Barnstable-Yarmouth, Massachusetts MSA (which has the highest self-employment rate among MSAs at 25.3 percent) and the Kokomo, Indiana MSA (which has the lowest self-employment rate among MSAs at 6.0 percent) gets reduced by 30 percent when adjusting for income underreporting by the self-employed (from -14.7 percent to -10.4 percent).

6E.   Potential Other Applications

The above results show that not accounting for the underreporting of income of the self-employed can lead to quantitatively important biases when comparing the income and savings behavior of the self-employed with wage and salary workers, when computing the economic importance of precautionary savings, when estimating lifecycle earnings profiles, and when computing earnings differentials across space. There are many other situations where not adjusting for the underreporting of income by the self-employed can potentially lead to quantitatively important biases. For example, given that self-employment propensities differ across races and genders, failing to account for the underreporting of

income could lead to biased estimates of racial and gender earnings gaps. Likewise, there is much interest in comparing the borrowing and default behavior of the self-employed to wage and salary workers. Not properly measuring income for the self-employed could lead to biased results in this setting as well. We leave an assessment of the quantitative importance of these potential biases to future work.

7. Conclusion

Essentially all empirical work using data from household surveys assumes that household income is not systematically mismeasured. However, there is reason to believe that this assumption may not hold, particularly for the self-employed. Research from tax audits finds that the self-employed substantially underreport their income to tax authorities. What was less known to the research community is whether and the extent to which the self-employed underreport their income within U.S. surveys. Our paper contributes to the literature by filling this gap.

Using data from both the Consumer Expenditure Survey and the Panel Study of Income Dynamics, we find that, on average, the self-employed underreport their income by about 25 percent within household surveys relative to wage and salary workers. The results are remarkably consistent across both surveys. As robustness analyses, we implement a sequence of alternative specifications that relax various elements of our assumptions. Our results are essentially unchanged throughout.

The results in this paper also have larger implications. There is an active literature showing that individuals respond to incentives when reporting information to administrative sources (like tax authorities) or alter their behavior when participating in experiments. Yet, researchers often ignore the potential of such problems spilling over into the household surveys that we use for much of our economic analysis. It is naive to assume that individuals who are willing to misconstrue their behavior to administrative sources would otherwise automatically provide accurate responses when participating in household surveys. While it is likely true that the benefits of providing distorted information to household surveys are small, it is also likely true that the costs of providing inaccurate information are also small. To the extent that individuals would have to exert effort to provide an accurate response to

household surveys or feel compelled to maintain consistency in light of concerns about confidentiality, economic theory suggests that they would continue to provide the erroneous information, even if there is no direct benefit of doing so. We show these effects are present with respect to the income reporting of the self-employed. Future work should try to identify and explore other situations in which individuals may provide systematically biased information to household surveys.

Finally, our paper speaks to the literature that attempts to infer the measurement error in household surveys by comparing the reports of individual level variables (e.g., income) in household surveys to administrative data for the same individual. A key assumption in much of this literature is that any potential errors in the administrative data are orthogonal to the errors in household surveys. The results of this paper show that such an assumption is violated in some cases. If the household is willing to misconstrue their behavior to tax authorities, there is no reason to believe that they will not misconstrue their behavior to household surveys as well.

References

Aguiar, Mark and Mark Bils (2010). "Has Consumption Inequality Mirrored Income Inequality," *University of Rochester Working Paper.*

Aguiar, Mark and Erik Hurst (2009). "Deconstructing Lifecycle Expenditures," *University of Chicago Working Paper.*

Aguiar, Mark and Erik Hurst (2007). "Lifecycle Prices and Production," *American Economic Review*, 97(5), pp. 1533-59.

Aguiar, Mark and Erik Hurst (2005). "Consumption vs. Expenditure," *Journal of Political Economy*, 113(5), pp. 919-48.

Andreoni, James (1992). "IRS as Loan Shark: Tax Compliance with Borrowing Constraints," *Journal of Public Economics*, 49, pp. 35-46.

Andreoni, James, Brian Erard, and Jonathan Feinstein (1998). "Tax Compliance," *Journal of Economic Literature*, 36(2), pp. 818-60.

Barsky, Robert, Miles Kimball, Thomas Juster, and Matthew Shapiro (1997). "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study," *The Quarterly Journal of Economics,* 112(2), pp. 537-79.

Becker, Gary (1965). "A Theory of the Allocation of Time," *The Economic Journal*, 75, pp. 493 – 517.

Bee, Adam, Bruce Meyer, and James Sullivan (2011). "Micro and Macro Validation of the Consumer Expenditure Survey," mimeo.

Blank, Rebecca, Kerwin Charles, and James Sallee (2009). "A Cautionary Tale about the Use of Administrative Data: Evidence from Age of Marriage Laws," *American Economic Journal: Applied Economics*, 1(2), pp. 128-49.

Blundell, Richard, Martin Browning, and Costas Meghir (1994). "Consumer Demand and the Life-Cycle Allocation of Household Expenditures," *Review of Economic Studies*, 61, pp. 57-80.

Bound, John, Brown, Charles, Duncan, Greg, and Rodgers, Willard (1994). "Evidence on the Validity of Cross Sectional and Longitudinal Labor Market Data," *Journal of Labor Economics*, vol. 12(3), pp. 345-68.

Bound, John, Charles Brown, and Nancy Mathiowetz (2001). "Measurement Error in Survey Data," in <u>*Handbook of Econometrics*</u>, Volume 5, eds. James Heckman and Edward Leamer, Elsevier Science.

Carroll, Chris and Samwick, Andrew (1998). "How Important is Precautionary Saving?" *Review of Economics and Statistics*, 80(3), pp. 410-9.

Clotfelter, Charles (1983). "Tax Evasion and Tax Rates: An Analysis of Individual Returns," *Review of Economics and Statistics*, 65(3), pp. 363-73.

Deaton, Angus (1997). *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy.* World Bank Publication.

Deaton, Angus and John Muellbauer (1980). "An Almost Ideal Demand System," *American Economic Review,* 70(3), pp. 312-26.

Engström, Per and Bertil Holmlund (2009). "Tax Evasion and Self-employment in a High-Tax Country: Evidence from Sweden," *Applied Economics*, 41(19), pp. 2419-30.

Feinstein, Jonathan (1991). "An Econometric Analysis of Income Tax Evasion and Its Detection," *Rand Journal of Economics,* 22(1), pp. 14-35.

Feldman, Naomi and Joel Slemrod (2007). "Estimating Tax Noncompliance with Evidence from Unaudited Tax Returns," *Economic Journal*, 117 (518), pp. 327-52.

Fitzgerald, John, Peter Gottschalk, and Robert Moffitt (1998). "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics," *Journal of Human Resources*, 33(2), pp. 251-99.

Gentry, William and R. Glenn Hubbard (2004). "Entrepreneurship and Household Saving," *Advances in Economic Analysis & Policy*, 4(1), Article 8.

Ghez, Gilbert and Gary Becker (1975). *The Allocation of Time and Goods over the Life Cycle,* Columbia University Press, New York.

Gibson, John, Bonggeun Kim and Chul Chung (2009). "Using Panel Data to Exactly Estimate Income under Reporting of the Self-employed," *Labor Economics Working Papers*, East Asian Bureau of Economic Research.

Glaeser, Edward (2007). "Entrepreneurship and the City." Harvard Institute of Economic Research Discussion Paper No. 2140.

Gottschalk, Peter and Robert Moffitt (1994). "The Growth of Earnings Instability in the U.S. Labor Market", Brookings Papers on Economic Activity, 1994(2), pp. 217-72.

Gourinchas, Pierre-Olivier and Jonathan Parker (2002), "Consumption over the Life Cycle," *Econometrica*, 70(1), pp. 47-89.

Haider, Steven and Gary Solon (2006). "Life-Cycle Variation in the Association between Current and Lifetime Earnings." *American Economic Review*, 96(4), pp. 1308-20.

Hamilton, Barton H. (2000). "Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment." *Journal of Political Economy*, 108 (3), pp. 604-31.

Hurst, Erik and Annamaria Lusardi (2004). "Liquidity Constraints, Household Wealth, and Entrepreneurship," *Journal of Political Economy*, 112(2), pp. 319-47.

Hurst, Erik and Ben Pugsley (2011). "What do Small Businesses Do?," *Brookings Papers on Economic Activity*, 43(2), pp. 73-142.

Johansson, Edvard (2005). "An Estimate of Self Employment Income Underreporting in Finland," *Nordic Journal of Political Economy*, 31(1), pp. 99 - 109.

LaLumia, Sara (2009). "The Earned Income Tax Credit and Reported Self-Employment Income," *National Tax Journal*, 62(2), pp. 191-217.

Levitt, Steve and John List (2009). "Was there Really a Hawthorne Effect at the Hawthorne Plant? An Analysis of the Original Illumination Experiments," NBER Working Paper 15016.

Meyer, Bruce, Wallace K.C. Mok, and James Sullivan (2009). "The Under-Reporting of Transfers in Household Surveys: Its Nature and Consequences," Harris School Working Paper #09.03,

Moskowitz, Tobias J. and Annette Vissing-Jorgensen (2002). "The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?" *American Economic Review* 92(4), pp. 745-78.

Pissarides, Christopher and Guglielmo Weber (1989). "An Expenditure Based Estimate of Britain's Black Economy," *Journal of Public Economics*, 39(1), pp. 17-32.

Sabelhaus, John, David Johnson, Stephen Ash, David Swanson, Thesia Garner, John Greenlees, and Steve Henderson (2011). "Is the Consumer Expenditure Survey Representative by Income," mimeo.

Saez, Emmanuel (2010). "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy*, 2(3),pp. 180-212.

Schuetze, Herb (2002). "Profiles of Tax Noncompliance Among the Self Employed in Canada:1969-1992," *Canadian Public Policy*, 28(2), pp. 219-37.

Slemrod, Joel (1985). "An Empirical Test for Tax Evasion," *The Review of Economics and Statistics*, 67(2), pp. 232-38.

Slemrod, Joel (2007). "Cheating Ourselves: The Economics of Tax Evasion," *Journal of Economic Perspectives,* 21(1), pp. 25-48.

Solon, Gary (1992). "Intergenerational Income Mobility in the United States," *American Economic Review,* 82(3) pp.393-408.

Table 1: Descriptive Statistics for Self-Employed and Wage and Salary Worker Samples, CE and PSID

| | Data Set | | | | | |
| | I. CE | | II. PSID: 1-Year | | III. PSID: 3-Year Average | |
Income Measure/Sample	Self Employed	Wage Workers	Self Employed	Wage Workers	Self Employed	Wage Workers
Panel A: Log Income Measures						
Labor Plus Business Income	10.81	10.83	11.01	10.90	11.14	10.97
Total Family Income	10.87	10.86	11.20	11.00	11.33	11.06
Total Family Disposable Income	10.83	10.75	11.04	10.86	11.15	10.91
Panel B: Log Expenditure Measures						
Food Expenditure	8.85	8.67	8.97	8.73	9.05	8.79
Nondurable Expenditure	9.98	9.78	-	-	-	-
Total Expenditure	10.74	10.51	-	-	-	-
Panel C: Other Labor Supply and Demographic Measures						
Average Age of Head	41.4	39.0	40.8	38.0	41.5	38.3
% Education = 12	26.9	29.2	20.5	24.5	22.1	25.0
% Education = Some College	23.3	24.4	30.8	30.7	31.6	30.5
% Education = College	20.4	19.0	21.6	19.2	21.0	19.5
% Education = Graduate School	17.9	12.5	15.1	10.0	15.1	10.4
% Black	3.3	8.0	2.9	9.9	1.7	9.4
% Married	83.2	80.6	86.3	77.9	90.2	79.4
Average Family Size	3.4	3.26	3.3	3.1	3.4	3.1
Average Head Work Hours	2,562	2,317	2,556	2,295	2,588	2,284
% Own Home	78.2	69.5	81.2	70.0	87.2	73.6
% Non-Positive Total Family Income	1.6	0.2	1.1	0.1	0.9	0.0
% Non-Positive Labor+Bus. Income	2.6	0.4	2.0	0.1	0.3	0.0
Sample Size	2,508	24,711	4,446	31,988	1,901	16,332

Notes: This table compares income, expenditure, and demographic variables between the self-employed and wage and salary workers within the CE and PSID . Each sample is restricted to include all households with male heads between the ages of 25 and 55 (inclusive) who are working full time and have positive levels of income and expenditure. Additional sample restrictions for each survey are discussed in the text. For the fraction of households with non-positive income, we compute this fraction assuming all sample restrictions hold aside from the restriction that the household have positive levels of income and expenditure.

Table 2: Log Expenditure Differences Between Self-Employed and Wage Workers, Conditional on Log Income and Demographics

Sample/Expenditure Measure	I. Income = Pre-Tax Total Family Income		II. Income = After-Tax Total Family Income	
	OLS	IV	OLS	IV
A. Estimates of γ				
CEX, Nondurables	0.17	0.18	0.14	0.13
	(0.01)	(0.01)	(0.01)	(0.01)
CEX, Total Expenditure	0.20	0.21	0.17	0.16
	(0.01)	(0.01)	(0.01)	(0.01)
CEX, Food	0.15	0.15	0.13	0.12
	(0.01)	(0.01)	(0.01)	(0.01)
PSID 1-Yr, Food	0.12	0.12	0.12	0.12
	(0.01)	(0.01)	(0.01)	(0.01)
PSID 3-Yr Average, Food	0.11	-	0.12	-
	(0.01)		(0.01)	
B. Estimates of β				
CEX, Nondurables	0.37	0.60	0.38	0.63
	(0.01)	(0.01)	(0.01)	(0.01)
CEX, Total Expenditure	0.42	0.76	0.43	0.80
	(0.01)	(0.01)	(0.01)	(0.01)
CEX, Food	0.26	0.40	0.27	0.42
	(0.01)	(0.01)	(0.01)	(0.01)
PSID 1-Yr, Food	0.29	0.31	0.31	0.36
	(0.01)	(0.01)	(0.01)	(0.01)
PSID 3-Yr Average, Food	0.32	-	0.36	-
	(0.01)		(0.01)	

Notes: This table shows the OLS and IV estimates of the log-linear Engel curve from equation (6). The income elasticity is denoted as β, and the coefficient on the self-employment dummy is denoted γ. We estimate the regression using different measures of expenditures and for different samples (indicated down the rows) and for different measures of household income (indicated across the columns). Robust standard errors are shown in parentheses. The samples are the same as the ones discussed in Table 1. The demographic controls are a series of five-year age dummies, a dummy if the household head is black, a dummy if the household head is married, and a series of family size dummies. We use education dummies as instruments for permanent income in the IV specifications in the CE and one-year PSID samples. For the PSID three-year average sample, we do not instrument. Instead our measure of permanent income is average household income over three years.

Table 3: Predicted Fraction of Income Underreported by Self-Employed, Baseline Specification

	Estimate of $(1-\kappa)$, in Percent	
Expenditure Measure	I. Income = Pre-Tax Total Family Income	II. Income = After-Tax Total Family Income
CEX, Nondurable Expenditure	25.5 (1.5)	18.8 (1.6)
CEX, Total Expenditure	24.5 (1.5)	17.8 (1.5)
CEX, Food Expenditure	31.7 (2.0)	25.2 (2.0)
PSID 1-Yr, Food	31.5 (2.1)	28.8 (2.0)
PSID 3-Yr Average, Food	29.4 (2.1)	27.8 (2.0)

Notes: This table shows the estimates of the amount of underreporting of income by the self-employed, denoted by $(1 - \kappa)$, using the estimates of β and γ from Table 2. We show the different estimates of $(1-\kappa)$ using different measures of expenditure, using data from different surveys, and using different measures of income.   For the CE and one-year PSID sample, we use the IV estimates from Table 2. For the PSID three-year average sample, we use the OLS estimates. Asymptotic standard errors are computed for the transformation $1-\kappa$ using a robust estimate of the variance covariance matrix.

Table 4: Predicted Fraction of Income Underreported by Self-Employed, Alternate Specifications

Expenditure Measure	Specification/Sample			
	1	2	3	4
CEX, Nondurables	27.5	18.4	16.9	19.8
	(1.6)	(1.6)	(2.6)	(1.9)
CEX, Total Expenditure	26.5	17.7	14.3	20.0
	(1.5)	(1.5)	(2.4)	(1.9)
CEX, Food	33.5	24.9	26.9	23.9
	(2.0)	(2.0)	(3.9)	(2.3)
PSID 1-Yr, Food	37.7	28.0	29.5	26.8
	(2.0)	(2.0)	(2.7)	(2.7)
PSID 3-Yr Average, Food	36.0	26.8	27.4	27.2
	(2.0)	(2.1)	(2.6)	(2.9)
Income Measure	Labor Earnings Plus Business Income	After-Tax Total Family Income	After-Tax Total Family Income	After-Tax Total Family Income
Specification	Base	Include Log Work Hours as a Control	Restrict To Heads With Age 25-40	Restrict To Heads With Age 40-55

Notes: This table shows the amount of underreporting of income by the self-employed $(1-\kappa)$ for alternate specifications of equation (6) from the text. In column (1), we re-estimate the specification in Table 3 with labor earnings plus business income as our measure of income. In column (2), we add log work hours to our vector of control variables. In columns (3) and (4), we split the samples by age (young vs. old). For all specifications, the sample is otherwise the same as the one described in the note to Table 1. Asymptotic standard errors are in parentheses.

Figure 1: Nonparametric Estimate of Food Engel Curve Using Three-Year PSID Panels

Notes: Figure shows the estimates of (3) from the text estimated nonparametrically for both wage and salary workers and for the self-employed. When creating the figure, we divide the support of log income residuals into equally spaced bins and fit a cubic spline through the median log average income and log average expenditure for each bin. Additionally, when making the figure, we use our three-year average PSID sample. Our measure of income is the three-year average of after tax family income. Our measure of expenditure is the three-year average of food expenditures. Our vector of additional controls is described in the text. See text for additional sample restrictions.