

NBER WORKING PAPER SERIES

ARE HOUSEHOLD SURVEYS LIKE TAX FORMS:
EVIDENCE FROM INCOME UNDERREPORTING OF THE SELF EMPLOYED

Erik Hurst
Geng Li
Benjamin Pugsley

Working Paper 16527
<http://www.nber.org/papers/w16527>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2010

We would like to thank the following for their helpful comments and discussions: Kerwin Charles, Raj Chetty, Steve Davis, Joshua Gottlieb, Bruce Meyer, Matt Notowidigdo, Jesse Shapiro, and Bob Willis. We would also like to thank seminar participants at Boston College, Chicago, Harvard Business School, and Stanford. The views presented in this paper are those of the authors and do not represent those of the Federal Reserve Board, its staff, or the National Bureau of Economic Research. Hurst and Pugsley gratefully acknowledge the financial support provided by the George J. Stigler Center for the Study of Economy and the State. Additionally, Hurst thanks the financial support provided by the University of Chicago's Booth School of Business.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2010 by Erik Hurst, Geng Li, and Benjamin Pugsley. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Are Household Surveys Like Tax Forms: Evidence from Income Underreporting of the Self Employed

Erik Hurst, Geng Li, and Benjamin Pugsley

NBER Working Paper No. 16527

November 2010

JEL No. C8,E21,H26,J3

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 systematically underreport their income to 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 the self employed underreport their income by about 30 percent. This result is remarkably robust across data sources and alternative model specifications. Aside from transportation expenditures, we find little evidence that the self employed misreport their expenditures to household surveys. We show that failing to account for such income underreporting leads to biased conclusions when comparing the earnings and saving behavior between the self employed and other workers as well as biased estimates of the importance of precautionary savings, the shape of lifecycle earnings profiles, and the magnitude of earnings differences across MSAs. Finally, our results show that it is naive for researcher to take it for granted that individuals will provide unbiased information to household surveys when they are simultaneously providing distorted information to other administrative sources.

Erik Hurst
Booth School of Business
University of Chicago
Harper Center
Chicago, IL 60637
and NBER
erik.hurst@chicagobooth.edu

Benjamin Pugsley
Department of Economics
University of Chicago
Chicago, IL 60637
benjamin.pugsley@gmail.com

Geng Li
Federal Reserve Board
MS 93
Washington, DC 20551
geng.li@frb.gov

1. Introduction

There is an abundance of evidence that finds that some individuals systematically underreport their incomes to tax authorities.¹ Likewise, there is a sizeable literature showing that some participants in experiments will distort their behavior as a reaction to being studied.² There is also evidence that some other types of administrative data, including Vital Statistics data, are not immune from potential problems of misreporting.³ Collectively, these literatures show 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).

An implicit assumption made in nearly all empirical work using household survey data, however, is that the data within the household surveys are immune to such systematic errors.⁴ 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 set out to address 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. We ask whether household surveys are akin to other types of data sources where systematic measurement errors can undermine data quality.

A natural question that arises is why would any individual purposefully misreport their earnings to household surveys? On the one hand, there is very little direct benefit for an individual to underreport their income to household surveys. Unlike when reporting to tax authorities, misreporting income to

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) compares 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.

⁴ There is an important literature examining measurement error in household surveys. We briefly discuss some of this literature below. When doing so, we discuss how our results contribute to this literature.

household surveys does not lessen the individual's tax burden. On the other hand, there is also essentially no penalty associated with misreporting income to household surveys.

However, individuals, particularly the self employed, may have other indirect net benefits associated with underreporting their income to household surveys. First, households may feel compelled to provide consistent measures of their income to household surveys if they believe that their reports to the household survey may not be completely confidential. Given that the self employed already have strong incentives to misreport their income to tax authorities and given that they face no penalty for misreporting to household surveys, even a small probability of self-incrimination may be enough to distort their survey responses. Second, unlike wage and salary employees who receive W-2 forms, the self employed have to expend effort to accurately account for their true income. If the self employed have already supplied a falsified report to tax authorities, it may simply be easier to reuse this report when responding to household surveys as opposed to computing an accurate one. With no penalty compelling an accurate response and potential net benefits from providing a distorted one, economic theory predicts that households, particularly those who already underreport their income to a tax authority, have incentives to also underreport their income to household surveys. The extent of underreporting of income to household surveys, therefore, becomes an empirical question.

We focus our attention on the self employed for two reasons. First, as noted above, there is a large literature showing that this group consistently and substantially underreports their income to the tax authorities. For example, for tax years 1985, 1988, and 1992, IRS random audits suggest that roughly 30-35 percent of nonfarm proprietor income was underreported.⁵ The degree of underreporting was even higher in recent years when the IRS estimated that 57 percent of nonfarm proprietor income was underreported in tax year 2001.⁶ Second, the self employed are a non-trivial fraction of household survey participants. They comprise between 10 and 14 percent of a sample of prime age working males, and

⁵ Source: "Individual Income Tax Gap Estimates for 1985, 1988, and 1992" which can be found at <http://www.irs.gov/pub/irs-soi/p141596.pdf>.

⁶ Source: "Reducing the Federal Tax Gap: A Report on Improving Voluntary Compliance" which can be found at http://www.irs.gov/pub/irs-news/tax_gap_report_final_080207_linked.pdf.

some surveys even over-sample the self employed.⁷ Furthermore, self employment propensities differ markedly over the lifecycle, across gender and race, and across geographic areas. Consequently, as we show below, the extent to which the self employed underreport their income has quantitatively important effects on many different types of questions addressed in empirical economics using household survey data.

Our goals in this paper are threefold. First, we will infer the extent to which income is underreported by the self employed within U.S. household surveys. As far as we know, this is the first paper that attempts to do so. To estimate the extent of underreporting, we follow a procedure similar 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 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 the self employed underreport their income to household surveys by about 30 percent. These magnitudes are nearly identical across both surveys for similarly defined specifications.

The above 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 the conceptual measures of income in household surveys are similar between wage and salary workers and the self employed. 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 discuss the validity of all of these assumptions, as well as provide a variety of additional robustness specifications, in later sections.

The second goal of the paper is to assess whether the extent of the underreporting of income correlates with observable economic or demographic characteristics. Research using data from actual tax returns finds that individuals are more likely to underreport their income when marginal tax rates are higher (Clotfelter, 1983) and that the self employed in professional industries are less likely to

⁷ For example, the share of households surveyed in the Survey of Consumer Finances that are self employed is significantly higher than the nation average, partly due to the survey's oversampling of the high-income households.

underreport their income relative to those in non-professional industries (Andreoni et al., 1998). Additionally, understanding both the extent of underreporting and whether the underreporting is correlated with observables is important for gauging when and how ignoring the underreporting could potentially bias different analyses. We do find greater estimated underreporting of income in the early part of our sample (prior to 1986) when marginal tax rates were higher relative to later parts of the sample. We also find evidence that the self employed with advanced degrees tend to underreport their income to household surveys to a much less extent.

The third and final 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. Not surprisingly, we show that studies that focus on earnings or savings differences between the self employed and wage and salary workers are severely biased if the underreporting of income by the self employed is not accounted for. For example, Hamilton (2000) finds that the earnings of the self employed are 35 percent lower than that of otherwise similar wage and salary workers. From this finding, he concludes that there are substantial non pecuniary benefits to self employment. However, accounting for the underreporting of income by the self employed dramatically reduces the earnings gaps estimated by Hamilton. Additionally, we show that adjusting for the misreporting of income by the self employed reduces the wealth gaps (conditional on earnings and demographics) between the self employed and wage and salary workers by roughly one third.

Moreover, we show that not accounting for the underreporting of income by the self employed can lead to quantitatively different results in a variety of other settings. For example, we show that estimated results on the importance of precautionary savings are 13 percent lower when we account the extent to which the self employed underreport their income. Similarly, 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 5 and 10 percent of the decline in earnings between the ages of 45 and 65 in the

Consumer Expenditure Survey and the Panel Study of Income Dynamics can be explained by the underreporting of income by the self employed.

Our work in this paper complements two different literatures. First, there is an important existing literature that looks at reporting errors within household surveys.⁸ For example, Bound et al. (1994) explore the extent to which classical measurement error is important for reported wage and earnings data in household surveys. Fitzgerald et al. (1998) examines the nature of survey attrition from household survey panels. Haider and Solon (2006) show that deviations between current and permanent income change over the lifecycle. Meyer et al. (2009) documents the extent to which transfer income is underreported in household surveys. Aguiar and Bils (2010) looks at differences in the underreporting of expenditures by different groups within household surveys and shows that these differences have been changing over time. Deaton (1997) outlines many potential data problems that could arise when using household surveys for empirical analysis. Our work adds to this literature by identifying another channel for 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 pass through to household survey responses.

Second, our work complements the large existing literature that estimates the extent of underreporting of income by the self employed using data on actual tax returns. Much of the work on the underreporting of income to tax authorities comes from the Taxpayer Compliance Measurement Program (TCMP) and the recent National Research Program (NRP). The TCMP and NRP are a stratified random sample of U.S. individual federal income tax returns that are subject to thorough examination by experienced IRS tax examiners.⁹ A separate strand of research uses data from actual tax returns (as opposed to random audits) to assess the amount of underreporting of income. For example, Feldman and Slemrod (2007) uses tax return data from 1999 to estimate the relationship between charitable giving and

⁸ Bound et al. (2001) surveys nearly fifty years of research examining the extent to which household survey data is measured with error.

⁹ There is much work using the TCMP and NRP data to assess the underreporting of income on U.S. tax returns. For a summary of this literature, see the recent surveys by Andreoni et al. (1998) and Slemrod (2007).

income for those with only wage and salary income. Again, our study differs from theirs in that we analyze the underreporting of income by the self employed to household surveys as opposed to tax authorities.¹⁰

Collectively, our work shows that it is naive for researchers to assume that individuals who are willing to misconstrue their behavior to other types of data sources would 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 that the benefits of providing accurate information are also small. Such potential biases need to be accounted for when interpreting data from household surveys.

2. Data Description: CE, PSID and CPS

We use two nationwide representative household surveys, the Consumer Expenditure Survey (CE) and the Panel Study of Income Dynamics (PSID), for the majority of our empirical analysis. To explore the quality of income data within both surveys, we also supplement our analysis with data from the Current Population Survey (CPS). In this section, we describe all three of the surveys and discuss our sample selection criteria. The Data Appendix provides background information on all three data sources.

2A. The CE Sample

To construct our sample, we start with the NBER CE extracts which pool together all the waves between 1980 and 2003.¹¹ We restrict the sample to households who report expenditures in all four quarters of the survey and sum the four quarterly responses to calculate an annual expenditure measure.

¹⁰ Several authors have examined the underreporting of income by the self employed within household surveys from other countries. For example, Pissarides and Weber (1989) uses 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. Gorodnichenko et al (2009) uses household expenditure data and changes in national tax policy to examine the extent to which taxes are evaded in Russia.

¹¹ See http://www.nber.org/data/ces_cbo.html for details on the NBER CE extracts.

We further restrict the sample to include only the households that have a male head between the ages of 25 and 55 (inclusive), and where the head reports currently working at least 30 hours in an average week and has worked at least 40 weeks during the previous year.¹² These latter restrictions ensure that we focus on prime age individuals who are working full time. To avoid some ambiguities in classifying households, we further exclude any households where the male head is a wage and salary worker but the spouse (if present) was self employed.¹³ We also restrict the sample by excluding any household who reported a positive amount of farm income. We exclude farmers because the relationship between income and food expenditures may differ for individuals who grow some of their own food. Finally, we exclude households with zero or negative reported measures of household income or zero reported measures of household expenditures.¹⁴

In the work below, we explore how the relationship between income and expenditures for wage and salary workers compares to the relationship between income and expenditures for the self employed. We define three measures of expenditures and two measures of income using the CE data. The three expenditure measures we explore are: total food expenditures, total nondurable expenditures, and total expenditures. Total food expenditure includes expenditures on food consumed at home as well as expenditures on food consumed away from home. Following Aguiar and Hurst (2009), we define nondurable expenditures to include spending on food, alcoholic beverages, tobacco, clothing and personal care, housing services, utilities, domestic care services, nondurable transportation, nondurable entertainment, and “other” nondurable expenditures. For a full definition of each of these categories, see

¹² Male heads are identified 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 themselves as being a head.

¹³ Within the CE, there are only 859 households where the male household head is working full time but is not self employed but the spouse within the household was self employed after satisfying all other sample restrictions. This constitutes 3.1 percent of the CE sample. Our results are not sensitive to whether or not we include such households within our self employed or wage and salary worker sample. For the results below, we simply exclude from our analysis all households where the wife is self employed (conditional on a wife being in the household) and the husband is not.

¹⁴ This last restriction is not overly binding. Given the other sample restrictions, only about 0.5 percent of our sample had missing income measures. All remaining households had positive expenditures measures.

Aguiar and Hurst (2009).¹⁵ 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 the paper, we use two different measures of household income. For our first measure of household income, we include the sum of all labor earnings from wage and salary employment plus all earnings generated by one's business. For business earnings, we include both the return to labor and the return to capital.¹⁶ When creating this income measure, we include both the earnings of the male heads and their spouses (if a spouse was present). Our second measure of household income is total family income, which includes all earnings, asset income, and transfer income received by the household. This latter measure is the total amount of pre-tax resources that flow into the household during a given year that can be allocated to consumption.

There are pros and cons for each of the income measures. According to standard theories of consumption, household expenditures are determined by the present value of household resources earned over their lifetime. In this sense, it is appropriate to account for the returns accrued to both labor and capital. Accounting for asset income is potentially even more important in our analysis given that it is well documented that the self employed have much higher levels of assets than wage and salary workers that can be used to fund current expenditures.¹⁷ However, asset levels, and consequently asset income, are in part determined by household preferences, the same preferences that determine household consumption behavior. We show below that these potential endogeneity concerns have only a limited effect on our results. Indeed, our estimates of the underreporting of income by the self employed are quantitatively similar regardless of the measure of income we use.

¹⁵ Following Aguiar and Hurst (2009), we do not include health and education expenditures in our nondurable goods expenditure measure.

¹⁶ The fact that both the return to business labor and the return to business capital is combined together in our base measure of earnings is done out of necessity. The CE does not separately distinguish between the two components of business income. A similar decision is made with respect to the PSID data (described below). While the PSID does attempt to distinguish business labor income from business capital income, such a distinction is imputed for most households.

¹⁷ See Hurst and Lusardi (2004) and the references cited within.

Finally, within the CE, all workers are asked to classify their employment type as either working for someone else in the private sector, working for the government, or being self employed. We define a self employed household as one where the male head reports their primary employment type as being self employed. We refer to all remaining households with a working male heads as being wage and salary workers. Our final sample includes 27,219 households, of which 2,508 are self employed.

2B. The PSID Sample

Compared with the CE, the PSID only collects limited information on household expenditures. Over the period of our analysis, the only expenditure category that is measured with any consistency is food expenditures. When computing PSID food expenditures, we sum together expenditures on food consumed at home with food consumed away from home. We define our two measures of household income in a way similar to how they are defined within the CE.

We use the PSID data only from 1980 to 1997 except the 1988 and 1989 waves, during which food expenditures were not collected. We do not use the PSID data after 1997 because the PSID started collecting data biannually after 1997. Prior to 1997, we can match food expenditures and income more closely. The way the PSID is set up, during a given survey year, households are asked to report their current weekly or monthly food expenditures and they are asked to provide extensive information about the income earned during the previous calendar year. As a result, we match the income reported by the household during survey year $t+1$ with the food expenditures reported by the household during survey year t , which is the standard approach in the literature.¹⁸

Similar to the CE sample construction discussed above, we restrict the PSID sample to those households with positive food expenditure during a given year, whose heads are male between the ages of 25 and 55 (inclusive), report working for pay, report working at least 30 hours in an average week, and

¹⁸ See, for example, Zeldes (1989).

report having worked at least 40 weeks during the year in which food expenditures are measured.¹⁹ As with the CE data, we also exclude households with zero or negative measures of household income. Lastly, we exclude all households where the head is a wage and salary worker but the spouse (if present) was self employed.²⁰

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. Again, for simplicity, we exclude households who report being both self employed and working for someone else from our analysis. None of our results are sensitive to whether these households are included or excluded from the analysis. Finally, after applying these restrictions, our sample has 33,339 households, 4,140 of which are self employed.

2C. The CPS Sample

To assess the quality of the income data within the CE and PSID, we supplement our analysis with data from the March Annual Supplement to the Current Population Survey (CPS). The CPS is a monthly survey of about 50,000 households conducted jointly by the U.S. Census Bureau and the U.S. Bureau of Labor Statistics. It is the primary source of labor force characteristics of the U.S. population. For our analysis, we use data from the March Annual Supplement to the CPS for the years 1980-2003. Participating households in March answer additional detailed questions about their annual income and

¹⁹ Because the PSID is a longitudinal survey, we are able to restrict the sample to those whose head had worked at least 40 weeks in the year during which the expenditure occurred. In contrast, in the CE sample, the annual working week restriction refers to the number of weeks in the year prior to expenditure.

²⁰ After making the other sample restrictions, there are only about 3 percent of households where the male was not self employed but the wife was self employed. As with the CE results, our results are not sensitive to how we classify such households. For simplicity, we exclude them from our analysis.

employment in the previous calendar year. We define the CPS sample to match the PSID and CE samples.²¹

3. Income and Expenditure Differences between Wage Workers and the Self Employed

In this section, we present descriptive results on how the income and expenditure of the self employed compare with those of the wage and salary workers. These data will form the basis of our formal empirical tests in the subsequent sections. We find that, conditional on demographics, the self employed tend to have lower levels of income while simultaneously having greater expenditures. In what follows, we first convert all income and expenditure measures to year 2000 dollars using the Consumer Price Index and express them in log levels. We compute all estimates using the sampling weights of the respective surveys.

Table 1 shows descriptive statistics separately for wage and salary workers and the self employed within the CE, PSID, and CPS samples. The table is divided into three panels. In panels A and B, respectively, we show statistics of the log real household income and expenditure measures. The PSID only has limited expenditure information and the CPS has no information on expenditure. In panel C, we show other descriptive statistics. The table shows that the differences in the unconditional averages between wage and salary workers and the self employed are mostly consistent across surveys. Noticeably, on average, the self employed have a somewhat higher level of total family income relative to wage and salary workers, but even greater differences across each measure of expenditures. Across all surveys, the self employed are older, more likely to be married, better educated, and on average work 11 to 12 percent more annual hours. The difference in hours worked is notable and we will return to it when we address whether the expenditure differences could be driven by varying tastes for leisure or difference in home production needs for the self employed. On balance, these demographic characteristics are strikingly similar across surveys. As we show below, the demographic composition identified in each of the samples explains much of the higher income of the self employed.

²¹ The data appendix details the exact sample selection criteria.

Next, we partial out all of the differences explained by the different demographic characteristics and examine the conditional differences in measures of income and expenditures. To find these conditional differences we approximate the conditional expectations of income and expenditures by estimating the regression models

$$\log y_{it} = \delta_0^y + \delta_1^y D_{it} + \Psi^y ' X_{it} + \mu_t^y + \eta_{it}^y \quad (1)$$

and

$$\log c_{it} = \delta_0^c + \delta_1^c D_{it} + \Psi^c ' X_{it} + \mu_t^c + \eta_{it}^c, \quad (2)$$

where y_{it} and c_{it} indicate real measures of income and expenditure of household i in year t , D_{it} is a dummy variable indicating whether or not i is self employed in year t , X_{it} is a vector of demographic controls, μ_t is a time fixed effect, and the η_{it} terms are orthogonal prediction errors. The vector of household controls include: a series of five-year age dummies, four dummies indicating the education of the household head (exactly high school, some college, exactly college, and more than college; less than high school was omitted), a dummy if the household head was black, a dummy if the household head was married, and a series of family size dummies.

The conditional differences in income and expenditure line up rather closely across the three surveys. Tables 2 and 3 show the estimated conditional differences in income and expenditure respectively across each survey. As seen from Table 2, conditional on demographic differences, the self employed typically have lower income than wage and salary workers. Controlling for the compositional differences between the groups makes the results closer across surveys than the unconditional differences shown in Table 1. For example, for labor earnings plus business income, the self employed earn 10 percent and 11 percent less than their wage and salary worker counterparts with the same demographic characteristics, respectively, within the CE and CPS samples.²² The gaps are largely similar for total

²² The results in Table 2 show differences in earnings between the self employed and wage and salary workers. As shown in Table 1, the self employed also work more than wage and salary workers. Given this, the difference in the average log hourly wage rate between the self employed and wage and salary workers is much lower than what is

family income in that the self employed earn 8 percent less and 7 percent less, respectively, within the CE and the CPS. However, total family income of the self employed is on average 7 percent higher than the wage and salary workers in the PSID sample.

In Table 3, we show the estimates from equation (2). While the self employed, all else equal, generally reported earning less income than wage and salary workers, they also reported spending substantially more than wage and salary workers. Specifically, within the CE, the conditional differences show that the self employed reported spending roughly 12 percent more on food, 13 percent more on nondurable expenditures, and 16 percent more on total expenditure. The conditional expenditure differences are nearly identical between the PSID and CE.

These purely descriptive measures of income and expenditure reveal an inconsistency. Across alternative household surveys, the self employed simultaneously earn less and consume more than their wage and salary counterparts. To determine whether underreporting in income by the self employed can account for the inconsistency we further examine the relationship between income and expenditure in both groups.

4. Detecting the Underreporting of Income of the Self Employed

In this section, we first examine, assuming income reports are accurate, how the conditional relationship between income and expenditures differs across the self employed and wage and salary workers. Previewing our results, we find a similar relationship for both the self employed and wage and salary workers except for a gap in levels of predicted expenditures. Income underreporting is one straightforward explanation of the gap. By relaxing the accurate reporting assumption, we propose a method closely related to Pissarides and Weber (1989) to use the vertical distances in the Engel curves to infer whether, and by how much, the self employed underreport their income relative to wage and salary workers. We use this method in Section 5 to quantify the extent of underreporting by the self employed.

shown in Table 2. Controlling for hours increases the differences in earnings between the self employed and wage and salary workers in all of our specifications.

Later, in Section 6 we show that the method is robust to a number of alternative model specifications that take into account other possible explanations for the expenditure gap.

Our central theme in understanding the conditional structural relationship between income and expenditure is the log-linear Engel curve. Our first goal is to understand whether the self employed and wage and salary workers have a substantially different relationship between income and expenditures. For each group, k , where k indicates either self employed ($k = S$) or wage and salary ($k = W$), we separately estimate log-linear Engel curves using alternative measures of income and expenditures. Here we assume that each household i has preferences that, when conditioned on our set of controls, are such that the intertemporal optimization, at least to a first order approximation, generates Engel curves of the following form:

$$\ln c_{ikt} = \alpha_k + \beta_k \ln y_{ikt}^p + \Theta_k' X_{ikt} + \mu_{kt} + \varepsilon_{ikt} \quad (3)$$

where β_k is the income elasticity, X_{ikt} is our previously defined vector of controls less any education dummies, and ε_{ikt} represents the cumulative effects of other unobserved determinants of the household's consumption. Here, 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.²³

Estimating (3) directly is of course problematic because even if the household's annual income reported to the CE and PSID were completely accurate it would only by chance coincide with y_{ikt}^p .²⁴ The

²³ 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 6, we explore whether differential liquidity constraints between the two groups can explain the patterns we see in the data.

²⁴ Given the nature of the PSID data, we could have exploited the panel nature of the PSID to construct a multi-year average income measure. In a prior draft of this paper, we did just that. We made a 3-year average income and expenditure PSID sample. Using that sample, we regressed log three year average food expenditure on three year average income, our vector of controls X , and the time dummies. The results from the three year average PSID sample and the instrumented version of (5) below using the one year PSID sample were nearly identical. As a result, to save space, we only show here the results using the one year PSID samples. However, if one is interested in the three year average results, they are available upon request from the authors.

household's annual income represents both y_{ikt}^p and a transitory component. Here we label the household's income report y_{ikt} and assume for the moment that all households accurately report their current income, y_{ikt} , which we model as $y_{ikt} = y_{ikt}^p \exp(\Omega_k' X_{ikt} + v_{ikt})$ so that:

$$\ln y_{ikt} = \ln y_{ikt}^p + \Omega_k' X_{ikt} + v_{ikt}, \quad (4)$$

where $\Omega_k' X_{ikt}$ and v_{ikt} represent respectively the predictable and unpredictable components of transitory 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 (3) in terms of (4) implies:

$$\ln c_{ikt} = \alpha_k + \beta_k \ln y_{ikt} + \Psi_k' X_{ikt} + \xi_{ikt}, \quad (5)$$

where $\Psi_k = \Theta_k - \beta_k \Omega_k$ and $\xi_{ikt} = \varepsilon_{ikt} - \beta_k v_{ikt}$. Even with accurate income reports, the transitory income introduces attenuation bias in our estimation of β_k from (5) since, by construction, $E[\ln y_{ikt} \xi_{ikt}] \neq 0$. To mitigate the effects of measurement error in our estimation of equation (5), we instrument for y_{ikt} using educational attainment dummies. As is the usual case, the education dummies are highly predictive of household reported annual income in the first stage regressions.

Figures 1, 2 and 3 plot simple income-expenditure Engel curves estimated separately for the two groups. Figure 1 shows the results using CE data where nondurable expenditure is used as the expenditure measure. Figure 2 shows the results using CE data where food expenditure is the expenditure measure. Figure 3 shows the results using PSID data. In all three of these figures, the measure of income is household total family income. To facilitate comparison, for each figure, we show the predictive relationship between income and expenditure estimated by OLS and by 2SLS instrumenting for $\ln y_{ikt}$.

A few things are of note of the estimated Engel curves shown in Figures 1-3.²⁵ First, for all specifications, the Engel curve for the self employed lies strictly above the Engel curve for the wage and salary workers over essentially the entire support of the data. In other words, for a given amount of income, the self employed spend more on both food and total nondurable expenditures than the wage and salary workers. These are consistent with the descriptive results in Tables 1-3.

Second, for the IV specifications, the slopes of the Engel curves for the wage and salary workers are roughly similar to the slopes for the self employed. For example, for food expenditures in the CE, the estimated income elasticities for the self employed and for the wage and salary workers are 0.38 and 0.40, respectively (Figure 2). The comparable income elasticities using food expenditures in the PSID were 0.31 and 0.32, respectively (Figure 3). The differences in income elasticities were somewhat bigger using CE nondurable expenditure data (0.51 and 0.61, respectively). When we constrain the effect of covariates to be the same across groups, we find these slopes to be statistically indistinguishable across groups. Although we only show the figures using total family income as our income measure, the results are nearly identical if we use labor plus business income as our income measure.

Collectively, the above descriptive results suggest that if income reports in the surveys are accurate, the self employed consume substantially more than wage and salary workers. We now outline a very simple model that relaxes the assumption of accurate reporting of income and allows us to estimate the magnitude of underreporting by the self employed.

We begin with the same household preferences and market arrangements that generate the log-linear Engel curve described by (3). However, now we relax our earlier assumption of accurate income reporting to allow for the possibility of misreporting of income by the self employed. For households of each type, we now assume their income reports to the household surveys are determined by

$$\ln y_{iWt} = \ln y_{iWt}^p + \Omega_W' X_{iWt} + v_{iWt} \quad (6)$$

$$\ln y_{iSt} = \ln \kappa_S + \ln y_{iSt}^p + \Omega_S' X_{iSt} + v_{iSt} \quad (7)$$

²⁵ To plot the Engel curves, we evaluate our estimate of (5) using means of pooled self employed and wage and salary workers for each sample.

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 .

Equations (6) and (7) 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, $\kappa_S \leq 1$. For now, we assume that κ_S is constant across self employed households.²⁶ Second, we assume that the wage and salary workers provide a systematically unbiased report of their income to household surveys i.e., $\kappa_W = 1$. This assumption is not overly restrictive. If wage and salary household also systematically mis-report 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 estimated parameters of the two Engel curves are constant between the self employed and the wage and salary workers. This amounts to imposing that $\alpha_W = \alpha_S$, $\beta_W = \beta_S$, $\Theta_W = \Theta_S$, and $\Omega_W = \Omega_S$. As seen in Figures 1-3, the slopes of the income-expenditure Engel curves do not differ dramatically between the two groups so this assumption is not at odds with the underlying data.²⁷ 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 by the same amount. We provide evidence for this assumption in Section 6.

²⁶ Later in the paper, we explore how κ varies across different demographic groups and across different tax regimes.

²⁷ Moreover, we cannot reject equality of the coefficients on the vector of controls between the two groups.

Given these assumptions, we can estimate of κ_S via the following expression formed from combining (5) with (6) and (7)

$$\ln c_{ikt} = \alpha + \beta \log y_{ikt} + \gamma D_{it} + \Psi X_{ikt} + \mu_t + \xi_{ikt}, \quad (8)$$

where D is again a dummy variable indicating whether the household head is self employed. As before $\Psi = \Theta - \beta' \Omega$ 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 (8). 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 $1 - \hat{\kappa}_S$.

When estimating (8) using the pooled sample, we again instrument for reported $\ln y_{ikt}$ using the set of education dummies. Estimating (8) with OLS would underestimate the income elasticity of consumption β and hence overestimate κ .

To summarize, our main identifying assumptions are: 1) household's preferences are such that log household expenditures can be expressed as a linear function of log household permanent income conditional on our control set, 2) self employed households underreport their true income by an unknown factor, κ_S , to household surveys relative to wage and salary workers, 3) expenditures are not differentially over or underreported between the two groups, 4) the underlying relationship between income and expenditures—absent the underreporting of income by the self employed—is similar between wage and salary workers and the self employed and 6) a set of education dummies is a valid instrument for the permanent component of household income.

In section 6, we assess in greater detail two of the above assumptions. First, we provide evidence in support of the assumption that there are no systematic differences in the measurement of expenditures between the two groups within household surveys. Second, aside from underreporting, we discuss and rule out several alternative explanations for the excessive expenditures of the self employed.

5. Estimating the Missing Income using Expenditures

In this section, we show the base results from estimating (8). Table 4 shows our estimates of both β and γ from equation (8) using different measures of expenditures separately for the CE and PSID samples. For comparison, we show both the OLS and IV results. Also, we separately show the results where we use total family income and household labor plus business income as our measure of household income. Remember that β estimates the slope of the income-expenditure log-linear Engel curves pooled across the two groups while γ estimates the average intercept gap between the log-linear Engel curves of the self employed relative to the wage and salary workers. A positive value of γ implies that, conditional on a given amount of reported income, expenditures for the self employed are higher than the expenditures of the wage and salary workers.

As seen from Table 4, across all expenditure measures and across both surveys, estimates of γ are positive and statistically different from zero (shown in panel B). 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. Moreover, the estimates are similar regardless of whether we use total family income or labor plus business income as our income measure. Finally, the estimates of γ for food expenditures are similar between the PSID and CE samples. For example, using the IV specification and focusing on labor plus business income, the estimate of γ in the CE sample was 0.15 (standard error = 0.01) while the estimate of γ in the PSID was 0.11 (standard error = 0.01).

The estimates of β from (8) are shown in Panel A of Table 4 and are similar to those shown in Figures 1-3. 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 (7) 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, the IV estimates of β are higher than the OLS estimates of β . Furthermore, the difference between the OLS and IV estimates is larger in the CE specifications relative to the PSID specifications. This is consistent with the fact that income within the

CE is measured with more error than in the PSID. The IV estimates of the income elasticities for food expenditures range from 0.31 to 0.40 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.²⁸

Table 5 is the central result of the paper, quantifying the self employed’s missing income. In this table, we use the estimates of β and γ shown in Table 4 to construct estimates of $(1-\kappa_s)$ —the estimated underreporting of income by the self employed relative to wage and salary workers. We estimate this reporting gap as $(1-\hat{\kappa}_s) = 1 - \exp\left(\frac{-\hat{\gamma}}{\hat{\beta}}\right)$ and compute standard errors using standard asymptotic approximations. In this table, we only show the results from the IV regressions, which are the appropriate way to estimate (8) given the measurement error in income.

The results of Table 5 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 and 30.1 percent in the PSID. Using labor plus business income and food expenditures yields estimates of 33.5 and 34.9 percent, respectively, in the CE and PSID. The estimated underreporting is slightly smaller using broader measures of consumption. For example, the estimated underreporting ranges from 25-28 percent using nondurable expenditures and total expenditures, across the various specifications.

The results in this section, given the estimating assumptions, 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.

²⁸ See, for example, Aguiar and Hurst (2009).

6. Examining the Validity of Key Identification Assumptions

In this section, we address 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, Pissarides and Weber (1989) included, have found this to be a reasonable approximation. Besides analytic convenience, we fail to reject models incorporating higher order terms for all of our income and expenditure measures.²⁹ Second, we constrain the effect of controls and the income elasticity to be identical across groups. In all samples, as noted above, 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 it is inconsistent across samples and small in magnitude.³⁰ Three, we rely on a set of educational attainment dummies to instrument for permanent income. While the dummies are highly predictive of income, their exogeneity is certainly an open question. Lacking a compelling alternative, and given the improvement in the estimated elasticities, we proceed under their assumed exogeneity.

Notwithstanding the preceding assumptions, underreporting remains the residual explanation for a difference in predicted expenditure across groups. It is useful to think about the separate Engel curves plotted in Figures 1—3. The measure of underreporting follows from converting the expenditure gap between the Engel curves into an income gap. The crucial identifying assumption of our model is attributing these expenditure gaps to income underreporting. However, there are other potential explanations for this expenditure gap even if income were accurately reported. In the remainder of this

²⁹ These specifications and all of the other robustness exercises are detailed in our online robustness appendix.

³⁰ We do this by estimating (8) and allowing β to vary across the two groups by additionally interacting $\ln y$ with D . This evidence is not robust across samples and alternative measures of income and expenditure. In both the PSID and CE, and using food expenditures and either measure of income, the data fail to reject equality of the income elasticities. More importantly, even when the differences are statistically significant, the estimated differences in income elasticities are economically small. We interpret this evidence as suggesting that, if anything, our underreporting results may be a lower bound on the actual level. In each sample, whether significant or not, the income elasticity of the self employed is less than the workers. If this is true, then smaller values of β would imply our estimates of $1-\kappa$ are biased downwards.

section we address several alternative explanations and show that our estimates are robust to specifications that take the alternatives into account.

6A. 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 mis-reporting 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. Yet, if the self employed also use that vehicle for personal use, 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 overreporting their personal expenditures to household surveys.

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 likely not 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 are 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 (7) 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 yields an underreporting of income by the self employed of roughly 30 percent, nearly identical to the food expenditure, nondurable expenditure, and total expenditure results.

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.

The results above, however, beg a different question. Is it possible that the self employed underreport both income *and* expenditures by uniformly misreport their spending on all categories (aside from evidence on nondurable transportation expenditures)? It is hard to tell a story that would generate this result. However, it is possible for us to test this potential hypothesis directly. To do so, we estimate the following relationship based on Deaton and Muellbauer's (1982) Almost Ideal Demand System such that:

$$\omega_{jikt} = \alpha + \mu_t + \beta_j \log C_{ikt}^* + \Theta' X_{ikt} + \varepsilon_{ikt}, \quad (9)$$

where ω_{jikt} is the share of expenditures allocated to category j , C_{ikt}^* represents the actual total expenditures of household i in period t and X_{ikt} is our maintained vector of controls. β_j , in this case, represents how the share of spending on a given category changes as total expenditures increase. β_j less than 1 (and greater than zero) reflect categories that are relative necessities while β_j greater than 1 reflect categories that are relative luxuries. There is a large variation in estimated β_j 's across different consumption categories (see Banks et al (1997), Aguiar and Hurst (2009) and Aguiar and Bils (2010)).

We can modify (9) to allow for systematic misreporting of total expenditure by the self employed (by a constant fraction across all categories) such that: $C_{ikt} = \kappa_k^C C_{ikt}^* \exp(v_{ikt})$ where C equals reported total expenditures, κ_W^C equals 1 for wage and salary workers ($k = W$), and κ_S^C different from 1 for the self employed ($k = S$). Assuming the parameters of (9) are otherwise similar between the self employed and wage and salary workers we can identify κ_S^C from:

$$\omega_{jikt} = \alpha + \mu_t + \beta_j \log C_{ikt}^* + \gamma D_S + \Theta' X_{ikt} + \varepsilon_{ikt} \quad (10)$$

where $\kappa_S^C = \exp(-\gamma / \beta_j)$. Underreporting all expenditures by a constant fraction κ_C of course leaves the share ω_{jikt} unchanged. The identification relies on the decreasing and increasing budget shares of various necessity and luxury good categories. This is similar to the idea underlying the procedure Aguiar and Bils (2010) uses to identify differential measurement error in expenditure reporting between the rich and poor using the CE data. Using this procedure, and aggregating the results across a large number of consumption categories, we find no evidence that the self employed systematically underreport their expenditures relative to wage and salary workers.³¹

The combined results of our two robustness exercises make us confident that any misreporting of expenditures is not biasing our results in any meaningful way.

6B. 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 desired expenditures due to differences in hours worked, 2) differences in desired saving propensities between the two groups, and 3) differences in the potential conceptual measures of income between the two groups.

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 is working full time. Depending on the sample, these differences in work hours average between 10 and 15 percent. 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

³¹ The complete estimation results can be found in the online robustness appendix that accompanies this paper.

of consumption.³² 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 the procedure similar to that of Blundell, Browning, and Meghir (1994) and add log work hours as a control variable in our estimation of (8). The results of including work hours directly into our regressions are shown in columns (1) and (2) of Table 6. Column (1) of Table 6 re-estimates (8) with the work hour controls using total family income as the income measure. Column (2) re-estimates (8) with the work hour controls using labor plus business income as the income measure. As seen from the results in columns (1) and (2), controlling for work hours does not change our conclusions in any way significant.

Differences in Saving Propensities

A potentially more significant problem is the extent to which the self employed have different saving preferences relative to wage and salary workers. If the self employed face more 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

³² Aguiar and Hurst (2009) show that non separable preferences between consumption and leisure are isomorphic to separable preferences between consumption and leisure with a home production sector that combines expenditures and time to produce the consumption good.

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

We can assess whether such possibilities affect the expenditure gap by segmenting the sample by age. In particular, 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, 2000). In columns (3) and (4) of Table 6 we re-estimate (8) 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. As seen from the table, the results with respect to the estimated amount of underreporting are 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.³³

As an additional robustness exercise, we examined the extent of underreporting by the level of household wealth. To do this, we augmented (8) 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 specification, available in the robustness appendix, shows that controlling for initial wealth does not undo the estimates of the underreporting nor do the estimates differ in any significant way across the wealth quartiles.

³³ 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.

Different Conceptual Measures of Income

Finally, we assess whether the above underreporting results could be driven by differences in the conceptual measures of income between the two types of individuals. We usually define income (net of taxes and transfers) as the economic resources that can be allocated to either consumption or savings. For wage and salary workers, the distinction is usually straight forward. Any dollar earned is given to the household and then the household decides how much to allocate to consumption and savings.³⁴ For the self employed, however, this distinction may be more complex. The self employed usually make a decision of how much of the return of the business to reinvest in the business (in the form of retained earnings) jointly with the decision of how much money they should withdrawal from the business in the form of self employed business income. The true income measure of the self employed should be any resources that they remove from their business to fund consumption and other types of saving as well as any retained earnings. Unfortunately in the CE and PSID we cannot observe any retained earnings.³⁵ 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 could explain the differential income-expenditure patterns documented above.

To test for the importance of the conceptual difference in income as an explanation of our results, 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 do not own business or own business with little value (see Hurst and Lusardi 2004). This is consistent with the facts documented in Hurst and Pugsley (2010) showing

³⁴ Fringe benefits such as pension contributions are an exception.

³⁵ Some household surveys, such as the Survey of Income and Program Participation (SIPP), include repeated observations on business capital that can be used to estimate retained earnings. Hamilton (2000), which measures self employed earnings from the SIPP, includes both the draw and an adjustment to account for retained earnings in the income measure. The SIPP however, does not include any information on expenditures.

that many of the self employed are in industries where very little physical capital is needed (skilled craftsmen, real estate agents, lawyers, etc.).

To test for the importance of potentially omitting retained earnings from our income measure, we re-estimate (8) including separate dummies for whether the business owner has business wealth equal to zero and whether the business owner has wealth above \$10,000. We can only do this for the PSID specification because the CE does not have a measure of business wealth. The PSID data includes measures of business wealth (and all wealth components) in survey years 1984, 1989, and 1994. As a result, we restrict our analysis to only the data from 1984 and 1994 given that the 1989 survey does not collect information on food expenditures. During these time periods, roughly one-third of the self employed had zero reported business wealth and fifty percent had business wealth more than \$10,000. The remaining had business wealth between \$0 and \$10,000. Although we estimate statistically significant difference in the estimated gaps between the two groups of self employed, the sizes of the gaps are limited (36 percent for the high business wealth group vs. 26 percent for the zero business wealth group). Furthermore, the estimated underreporting gap is still substantial 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.

7. Differences in Underreporting Across Different Groups and Tax Regimes

To bolster our case that households underreport their income to household surveys in a way similar to how they underreport their income to tax authorities, we explore whether factors that predict underreporting in the tax data are also found to predict underreporting within the household surveys. In particular, the literature exploring underreporting of income by the self employed to tax authorities has found that underreporting is higher when marginal tax rates are higher (Clotfelter, 1983) and that there is some variation in the extent of underreporting across the self employed in different industries. For example, Andreoni et al. (1998) finds that the self employed in professional industries underreport their

income the least while those who are shopkeepers and restaurateurs underestimate their income the most. We find suggestive evidence of similar patterns within the household surveys.

To get a better sense of the changing nature of the underreporting over time, we look at the estimated underreporting separately for each year. To do this, we re-estimate (8) but instead of putting in the self employment dummy alone we include the self employment dummy interacted with individual year dummies. In Figure 4, we show the results using non durable expenditure as our measure of expenditure and using total family income as our measure of income. As seen from the figure, underreporting was high during the early 1980s (prior to 1986), fell between the late 1980s and mid 1990s and then started increasing during the late 1990s and early 2000.

Does the year-to-year variation in the extent to which the self employed underreport their income correlate with prevailing tax regime within the U.S.? To explore this, we start by correlating the estimated coefficients of year dummies with measures of the aggregate tax regime. For our tax measure, we use the annual average marginal tax rate within the U.S. (inclusive of federal and state taxes) made available by the National Bureau of Economic Research.³⁶ For reference, we plot this series in Appendix Figure A1. It is well known that the average marginal tax rate was high during the early 1980s, fell by roughly 15 percent (roughly 5 percentage points) through the mid 1980s due to the 1981 and 1986 tax reductions, increased gradually during the 1990s, and then fell again with the tax reductions in the early 2000s. A simple correlation of the estimated underreporting by the self employed by year (as shown in Figure 4) with the average marginal tax rates within the U.S. (as shown in Figure A1) yields a correlation coefficient of 0.37 which is statistically significant. Most of this is driven by the high amount of underreporting prior to 1986, the sharp decline in underreporting between 1986 and the early 1990s, and the relative stability of the underreporting rate throughout most of the 1990s (aside from 1993 and 1994).

In Table 7, we explore this relationship more extensively. In this table, we show how underreporting is related to the average marginal tax rate within the U.S. along with various other

³⁶ See <http://www.nber.org/~taxsim/ally/>. The series was computed by Daniel Feenberg and is based on micro data file provide by the Statistics of Income Division of the IRS.

demographic controls. As above, we do this using nondurable expenditure from the CE (column 1), food expenditure from the CE (column 2), and food expenditure from the PSID (column 3). All specifications use log total family income as the income measure. To estimate these coefficients, we follow the same procedure discussed above for estimating the year dummies. In particular, we estimate (8) but add interactions between the self employment dummy and the relevant demographic variables. We then use the estimated coefficients on the interactions to compute the gap of underreporting, $1 - \kappa$. Again, asymptotic standard errors are reported in parentheses.

As seen from Table 7, very few variables enter significantly. Moreover, the standard errors are large in all the specifications, making it hard to rule out either large negative or large positive effects with most of the control variables. A few patterns, however, do emerge. First, even conditional on control variables, the average marginal tax rate is still a significant predictor of the extent of underreporting. Most of this identification comes from the fact that underreporting was much higher in pre-1986 when average marginal tax rates were higher. After 1986, the variation between the average marginal tax rate and the extent of income underreporting by the self employed was essentially zero. Given the power issues, the results are only suggestive. But, the results do show that underreporting was markedly higher prior to 1986 when average marginal tax rates were higher. This evidence is broadly consistent with the results using actual tax return data.

The only additional significant result that emerges from Table 7 is that those with advanced educational degrees are less likely to underreport their income. We do not have enough power to look at individual industries in isolation. However, most of the self employed individuals with advanced degrees are professionals such as doctors, dentists, lawyers, and accountants. In all of these professions, the self employed may find it more difficult to hide their income given that there is usually a detailed paper trail on the other side of the transaction. This is particularly true for doctors, dentists, and other medical professions given that they often bill insurance companies for their services rendered. The point estimates from columns 1 and 3 suggest that those with professional degrees only underreport their income by 12 percent and 23 percent, respectively, which represents a significant reduction relative to households with

a high school degree or less. Again, the evidence is broadly consistent with the literature using tax filings that find professionals are less likely to underreport their income.

8. 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 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.³⁷ 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.

8A. Differences in Earnings 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 carefully constructs lifecycle wage profiles separately for wage and salary workers and for the self employed under several alternative measures of self employed income.³⁸ He finds the median self employed individual's wages, cumulating over 10 years, are 35 percent lower than a comparable wage and salary worker. He assigns a prominent role to non pecuniary returns to self employment over alternative explanations.³⁹

³⁷ A full discussion, including details on specifications and sample restrictions can be found in the online robustness that accompanies this paper.

³⁸ His favored measure adjusts the draw from the business to account for retained earnings.

³⁹ Using the panel dimension of the SIPP Hamilton (2000) also supplies convincing evidence that the gap cannot be explained by selection, finding that the self employed earned higher than average wages before transitioning to self employment.

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 formally 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 to 30 percent. This difference is only mitigated slightly once we control for differences in work hours between the two groups. 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.⁴⁰

8B. Difference in Wealth Holdings Between the Self Employed and Wage and Salary Workers

There is a separate literature that focuses on differences in wealth between the self employed and wage and salary workers. Most of this work looks to examine the importance of financial frictions facing the self employed. The argument goes that because the self employed face binding liquidity constraints they must accumulate additional wealth to self finance their entrepreneurial projects. For example, both Gentry and Hubbard (2004) and Cagetti and De Nardi (2006) document that the self employed hold a disproportionate amount of the aggregate wealth within the U.S.. Gentry and Hubbard (2004) goes further and shows that even conditional on income, the self employed hold higher wealth than wage and salary workers. Cagetti and De Nardi (2006) uses such relationships to calibrate a model of financial constraints facing business owners to explain the skewed wealth distribution within the U.S..

The differential relationship between wealth conditional on income between the self employed and wage and salary workers depends on the extent to which the self employed income is reported without error. To the extent that the self employed underreport their income, the conditional wealth holdings of the self employed relative to wage and salary workers will be overstated.

⁴⁰ 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.

To illustrate this point, we use data from the 1984, 1989, and 1994 PSID. Our sample is otherwise the same as we used above. Using this data, we regress log total wealth (inclusive of home equity) on a self employment dummy, current household income, and the demographic variables included above when estimating equations (1) and (2). With this specification, we exclude those households with zero or negative wealth. Similar to the results documented by others in the literature, we find that the net worth of the self employed is 93 log points higher than wage and salary workers (a coefficient on the self employment dummy of 0.93 with a standard error of 0.05). If, however, we adjust for the potential misreporting of income by the self employed, the estimated wealth differences fall dramatically. To do this, we simply add 30 percent to the current income reports of the self employed and then re-estimate the above wealth equation. Modifying the income of the self employed reduced the estimated wealth gap to 63 log points (standard error = 5 log points). As a result, roughly one third of the estimated wealth differences between the self employed is simply due to the underreporting of income by the self employed. The results are nearly identical if we estimate the regression in levels and include households with non-positive wealth.

8C. The Empirical Importance of Precautionary Savings

The first two examples show 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) uses 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 their same specification, 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.⁴¹ However, if we increase the income of the self employed by 30 percent to adjust for the underreporting, the estimated importance of precautionary savings is 13.5 percent lower (from 47.5 percent to 41.1 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 substantial upward biases in the importance of precautionary savings in explaining total wealth accumulation of younger households.

8D. Adjustments to Life Cycle Earnings Profiles

Self employment probabilities change over the lifecycle. For example, within our PSID, CE and CPS sample, the fraction of individuals that are self employed rises from 5.0 percent, 5.9 percent and 5.1 percent respectively when households are 25 to 19.1 percent, 11.2 percent and 16.5 percent respectively when households are 45. They rise further to 39.5 percent, 21.6 percent and 28.3 percent when households are 65.⁴² 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.

⁴¹ The details of the specification and sample can be found in the online robustness appendix.

⁴² We extend our sample to include households up to the age of 65 as long as they meet our other sample criteria. The number of households whose head's age is close to 65 is small in both the PSID (n=2,597) and the CE sample (n=120). However our estimates are broadly consistent with the more precisely estimated versions from the CPS where the number of households at age 65 is large (n=2,831).

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, CE and CPS samples, respectively, average earnings fell by 27.3 percent, 21.5 percent and 14.3 percent between the ages of 45 and 65.⁴³ Simply adjusting for underreporting reduces these declines by 1.4, 3.2, and 3.6 percentage points. Taken together, between 5 and 10 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.

8E. 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 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

⁴³ 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.

dummies. We then compare these residuals, with and without adjusting for the underreporting of income by the self employed, across MSAs.

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 30 percent, increases the Nassau-Detroit earnings differential by nearly 20 percent (from 8.9 percent to 10.6 percent). A similar comparison can be made between San Francisco (where 16.2 percent of individuals are self employed) and Buffalo (where 10.2 percent of individuals are self employed). Adjusting for the underreporting of income of the self employed increases the San-Francisco-Buffalo earnings gap by 6.7 percent (from 23.9 percent to 25.5 percent). The earnings differential between the Barnstable-Yarmouth, Massachusetts MSA (which has the highest self employment rate among MSAs at 25.5 percent) and the Kokomo, Indiana MSA (which has the lowest self employment rate among MSAs at 6.0 percent) gets reduced by 35 percent when adjusting for the underreporting of income by the self employed (from -14.7 percent to -9.6 percent).

8F. 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.

9. 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 the extent to which the self employed also 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 self employed underreport their income by about 30 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. Consistent with data from the tax audits, we also find that the amount of underreporting was higher when marginal tax rates were higher and was lower among those with an advanced degree.

The results in this paper have implications for many literatures that use the income data reported in U.S. household surveys. In particular, we show that research that compares the earnings or saving behavior between the self employed and wage and salary workers will yield substantially biased results if the research does not account for the underreporting of income by the self employed. Moreover, we show that in many other settings, failing to account for the underreporting of income by the self employed can lead to quantitatively biased conclusions. In particular, we show this to be true with respect to studies of the importance of precautionary savings, lifecycle earnings profiles, and differences in earnings across MSAs.

The results in this paper also point to a potentially bigger concern. There is an active literature showing that individuals will 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 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 benefits of providing accurate 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 would suggest 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.

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, Brian Erard, and Jonathan Feinstein (1998). "Tax Compliance," *Journal of Economic Literature*, 36(2), pp. 818-60.
- Banks, James, Richard Blundell, and Arthur Lewbel (1997). "Quadratic Engel Curves and Consumer Demand," *The Review of Economics and Statistics*, 79(4), pp. 527-39.
- 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.
- 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, 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 *Handbook of Econometrics*, Volume 5, eds. J.J. Heckman and E. Leamer, Elsevier Science.
- Cagetti, Marco, and De Nardi, Mariacristina (2006). "Entrepreneurship, Frictions, and Wealth", *Journal of Political Economy*, vol. 106(5), pp. 835-70.
- Carroll, Chris and Samwick, Andrew (1998). "How Important is Precautionary Saving?" *Review of Economics and Statistics*, vol. 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 Microeconomic 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.
- Feldman, Naomi and Joel Slemrod (2007). "Estimating Tax Noncompliance with Evidence from Unaudited Tax Returns", *Economic Journal*, vol. 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), 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.

- Gorodnichenko, Yuriy, Jorge Martinez-Vazquez, and Klara Sabirianova Peter (2009). "Myth and Reality of Flat Tax Reform: Micro Estimates of Tax Evasion Response and Welfare Effects in Russia," *Journal of Political Economy*, 117(3), pp. 504-554.
- Gourinchas, Pierre-Olivier and Jonathan Parker (2002), "Consumption over the Life Cycle," *Econometrica*, January 2002, 70, 1, 47-89.
- Haider, Steven and Gary Solon (2006). "Life-Cycle Variation in the Association Between Current and Lifetime Earnings". *American Economic Review*, 96(4), 1308-1320
- Hamilton, Barton H. (2000). "Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment." *Journal of Political Economy* vol. 108 (3), pp. 604-31.
- Hurst, Erik and Lusardi, Annamaria (2004). "Liquidity Constraints, Household Wealth, and Entrepreneurship", *Journal of Political Economy*, vol. 112(2), pp. 319-47.
- Hurst, Erik and Ben Pugsley (2010). "The Non Pecuniary Benefits of Small Business Ownership", University of Chicago Working Paper.
- Johansson, Edvard (2005). "An Estimate of Self Employment Income Underreporting in Finland", *Nordic Journal of Political Economy*, vol 31(1), pp. 99 - 109.
- 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 Vissing-Jorgensen, Annette (2002). "The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?" *American Economic Review* vol. 92(4), pp. 745-78.
- Pissarides, Christopher and Guglielmo Weber (1989). "An Expenditure Based Estimate of Britain's Black Economy", *Journal of Public Economics*, vol. 39(1), pp. 17-32.
- Schuetze, Herb (2002). "Profiles of Tax Noncompliance Among the Self Employed in Canada:1969-1992." *Canadian Public Policy*, vol. 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.

Zeldes, Steve (1989). "Consumption and Liquidity Constraints: An Empirical Investigation", *Journal of Political Economy*, vol. 97(2), pp. 305-46.

Data Appendix

This appendix gives background information on all the data sources used in this paper.

CE Data

The CE is conducted by the Bureau of Labor Statistics. The primary objective of the survey is to collect detailed household expenditure information to update the weights used to construct the Consumer Price Index. However, the household level data of the CE have also been used extensively in academic research to explore household consumption behavior. The CE is a quarterly survey and its modern version started in 1980. In addition to detailed information on household expenditures, the CE collects key information on household demographics, income, and employment status. The CE interviews a consumer unit up to five times, once every 3 months. The first interview is conducted to establish contact with, and collect such data on, the interviewees; the subsequent four interviews are carried out to collect most of the expenditure data. In the fifth interview, demographic, income, and labor supply variables are updated. After the fifth interview, the consumer unit leaves the sample and new units are added to the sample, giving the CE a rotating sample structure.

PSID Data

The PSID is a longitudinal survey conducted by the Institute of Social Research at the University of Michigan. The PSID collects extensive information on employment history, demographic characteristics, and household income.

CPS Sample

To match the PSID and CE sample selection criteria, we restrict our analysis to households with a civilian male household head, whose age is between 25 and 55 (inclusive), who worked at least 40 weeks during the last calendar year and usually worked at least 30 hours per week in paid or self employment. We further restrict the sample by excluding households with a male wage and salary worker, but whose spouse (if present) was self employed. Lastly, we exclude any remaining households who had non-

positive labor plus business family income or who had any farm income. We define two measures of household income within the CPS: family labor earnings plus family business income and total family income. The two income measures are defined analogously to the way they were defined in the CE and the PSID. Finally, self employed households are defined as those households whose male head reports that for his longest job during the last year he was self employed.⁴⁴ Over the sample period of 1980 to 2003 applying these restrictions yields a sample of 511,306 households, with 70,423 of those classified as self employed.

⁴⁴ To determine the class of employment, the CPS interviewer asks a series of questions about the individual's employer to determine whether he worked for a private firm, for a government (federal, state, or local), worked without pay in family business, or were self employed (incorporated or unincorporated). Since we have already excluded those individuals working without pay, wage and salary workers are defined as those working for a private firm or for a government.

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

Income Measure/Sample	Data Set					
	CE		PSID		CPS	
	Self Employed	Wage Workers	Self Employed	Wage Workers	Self Employed	Wage Worker
Panel A: Income Measures						
Total Family Income	10.87	10.86	11.20	11.00	10.97	10.93
Labor Plus Business Income	10.81	10.84	11.02	10.91	10.87	10.88
Panel B: Expenditure Measures						
Food Expenditure	8.85	8.67	8.97	8.73	-	-
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	40.7	37.9	41.7	39.2
% Education = 12	26.9	29.2	20.5	24.0	29.7	33.8
% Education = Some College	23.2	23.3	31.2	31.2	22.7	23.8
% Education = College	20.3	18.9	21.4	19.2	20.1	18.7
% Education = Graduate School	17.9	12.5	14.9	9.9	17.8	11.6
% Black	3.4	7.8	2.7	9.7	3.2	8.9
% Married	83.2	80.6	86.7	78.6	86.1	81.8
Average Family Size	3.4	3.26	3.30	3.10	3.36	3.22
Average Head Work Hours	2,561	2,317	2,560	2,293	2526	2280
Sample Size	2,508	24,704	4,207	29,639	70,423	440,883

Notes: This table compares income, expenditure, and demographic variables between the self employed and wage and salary workers within the CE, PSID, and CPS. Each sample is restricted to include all households with male heads between the ages of 25 and 55 (inclusive) who are working full time. Additional sample restrictions for each survey are discussed in the text.

Table 2: Conditional Log Income Differences for Full Time Wage and Salary Workers and Self Employed

Variables	Data Sets/Specification					
	I. Income = Total Family Income			II. Income = Labor Earnings + Business Income		
	(1)	(2)	(3)	(4)	(5)	(6)
Dummy: Self Employed	-0.08 (0.02)	0.07 (0.01)	-0.07 (0.01)	-0.10 (0.02)	-0.01 (0.01)	-0.11 (0.01)
Demographic Controls Included	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies Included	Yes	Yes	Yes	Yes	Yes	Yes
Sample	CE	PSID	CPS	CE	PSID	CPS

Notes: The table shows the regression of household income on a self employment dummy and additional demographic and year controls. In panel I, our measure of income is total family income for the household. In panel II, our measure of income is labor earnings of the household plus household business income. We run the regressions separately within each of the three surveys. For example, columns (1) - (3) show the results from the same regression, but using data from the CE, PSID, and CPS surveys, respectively. All regressions control for: a series of five year age dummies, four dummies indicating the education of the household head (exactly high school, some college, exactly college, and more than college; less than high school was omitted), a dummy if the household head was Black, a dummy if the household head was married, and a series of family size dummies. The CE, PSID, and CSP samples are the same as the ones defined in Table 1. Robust standard errors are in parenthesis.

Table 3: Conditional Log Expenditure Differences For Full Time Wage and Salary Workers and Self Employed

Variables	Data Sets/Dependent Variable			
	I. CE			II. PSID
	Food	Nondurable	Total	Food
Dummy: Self Employed	0.12 (0.01)	0.13 (0.01)	0.16 (0.01)	0.14 (0.01)
Demographic Controls Included	Yes	Yes	Yes	Yes
Year Dummies Included	Yes	Yes	Yes	Yes

Notes: The table shows the regression of household expenditure on a self employment dummy and additional demographic and year controls. In panel I, we show the estimates using CE data. In panel II, we show the estimates using PSID data. For the CE sample, we explore conditional differences between the self employed and the wage and salary workers using three measures of expenditures: food (column 1), nondurable expenditures (column 2), and total expenditures (column 3). The PSID sample has information only on food expenditures. All regressions control for: a series of five year age dummies, four dummies indicating the education of the household head (exactly high school, some college, exactly college, and more than college; less than high school was omitted), a dummy if the household head was black, a dummy if the household head was married, and a series of family size dummies. The CE and PSID samples are the same as the ones defined in Table 1. Robust standard errors are in parenthesis.

**Table 4: Log Expenditure Differences Between Self Employed and Wage Workers,
Conditional on Log Income and Demographics**

Sample/Expenditure Measure	I. Income = Total Family Income		II. Income = Labor Earnings Plus Business Income	
	OLS	IV	OLS	IV
	Panel A: Estimates of β			
CEX, Nondurable Expenditure	0.37 (0.01)	0.60 (0.01)	0.34 (0.01)	0.61 (0.01)
CEX, Total Expenditure	0.42 (0.01)	0.76 (0.01)	0.39 (0.01)	0.76 (0.01)
CEX, Food Expenditure	0.26 (0.01)	0.40 (0.01)	0.25 (0.01)	0.40 (0.01)
PSID, Food Expenditure	0.30 (0.01)	0.32 (0.01)	0.29 (0.01)	0.31 (0.01)
Panel B: Estimates of γ				
CEX, Nondurable Expenditure	0.17 (0.01)	0.18 (0.01)	0.18 (0.01)	0.19 (0.01)
CEX, Total Expenditure	0.20 (0.01)	0.21 (0.01)	0.21 (0.01)	0.23 (0.02)
CEX, Food Expenditure	0.15 (0.01)	0.15 (0.01)	0.15 (0.01)	0.16 (0.01)
PSID, Food Expenditure	0.12 (0.01)	0.11 (0.01)	0.14 (0.01)	0.13 (0.01)

Notes: This table shows the OLS and IV estimates of equation (8) from the text. Equation (8) estimates the relationship between log household expenditures and log household income conditional on a self employment dummy and other demographics. The coefficient on log household income is denoted as β and the coefficient on the self employment dummy is denoted γ . The demographic controls are the same as the controls used to estimate the results in Tables 2 and 3, aside from the education dummies. We use the education dummies as the instruments for income in the IV specifications. As with the results in Tables 2 and 3, 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.

**Table 5: Predicted Fraction of Income Underreported by Self Employed,
Baseline Specification**

Expenditure Measure	Estimate of $(1-\kappa)$, in Percent	
	Income = Total Family Income	Income = Labor Earnings Plus Business Income
CEX, Nondurable Expenditure	25.5 (1.5)	27.5 (1.6)
CEX, Total Expenditure	24.5 (1.5)	26.5 (1.5)
CEX, Food Expenditure	31.7 (2.0)	33.5 (2.0)
PSID, Food Expenditure	30.1 (2.2)	34.9 (1.9)

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 4. We show the different estimates of $(1-\kappa)$ using different measures of expenditure, using data from the different surveys, and using different measures of income. In this table, we only show the estimates using the β and γ estimates from the IV specifications. Asymptotic standard errors are computed for the transformation $1-\kappa$ using a robust estimate of the variance covariance matrix.

Table 6: Predicted Fraction of Income Underreported by Self Employed, Alternate Specifications

Expenditure Measure	Specification/Sample			
	1	2	3	4
CEX, Nondurable Expenditure	25.2 (1.5)	27.4 (1.6)	26.6 (2.7)	24.4 (1.8)
CEX, Total Expenditure	24.5 (1.5)	26.7 (1.5)	24.2 (2.6)	24.6 (1.7)
CEX, Food Expenditure	31.5 (2.0)	33.5 (1.9)	35.7 (3.7)	28.6 (2.2)
PSID, Food Expenditure	28.6 (2.2)	34.1 (1.8)	30.1 (3.0)	28.4 (3.0)
Income Measure	Total	Labor Plus Business	Total	Total
Specification	Include Log Work Hours as a Control	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- κ) for alternate specifications of equation (8) from the text. All specifications in this table are based on the IV specification using CEX data with total nondurable expenditures as the expenditure measure. As a result, these estimates should be compared to row (1) of Table 5. In columns (1) and (2), we add log work hours to our vector or control variables. The first column uses total family income as the income measure while the second column uses labor plus business income as the income measure. In columns (3) and (4), we split the samples by age (young vs. old). In columns (5) and (6), we only use data from the 1984, 1989, and 1994 PSID because they are the only years that collect information on household business wealth. In column (5), we only include the self employed with zero business wealth. In column (6), we only include the self employed with more than \$10,000 in business wealth. For all specifications, the sample is otherwise the same as the one described in the note to Table 1. Asymptotic standard errors are computed for the transformation 1- κ using a robust estimate of the variance covariance matrix.

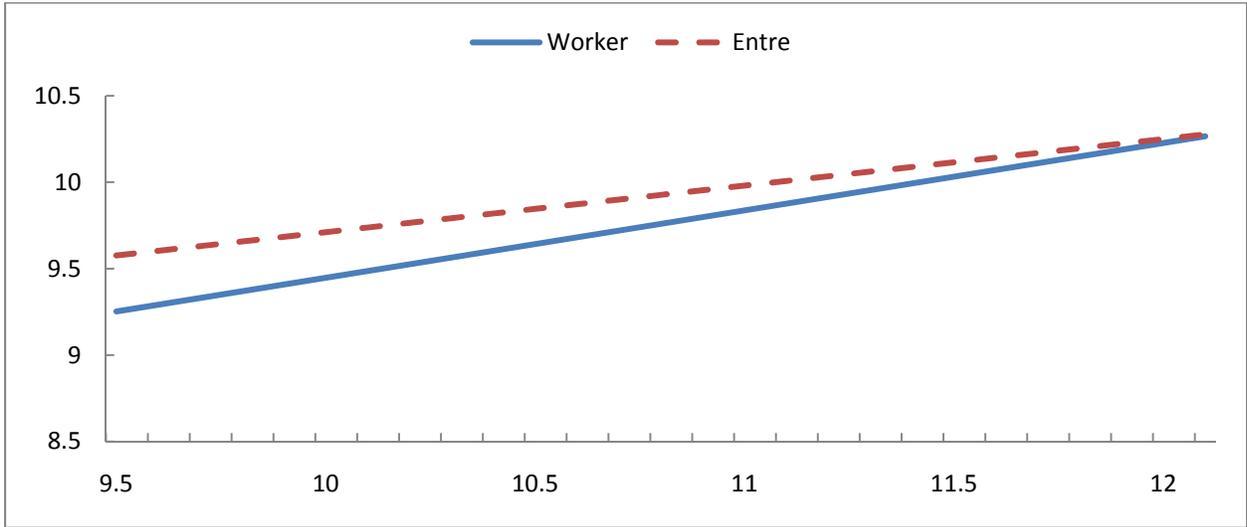
Table 7: Explaining the Income Gap Among the Self Employed, Percentage Changes to $(1-\kappa)$

Variable	I. CE		II. PSID
	(1) Nondurable Expenditure	(2) Food Expenditure	(3) Food Expenditure
Average marginal federal tax rate (percent)	2.24 (0.89)	2.76 (1.18)	1.56 (0.84)
Age: 40-55	1.91 (3.86)	0.62 (5.32)	3.45 (4.63)
Ed = Some College	4.94 (5.15)	6.04 (6.72)	14.51 (4.74)
Ed = College	-5.62 (5.08)	-3.03 (6.81)	2.20 (6.16)
Ed = More than College	-11.48 (6.00)	-2.89 (7.47)	-23.03 (9.74)
Black	10.39 (16.30)	2.98 (18.43)	6.03 (11.25)
Married	-8.19 (5.47)	-21.05 (8.63)	-4.11 (9.39)
Income Measure	Total Family	Total Family	Total Family

Note: This table shows how the estimates of $(1-\kappa)$ change with economic and demographic variables. To get these estimates, we re-estimate equation (8) with the relevant variable of interest interacted with a self employment dummy. To get this specification, we also included the average marginal federal tax rate as a separate control. The results in this table are based on the IV specification using the relevant specifications. All specifications use total family income as the income measure. Panel I shows the results using CE data while panel II shows the results using PSID data. The samples are the same as the ones used in Table 1. Asymptotic standard errors are computed for the transformation $1-\kappa$ using a robust estimate of the variance covariance matrix.

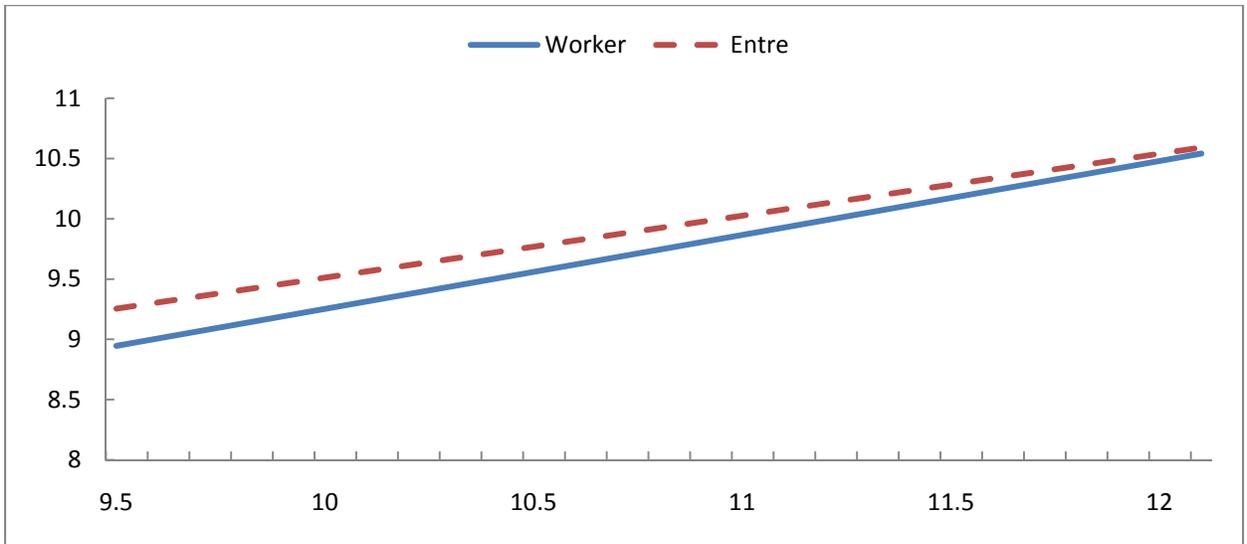
Figure 1: Relationship Between Total Family Money and Nondurable Expenditure, CE

Panel A: OLS Estimation



Notes: This figure shows the OLS estimates of the Engel curves as specified in equation (3) in the text, separately for wage/salary workers (solid line) and the self employed (dashed line). We evaluate both curves at the sample means of the demographic control for the pooled sample. The estimated intercepts for the self employed and the wage/salary workers, respectively, are 7.02 and 5.55. The estimated slopes for each group, respectively, are 0.27 and 0.39. The sample for each specification are the same as the ones described in the notes to Table 1.

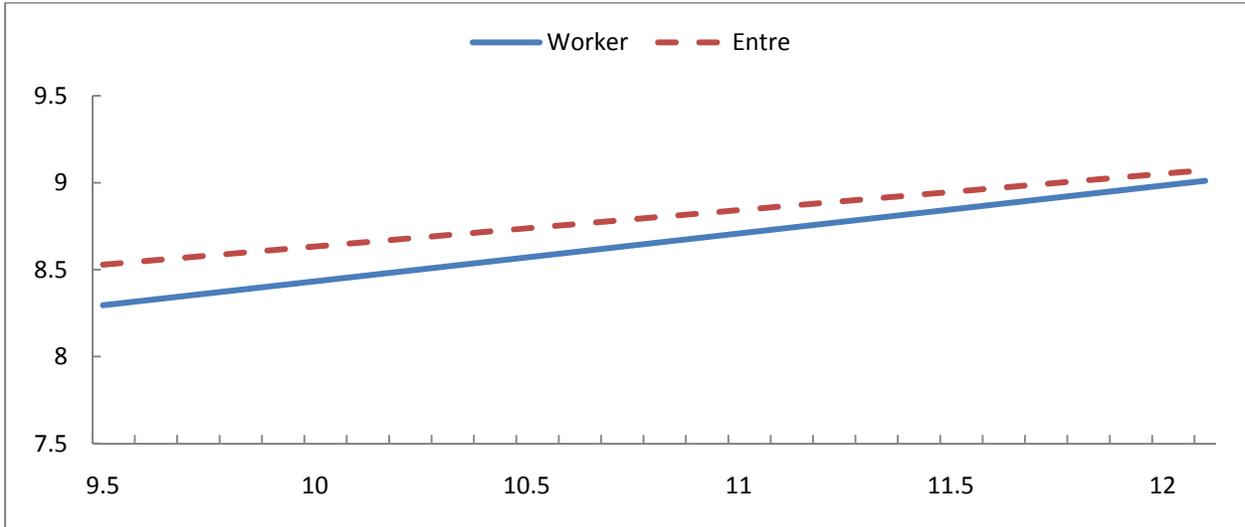
Panel B: IV Estimation



Notes: This figure shows the OLS estimates of the Engel curves as specified in equation (3) in the text, separately for wage/salary workers (solid line) and the self employed (dashed line). We evaluate both curves at the sample means of the demographic control for the pooled sample. The estimated intercepts for the self employed and the wage/salary workers, respectively, are 4.37 and 3.12. The estimated slopes for each group, respectively, are 0.51 and 0.61. The sample for each specification are the same as the ones described in the notes to Table 1.

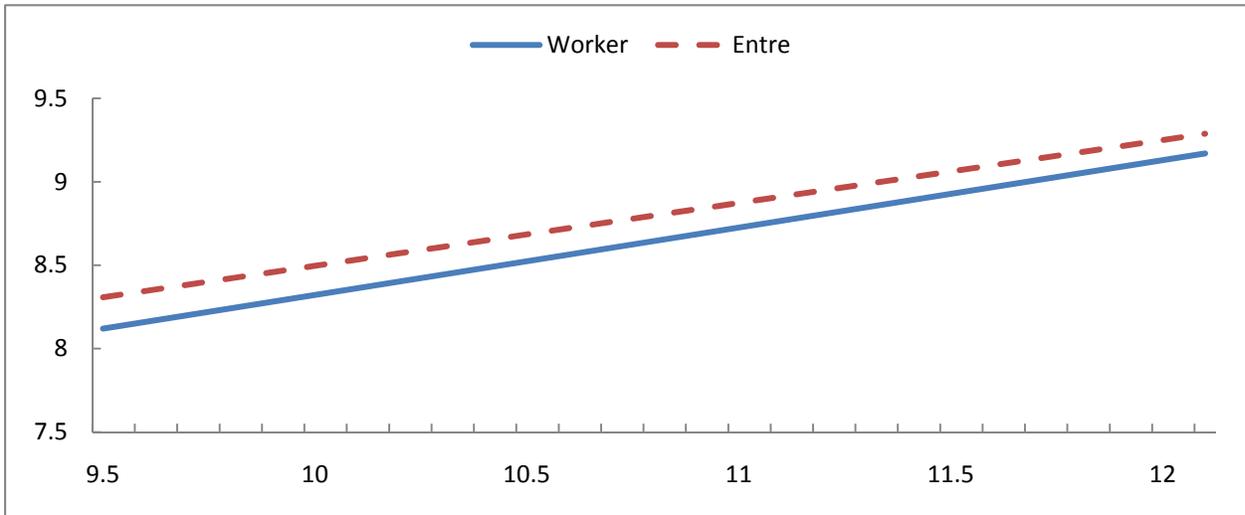
Figure 2: Relationship Between Total Family Money and Food Expenditure, CE

Panel A: OLS Estimation



Notes: This figure shows the OLS estimates of the Engel curves as specified in equation (3) in the text, separately for wage/salary workers (solid line) and the self employed (dashed line). We evaluate both curves at the sample means of the demographic control for the pooled sample. The estimated intercepts for the self employed and the wage/salary workers, respectively, are 6.54 and 5.68. The estimated slopes for each group, respectively, are 0.21 and 0.28. The sample for each specification are the same as the ones described in the notes to Table 1.

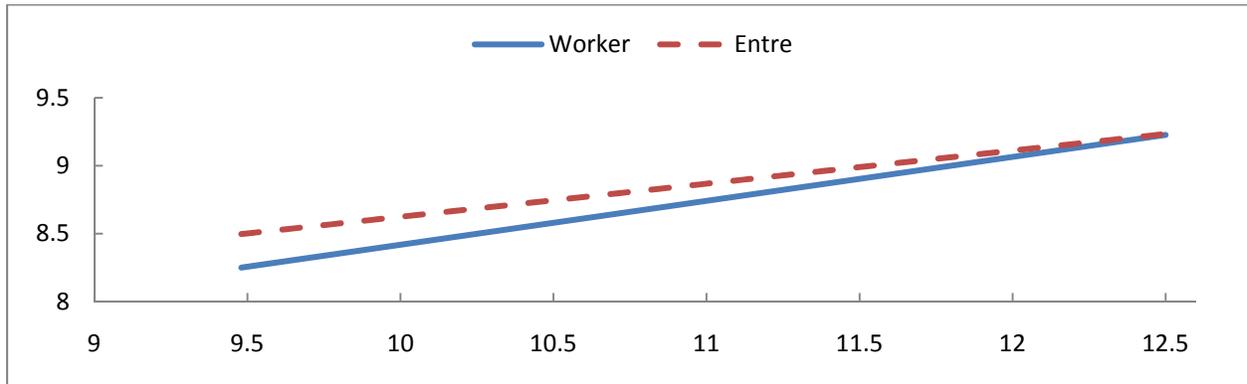
Panel B: IV Estimation



Notes: This figure shows the OLS estimates of the Engel curves as specified in equation (3) in the text, separately for wage/salary workers (solid line) and the self employed (dashed line). We evaluate both curves at the sample means of the demographic control for the pooled sample. The estimated intercepts for the self employed and the wage/salary workers, respectively, are 4.73 and 4.29. The estimated slopes for each group, respectively, are 0.38 and 0.40. The sample for each specification are the same as the ones described in the notes to Table 1.

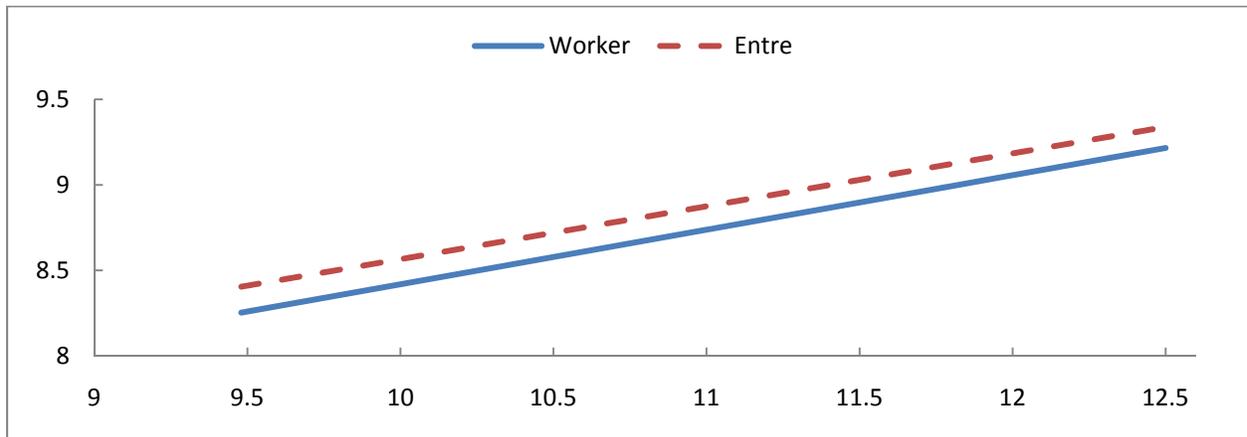
Figure 3: Relationship Between Total Family Money and Food Expenditure, PSID

Panel A: OLS Estimation



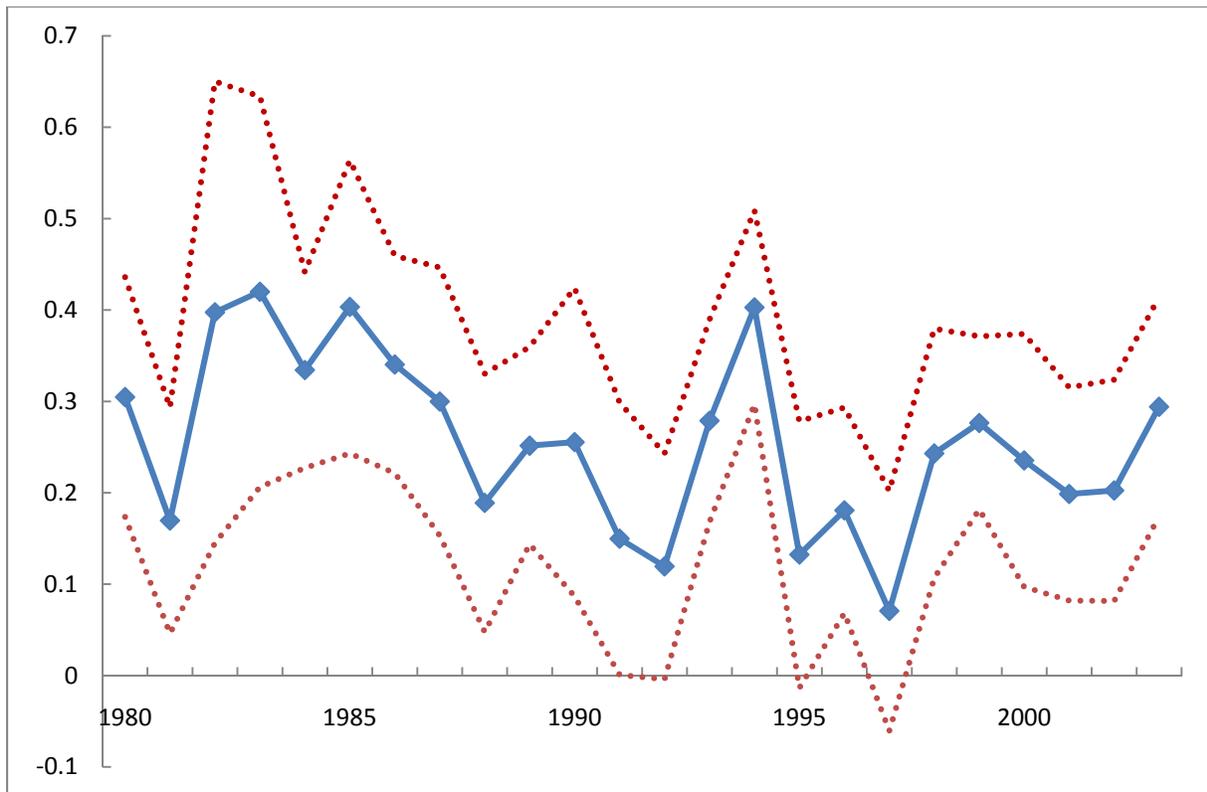
Notes: This figure shows the OLS estimates of the Engel curves as specified in equation (3) in the text, separately for wage/salary workers (solid line) and the self employed (dashed line). We evaluate both curves at the sample means of the demographic control for the pooled sample. The estimated intercepts for the self employed and the wage/salary workers, respectively, are 6.18 and 5.18. The estimated slopes for each group, respectively, are 0.24 and 0.32. The sample for each specification are the same as the ones described in the notes to Table 1.

Panel B: IV Estimation



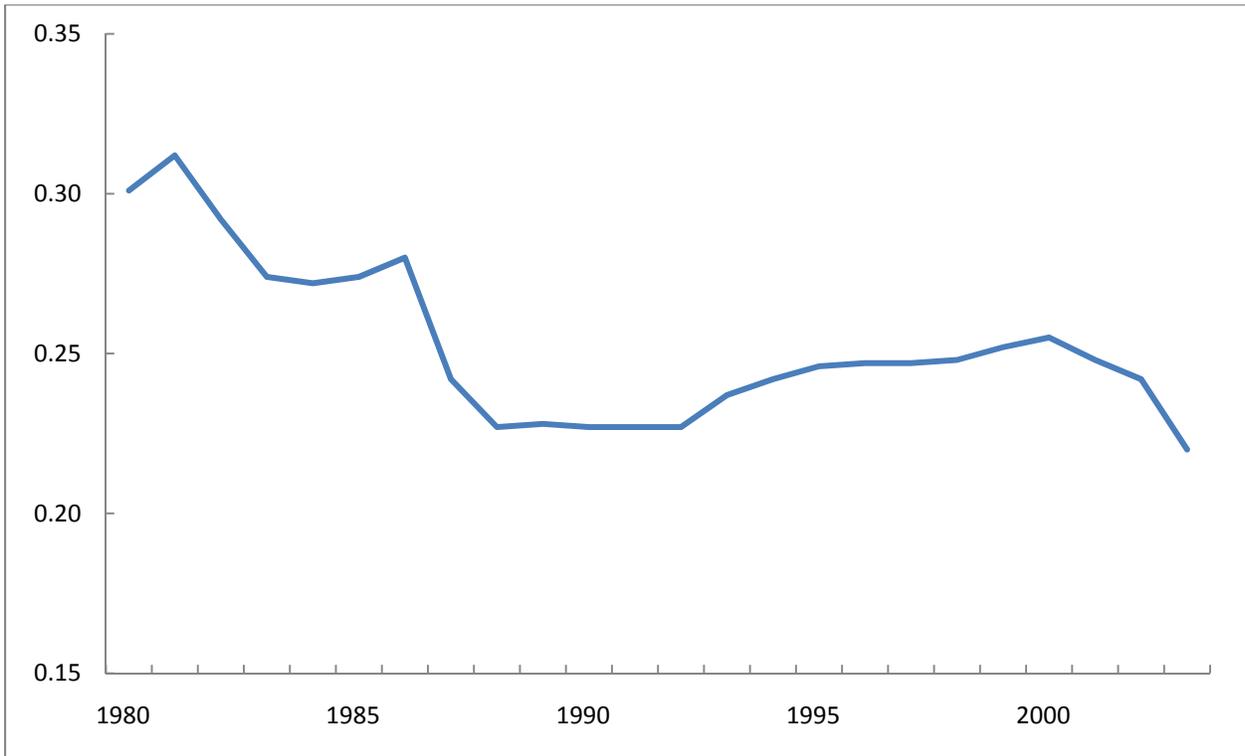
Notes: This figure shows the OLS estimates of the Engel curves as specified in equation (3) in the text, separately for wage/salary workers (solid line) and the self employed (dashed line). We evaluate both curves at the sample means of the demographic control for the pooled sample. The estimated intercepts for the self employed and the wage/salary workers, respectively, are 5.47 and 5.23. The estimated slopes for each group, respectively, are 0.31 and 0.32. The sample for each specification are the same as the ones described in the notes to Table 1.

Figure 4: Estimates of the Underreporting of Income By Self Employed Over Time, CEX Data



Notes: This table shows the estimates of underreporting ($1-\kappa$) by year using the CEX data (IV specification) where nondurable income is the expenditure measure and total family money is the income measure. To get these estimates, we re-estimated equation (8) including year dummies interacted with the self employed dummy. Robust standard error bands are shown as the dashed lines.

Appendix Figure A1 - The Average Marginal Tax Rate in the U.S.



Notes: Figure shows the time series variation in the average marginal tax rate. See text for details.