Profiles of Tax Noncompliance Among the Self-Employed in Canada: 1969-1992

Herb J. Schuetze*
University of Victoria
PO Box 1700 STN CSC, Victoria BC
E-mail: hschuetz@uvic.ca

Abstract:

This paper utilizes an expenditure based approach to provide estimates of the degree of income tax noncompliance by Canadian self-employed households using a series of Canadian Family Expenditure Surveys from 1969 to 1992. These estimates are disaggregated across years, demographic characteristics and occupation to shed some light on the determinants of such activities and to provide guidance to tax enforcement policy makers. The findings suggest that the degree of noncompliance by the self-employed varies significantly by occupation, age and the number of household members self-employed.

* I am grateful to Tom Crossley, Peter Kuhn, Lonnie Magee, Les Robb and Jeffrey Smith who provided helpful advice and comments on earlier drafts. I also thank four anonymous referees for insightful comments. Funding was provided in the early stages by the Canadian International Labour Network (CILN).

(1) Introduction

Because of the absence of a third party reporting income, the self-employed are often able to conceal income from the tax authorities. Public finance economists argue that the presence of such "underground" activity is economically distortionary and leads to an inequitable distribution of the tax burden (Auerbach and Feldstein 1998 and Slemrod 1992). The distribution of the tax burden is shifted away from individuals who have the ability and choose not to comply with the tax rules toward those who have little choice but to comply. These distortions have become increasingly important in light of the fact that self-employment as a share of overall labour market activity has grown in Canada. Between 1981 and 1998 the percentage share of self-employment in total employment increased from 12.8 percent to 17.1 percent. Nearly 2.5 million Canadians reported working for themselves by 1998 compared to just 1.4 million in 1981 -- a 68 percent increase. Over the same period the number of wage employees in Canada grew by only 19 percent[1]. While still significant, the contributions of self-employment to overall declared employment incomes in Canada did not match the contributions entrepreneurs made to total employment. In 1998, the most recent year for which there is data, 7.5 percent of total declared employment income or 33 billion dollars in income was derived from self-employment[2].

In previous work (Schuetze 2000), I presented evidence that suggests that one of the main reasons for the growth in self-employment in Canada between 1983-1994 is the concurrent rise in income taxes which made self-employment more attractive to potential entrepreneurs. To the extent that these potential entrepreneurs are attracted to self-employment by the relative ease of concealing income, this result suggests that the *incidence* of tax noncompliance has risen in Canada. However, very little is known about the *extent* to which the self-employed conceal income in Canada. Even

less is known about the factors which influence the degree of income concealing among the self-employed. Previous research has been limited to "aggregate" estimates of income tax noncompliance that do not give much of an indication of the determinants of such activity (Mirus and Smith 1996) and , in some cases, do not allow one to focus on the self-employed (Spiro 1997, Mirus, Smith and Karoleff 1994, and Hill and Kabir 1996). In addition, those studies which do provide estimates over time are subject to large measurement error that make time trend analysis difficult to interpret.

In this paper, the limitations of previous research are overcome by utilizing an expenditure based approach developed by Pissarides and Weber (1989) (referred to as P&W throughout this paper) to estimate the degree of noncompliance by the self-employed. The data used are taken from a series of Canadian Family Expenditure Surveys for six years between 1969 and 1992, which allow one to estimate noncompliance across various dimensions. Essentially, this approach uses estimates of the relationship between food consumption and income obtained from a sample of wage and salary earners, who are assumed to report income correctly, in order to impute "true" income levels for the self-employed. When compared to reported incomes these imputed incomes provide an estimate of the degree of noncompliance among the self-employed. Unlike other measures of noncompliance, the estimates obtained using this procedure take into account both under-reporting of gross income and overstatement of expenses and deductions (see section 3 for a discussion).

I obtain estimates of the degree of noncompliance by households obtaining a significant amount of income from self-employment across years, demographic characteristics and occupations. By disaggregating across these dimensions I hope to achieve two goals. The first goal is to identify the determinants of income tax noncompliance by self-employed households in Canada. The

characteristics of noncompliance among the self-employed might provide clues to the reasons for this activity. For instance, one might expect that the amount of income concealed by entrepreneurs varies with changes in the tax environment and, therefore, also may vary with time. Put simply, an increase in the marginal benefit of concealing self-employment income (the effective marginal tax rate) is likely to lead to an increase in noncompliance. Further, one's opportunity to under-report income from self-employment is likely to vary by occupation. Individuals who provide products or services in occupations which frequently involve cash transactions have greater opportunity to under-report income.

The second is to utilize those determinants which may be observed by tax authorities to provide guidance to tax enforcement policy makers in designing effective tax policy schemes. Tax enforcement agents often have difficulty enforcing tax laws because of an imperfect ability to monitor tax compliance. One tool used by tax authorities to enforce compliance is a taxation audit scheme. By identifying the observed patterns of tax noncompliance across characteristics which are commonly included on income tax returns (occupation, age of tax-filer, and number of household adults self-employed) this paper will provide guidelines to aid the tax authority in enforcing income tax compliance. Audits can be directed towards those who are most likely to be in noncompliance with the income tax code. The results of such audits could then be used to identify the source of noncompliance – overstated expenses or understated income – which would aid in designing tax policy focussed on reducing this activity.

My main findings are as follows. The estimates obtained in this study suggest that households obtaining 30 percent or more of household income from self-employment concealed on average between 11 and 23 percent of total household income between 1969 and 1992. There

appears to be no clear time trend in the degree of income concealing by self-employed households over this period and no clear pattern of non-compliance arises when looking across the household head's level of education. With regard to the insights provided by looking across characteristics observed by the tax authority, there is a great deal of variation in the degree of noncompliance among the self-employed across these dimensions. For example, I find evidence that the degree of noncompliance varies significantly by the occupation of the entrepreneur. This suggests that one's "opportunity" to under-report income may be a determinant of the degree of noncompliance among entrepreneurs. Perhaps not surprisingly, noncompliance among self-employed households is highest among those in occupations such as construction and services which commonly involve cash transactions that allow income under-reporting with relative ease. On the other hand, households with members engaged in self-employment activities in the product fabrication industries– an occupation likely involving very few cash transactions– had the lowest estimates of noncompliance. Further, households with significant amounts of self-employment income which are headed by younger males tend to conceal income to a greater extent than those headed by older males. Thus, a tax audit scheme targeting groups which have been found to conceal income the most (such as self-employed households headed by younger males or those in the construction and service occupations) is likely to be an effective tool in reducing noncompliance. The implications of these results are discussed in greater detail below.

The remainder of this paper is organized as follows. Section 2 briefly reviews the existing literature on income tax compliance– focusing on the various techniques used to derive estimates of income concealing among the self-employed. Section 3 outlines the methodology used to derive estimates of tax noncompliance by the self-employed in Canada. In Section 4 the data used in

estimation, the sample restrictions and the empirical approach are described. Section 5 outlines the results and Section 6 concludes.

(2) Previous Research

This section briefly summarizes previous attempts to provide estimates of the extent to which income is concealed, with special attention given to the few studies which provide estimates of noncompliance for the self-employed. For a more thorough review of the literature see, for example, Erard (1997) on Canada and Smith (1986) on the UK. The various sources of information on tax evasion used by researchers to this point are derived from i) national accounts data, ii) tax audit data, iii) tax amnesty data, and iv) survey data.

A common aggregate measure of underground activities employed by researchers is the "GDP gap". The GDP gap is a measure of unrecorded GDP and is calculated as "the difference between total market-based income (whether from legal or illegal sources) in the domestic economy and recorded GDP".[3] Estimates of unrecorded GDP for Canada range, depending upon the year and method used to obtain an estimate, from 2.5% of GDP in the 1970's (Pouftis 1993) to 21.6% of GDP in 1990 (Mirus, Smith and Karoleff 1994). In addition to the wide variation in estimates, there is also no consensus on the trends in unrecorded GDP in Canada. At least two authors (Ethier 1985 and Pouftis 1993) find evidence that the fraction of GDP unrecorded rose from the 1960's to the early 1990's but fell in the early 1980's. At the same time, estimates provided by Mirus and Smith (1981) and Mirus, Smith and Karoleff (1994) suggest that the percent of GDP that was unrecorded did not change much between 1976 and 1990.

There are several problems with this approach as it relates to the focus of the current paper. Because the GDP gap is measured at the aggregate level one cannot get a breakdown of the estimate

of underground activity for individual groups in the economy. For example, no estimate of underground activity by the self-employed can be obtained using this approach. Further, as Erard (1997) points out, measuring unrecorded GDP is not necessarily equivalent to measuring tax non-compliance– which is the focus of this paper. For instance, not all income included in unrecorded GDP is taxable. Some portion of the GDP gap may be attributable to individuals earning income below that which is taxable. Also, because deductions are included in gross income, the GDP gap captures only the under-reporting of gross income and leaves out overstated tax deductions and credits. Finally, as is suggested by the large variation in estimates of the GDP gap, the potential for measurement error inherent in this approach does not lend itself well to investigations of time trends in under-reporting.

A second source of information on underground activity that is available, though only for the US, is data collected through tax audits. Through the Taxpayer Compliance Measurement Program (TCMP) the US Internal Revenue Service publishes estimates of the difference between federal tax liabilities reported and assessed liabilities owed by individuals and small corporations. Under the TCMP a stratified random sample of individual and corporate income tax returns are subjected to intensive audits. Estimates based on these audits are produced and published for various groups. Most notably, a breakdown by industry for sole proprietors provided by the US General Accounting Office (1990) suggests that the degree of underground activity by sole proprietors varies greatly by industry. Among the industries in which the estimated extent of noncompliance by sole proprietors was the greatest include: retail sales (fixed location) where it is estimated that sole proprietors payed 39% less tax on average than should have been paid, transportation 36%, retail sales (no fixed location) 31% and production (including construction)

24%. These estimates of noncompliance were lowest among propietors in wholesale trade 19%, and finance/insurance/real estate 16%. In comparison the estimate of the extent of noncompliance among wage and salary employees is low. It is estimated that wage and salary employees paid only 1% less tax than should have been paid. Although these estimates are based on data from the US, given the similarities between Canada and the US in terms of their labour market characteristics and overall tax structure, to the extent that they provide a general pattern of noncompliance, these estimates are relevant to the current study.

While audit data provides estimates of both under-reported income and overstated tax deductions and credits, these estimates are not very reliable. Estimates based on audits rely heavily on the auditors' ability to identify under-reporting and interpretation of the tax laws as they pertain to tax deductions and credits. It is highly unlikely that income tax auditors are able to identify all income that is concealed from the tax authorities. Also, no audit data or estimates are available for Canada.

Tax amnesty data also provide a direct source of information on tax noncompliance. Data on self-reported evasion by those choosing to participate in the amnesty program have been used to obtain estimates of noncompliance (see Andreoni, Erard and Feinstein 1998 for a discussion). The biggest problem with such data is that of sample selection bias. Those who participate in such programs are not likely to be representative of the entire population. Further, they might provide information only on sources that were likely to be traced later.

Finally, a unique approach which uses household expenditure micro-survey-data has been developed by P&W (1989). Essentially, this approach uses unbiased estimates of the marginal propensity to consume food in order to impute true income levels for the self-employed. A more

detailed description of this approach is provided in the next section. This method has been applied recently by Mirus and Smith (1996) using the 1990 Canadian Survey of Family Expenditures. They estimate that in 1990 households obtaining a significant percent of total income from self-employment concealed 12.5% of their incomes, on average, or $6 billion, in aggregate. While informative, the fact that their estimate is aggregated across the economy and based on a single year is limiting. As noted earlier, disaggregating is likely to reveal clues about the determinants of noncompliance among the self-employed and provide guidelines in designing effective tax policies. In addition, the results presented below suggest that estimates based on a single year of data without controlling for hours worked can be misleading[4].

(3) Methodology

Following closely the strategy developed by P&W for a single year of data, the approach taken here is to estimate the extent of noncompliance among the self-employed based on household expenditure patterns pooling across six years of data covering the period 1969 to 1992. Estimation is made possible by a few key assumptions. It is assumed that: i) all respondents report expenditures on food correctly in the FAMEX surveys, ii) all earners report the same income in the survey as they do on their tax return[5], iii) wage and salary earners report income correctly while the self-employed may not, and iv) the relationship between family expenditure and after-tax income is the same for all families in the survey– whether self- or wage and salary employed.

Based on these assumptions a two-stage estimation technique can be applied to obtain estimates of the extent of noncompliance. Essentially, a prediction of the relationship between household food consumption and after-tax income, controlling for household characteristics and year effects, is obtained for all years in the first stage. This undistorted marginal propensity to consume

food is calculated using data on households obtaining all of their income from wage and salary employment. These households have very few opportunities to conceal income. In the second stage this estimate of the underlying relationship between food consumption and after-tax income is used to impute estimates of "true" income for households obtaining a substantial fraction of household income from self-employment in each year, demographic and occupational group. If the self-employed are concealing income we would expect household food consumption among these households to be high relative to the wage employed with similar reported incomes and similar characteristics. When compared to reported income these imputed incomes provide estimates of non-compliance among these self-employed households. Because this method yields a measure of *any* difference between "true" and reported *after-tax* income, both unreported income and overstated deductions will be captured in the estimate of noncompliance[6].

The expenditure equation can be written:

$$\ln Xf_i = \alpha Z_i + \beta \ln Y_i^P + e_i \qquad \textbf{(1)}$$

where $\ln Xf_i$ is the natural log of food expenditure for household i, Z_i is a set of household characteristics and year indicators, and Y_i^P is permanent after-tax income.

However, as P&W suggest, reported income will differ from permanent income for two reasons. First, following the discussion above, true income will differ from reported income for the self-employed because of noncompliance– but not for wage and salary employed. The relationship between true income (Y_i^T) and reported income (Y_i^R) can be written:

$$Y_i^T = k_i Y_i^R \qquad \textbf{(2)}$$

where by assumption $k_i=1$ for wage and salary earners and $k_i>1$ for the self-employed. A larger k_i indicates a greater degree of noncompliance. Next, true income (Y_i^T) will differ from "permanent" income (Y_i^P) because of transitory shocks. P&W write the correspondence between true income and

permanent income as:

$$Y_i^T = p_i Y_i^P \qquad \textbf{(3)}$$

where p_i is a random variable measuring the effect of aggregate events on current income. Therefore, k_i and p_i will enter the expenditure equation as follows:

$$\ln Xf_i = \alpha Z_i + \beta \ln Y_i^R - \beta \ln p_i + \beta \ln k_i + e_i \qquad \textbf{(4)}$$

Thus, $\ln(p_i)$ and $\ln(k_i)$ enter the expenditure equation along with reported income and each of these regressors enters with the same coefficient. However, p_i and k_i are not observed. Further, it is assumed that, while the mean transitory shock (mean of p_i) is the same for both the wage and self-employed, the variance of p_i is likely to differ across employment sectors. To make estimation possible, therefore, it is assumed that p_i and k_i have log normal distributions[7].

By introducing a self-employment indicator SE_i (=1 if household i is self-employed, 0 otherwise), equation (2) can be altered to isolate differences in expenditure between wage and self-employed households. In the following estimation equation:

$$\ln Xf_i = \alpha Z_i + \beta \ln Y_i^R + \delta SE_i + e_i \qquad \textbf{(5)}$$

α and β are assumed to be the same for all households. The intercepts for the two groups will differ if the self-employed conceal income and because the variance of transitory shocks is likely to be higher among the self-employed than wage workers. P&W conclude that δ ' $\beta[\mu_k \%1/2(\sigma^2_{uSE} \& \sigma^2_{uEE})]$ where μ_k is the mean of the natural log of k_i and σ^2_u is the variance of p_i (subscripts SE and EE refer to the self- and wage-employed respectively). From the estimate of δ one cannot isolate μ_k, because of the differences in the variance of log p_i between the self- and wage-employed.

What one would like to estimate is the average degree of noncompliance, the average value of k_i, denoted ($\bar{k}$). Assuming log normality, it can be shown that the log of this value can be written:

$$\ln \bar{k} = \mu_k + \frac{1}{2}\sigma^2_{vSE} \qquad \textbf{(6)}$$

where σ^2_v is the variance of k_i. By substituting the interpretation of δ given above into equation (6), $\ln \bar{k}$ can be rewritten:

$$\ln \bar{k} = \frac{\delta}{\beta} + \frac{1}{2}(\sigma^2_{vSE} - \sigma^2_{uSE} + \sigma^2_{uEE}) \qquad \textbf{(7)}$$

However, because p_i and k_i are not observed, it is not possible to obtain estimates of the variances related to aggregate shocks and noncompliance, separately. Hence, no single estimate of the degree of noncompliance can be calculated. Instead, an income equation is estimated in order to obtain estimates of the income variance of errors for the self-employed ($\sigma^2_{\lambda SE}$) and wage workers ($\sigma^2_{\lambda EE}$), separately. P&W show that, under a reasonable set of assumptions, this aggregate[8] measure of the income variance will allow estimates of upper and lower bounds on the extent of noncompliance to be computed[9]. It can be shown that the upper bound $\bar{k}_u$ can be estimated by substituting estimated values of $\delta, \beta, \sigma^2_{\lambda SE}$, and $\sigma^2_{\lambda EE}$, in:

$$\ln \bar{k}_u = \frac{\delta}{\beta} + \frac{1}{2}(\sigma^2_{\lambda SE} - \sigma^2_{\lambda EE}) \qquad \textbf{(8)}$$

An estimate of the lower bound ($\bar{k}_l$) can be obtained by substituting estimated values in:

$$\ln \bar{k}_l = \frac{\delta}{\beta} - \frac{1}{2}(\sigma^2_{\lambda SE} - \sigma^2_{\lambda EE}) \qquad \textbf{(9)}$$

(4) Data and Estimation

The data used in this analysis are taken from a series of Canadian Family Expenditure Surveys (FAMEX) from 1969 to 1992. Because the FAMEX surveys are conducted somewhat sporadically over time and some key variables are absent in some years of data, only the following six years of data could be used in estimation: 1969, 74, 84, 86, 90, and 92. The FAMEX surveys are conducted in the first few months of the calendar year and provide retrospective information for the previous year on household's income, demographics and detailed expenditures. The unit of analysis used throughout this study is the household spending unit, which is assumed to act as a single decision maker for decisions about income reporting and expenditures[10]. In all years, the sample is restricted to private households in 15 metropolitan areas in Canada. The sample studied

is further restricted to i) households containing two adults where the spouse of the head is the wife, ii) the household had positive after-tax income, iii) the head was employed full-time full-year (49 weeks or more) in the previous year-- either in the wage and salary or self-employment sector, iv) the head was aged 25 to 64, and v) the head was employed in a non-primary occupation. This was done to eliminate households from the sample deriving much of their self-employment income from farming. Farm households are likely to have much different expenditure patterns on food than those in other occupations. These restrictions left a total of 8,463 households, of which 8.1%, or 682 households were "self-employed" according to the definition which follows.

In a regression setting using the data files pooled across the years, equation (5) was estimated for a number of different specifications. The dependent variable is the natural log of expenditure on food (converted to real 1986 dollars using the Consumer Price Index-- All Items), whether purchased for home or in restaurants. Food was chosen as the dependent variable in this analysis for several reasons. First it is likely the case that tastes for food are more likely to be uniform across employment groups because food is a necessity. Second, households are typically unable to postpone food consumption because of transitory shocks. Third, expenditure on each of the many individual food items surveyed tends to be small. Thus, individuals who under-report income are not likely to be concerned that expenditures on each of these items alone will raise suspicion. One might argue that by including food in restaurants in total food consumption there is potential for upward bias in the estimates of noncompliance if it is the case that the self-employed simply eat in restaurants more frequently where the price of food tends to be higher. However, similar to what Mirus and Smith (1996) found there does not appear to be a relationship between the share of income from self-employment and the ratio of restaurant to store bought food in my sample[11]. In addition, controls

for differences in household characteristics that may influence the ratio of store to restaurant bought food are included in equation 5. As a final check, I experimented with a number of different expenditure items. These results yielded similar patterns of income under-reporting[12].

The independent variables include the natural log of net after-tax family income[13] at real 1986 values, an indicator variable for self-employment status (defined below), the self-employment indicator interacted with year or household characteristic dummies as indicated below, a series of controls for household characteristics, and year indicator variables. The implicit restriction made in this pooled setting is that the marginal propensity to consume food is the same across time[14]. Controls for household characteristics (Z_i) include: the age of the head, the head's age squared, the number of children present in the household-- separately for young children (aged 6 or less), "middle aged" children (7-12) and older children (13-17), the wife's level of education, indicator variables for region, the number of rooms in the household's dwelling, the value of the house resided in if owned, a dummy variable indicating households that rent and the year that the household was surveyed.

The self-employment indicator takes on the value of one for those households classified as self-employed and zero otherwise. Individuals are not asked to self-report self-employment status in the FAMEX surveys. Instead, in this study self-employment status is determined by the fraction of total household income from all sources coming from self-employment. Income from self-employment includes: net income from non-farm self-employment, net income from farm self-employment, and gross income from roomers and boarders not related to the head. "Self-employed households" are defined as households deriving 30 percent or more of total income from self-employment.[15] With this definition I attempt to capture households with sufficient income from self-

employment to allow for the opportunity to conceal income from the tax authorities[16].

Table 1 presents sample characteristics for households classified as self-employed and wage and salary, separately. The self-employed in this sample tended to be older, had more children and were less likely to rent, on average, than those in wage and salary employment. In addition, the self-employed were more likely to reside in British Columbia and the Prairies and less likely to reside in the Atlantic Provinces and Quebec than wage and salary workers. Incomes were higher among the self-employed as compared to wage and salary earners. Somewhat surprisingly, however, the standard deviations of the natural log of incomes were no higher for the self-employed relative to wage and salary earners, as was expected. Finally, average log expenditures on food by the self-employed were higher than the same expenditures by the wage and salary employed.

Because permanent after-tax income is measured with error and the error term is assumed to be heteroscedastic, equation (5) is estimated by generalized two-stage least squares.[17] Reported income is treated as endogenous and estimated as appendix equation (A3) , where the instruments, X_i, are a series of indicators for the head's education level, indicators for the wife's work intensity and the self-employment indicator interacted with the number of children, the wife's work intensity and the number of rooms in the dwelling.[18] The results of this particular specification are presented because the test of the over-identifying restrictions was not rejected. However, several alternative specifications were tried with similar outcomes.

(5) Results

In order to calculate the upper and lower bounds on the estimates of noncompliance in equations (8) and (9) one requires estimates of the coefficients β and δ from equation 5, and the residual variance of income for the self- and wage-employed, separately. These coefficients, the

residual variances obtained from the first stage regression, and the resulting estimates of noncompliance by year are presented in Table 2. Estimates of the upper and lower bounds on $\overline{k}$ are calculated for each year of data from the coefficients on the self-employment status indicator interacted with year indicators and estimates of the residual variance of income for the two groups. The estimates of the coefficient on log reported income and the residual variances from income for the two groups are restricted to be unchanged over the period.[19] (For the full set of coefficients for an end-stage regression see appendix Table 1). The estimated marginal propensity to consume food, $\hat{\beta}$, is positive and significantly different from zero at the five percent level and is estimated to be about 0.39. The estimated coefficient on the self-employment dummy, $\hat{\delta}$, is positive and significant in the base year (1969). While individually each of the interactions between the self-employment indicator and year interactions is statistically insignificant, these coefficients are jointly significant at standard levels.

The estimates of $\overline{k}$ give the numbers, on average, by which reported self-employed household incomes must be multiplied to derive true income. For example, the upper bound estimate for 1969 suggests that true household incomes among households deriving thirty percent or more of total income from self-employment were 26 percent higher, on average, than reported. For self-employed households in which income is comprised of both wage and self-employment income these numbers do not reflect the number by which *self-employment income* alone must be multiplied to derive true income[20]. One advantage to the estimates of noncompliance based on total household income is that they account for forms of noncompliance, such as income splitting, that may distort the true fraction of self-employment income[21]. There is, however, the potential for bias if the fraction of self-employment income varies systematically by year or household characteristic.

To test for such a bias estimates based on the households self-employment income are computed using the average fraction of household income derived from self-employment by year in Table 2 (and later in the paper by occupation and household characteristic Tables 3 and 4). These estimates are provided in square brackets in Tables 2 to 4 for comparison. While these estimates tend to be higher than those based on total household income, similar patterns of noncompliance are observed using either estimates[22]. Because of this result and the potential advantage noted above this analysis will focus on the estimates that yield the fraction of total household income that is concealed.

In Table 2 the implied effect of self-employment on food expenditure is positive in all years. This suggests that, controlling for income and household characteristics, households with significant self-employment income consume more food than wage and salary households. There is, however, no discernible time trend in the estimated degree of income concealing by self-employed households over this period. In 1969 the upper and lower bounds on $\overline{k}$ fall in a somewhat narrow range between 1.14 to 1.26. Thus, self-employed households concealed, on average, between 14 and 26 percent of household income from the tax authorities in 1969. These estimates drop slightly after 1969 and, if anything, trend upward after 1974. Interestingly, the introduction of the Goods and Services Tax (GST) in 1991 appears to have had no effect on the average degree of noncompliance among the self-employed. One might have expected a reduction in under-reporting given that self-employed workers who simply do not report income would not be able to deduct GST paid on certain inputs used in production. The evidence presented in Table 1 suggests that this was not sufficient enough of a disincentive to reduce noncompliance. Perhaps, the decrease in the net benefits to under-reporting were insufficient to merit a decrease in such activity. Alternatively, the self-employed may have increased the amount of income that was hidden through the overstatement of expenses,

which would not alter the ability to deduct GST payments.

It is unclear as to how educational attainment, controlling for household income, would affect the degree of noncompliance by the self-employed. It might be the case that those who are more highly educated are more familiar with the tax laws and are, therefore, more able to avoid paying taxes. On the other hand, those with lower levels of education might be more likely to enter industries, such as construction, which provide a greater opportunity to conceal income. Table 3 presents estimates of noncompliance by the self-employed by the head's level of education applying a similar estimation technique to that used above. One difference is that the self-employment indicator is no longer allowed to vary by year or occupation. Instead, the self-employment indicator is interacted with education category indicators. The results presented in Table 3 show no clear pattern of noncompliance by head's education level. If anything, the degree of noncompliance rises with the level of education up to some post-secondary education. Households where the head has a post-secondary degree tend to conceal the least, while those where the head has a university degree conceal approximately the same fraction of income as those whose head has a secondary degree.

Next, to further identify the possible determinants of noncompliance and to provide guidance to policy makers involved in developing tax policy I provide estimates of noncompliance by characteristics which are typically observed on individuals' income tax returns. The Canadian federal tax authority observes the tax filer's occupation and age as well as information about the tax filer's spouse– including whether or not the spouse is self-employed. Table 4 provides estimates of noncompliance among self-employed household's by head's occupation, head's age category and which of the adult members in the household are self-employed. Here the estimates are derived, as with the previous estimates, by allowing the self-employment indicator in equation 5 above to vary

by certain characteristics of the household but maintaining all of the same control variables. The results in the upper section of Table 4 are derived by interacting the self-employment indicator with indicator variables for the head's occupation instead of year or education indicators. The occupation groupings are: managerial and administration, professional and technical, sales, services, product fabricating, construction, and other occupations. Due to changes in the coding of the occupation variable, only the years 1984, 86, 90 and 92, which maintained the same coding, were included in the regression. Further, the self-employment indicator is interacted with indicator variables for the head's age category and for which of the adult members of the household are self-employed, in the middle and lower sections of Table 4 respectively. For these two estimates all six years of data were employed again.

Although there appears to be no pattern of noncompliance over time or across education categories there is substantial variation across occupations. The upper section of Table 4 presents estimates of the upper and lower bounds on $\overline{k}$, by occupation for these years. Again, a joint test of the industry dummies crossed with self-employment status suggests that these coefficients are jointly significant at standard levels and significantly different from each other. The estimated bounds on $\overline{k}$ vary from a low of 0.96-1.05 in product fabricating occupations– which suggests that entrepreneurs in these occupations report their incomes correctly– to a high of 1.39-1.53 in construction occupations. The high degree of noncompliance among self-employed households in the construction occupations is followed closely by those in services with estimated bounds of 1.38-1.52. Interestingly, while not directly comparable, the results here are similar to those found in the US, presented in section 2, which suggest that construction is among the industries in which the self-employed conceal income the most, and product fabricating is among those with the least amount

of income concealing. One exception in which the results in section 2 differ from those in Table 4 arises when the retail sales industries in the US and the sales occupations in Canada are compared. There appears to be a great deal of noncompliance by entrepreneurs in the US in the retail sales (fixed location) industry while the estimates in Table 4 suggest that this is not the case for the sales occupations in Canada. These differences arise, most likely, because the industry category for the US is much more narrowly defined than the occupation category in Canada. The occupation category in Canada includes occupations such as wholesale trade– an occupation which has a much lower estimated degree of noncompliance among the self-employed in the US.

These results are consistent with the hypothesis that an entrepreneur's "opportunity" to conceal income as determined by industry/occupation characteristics is correlated with the actual degree of noncompliance. For instance, those involved in the construction trades typically provide services through informal arrangements frequently involving cash transactions. This type of self-employment undoubtably provides greater opportunity for under-reporting. On the other hand, these characteristics do not typify production industries. As the estimates in Table 4 show, entrepreneurs in construction tend to conceal a greater fraction of their income than those in product fabrication. This finding has important implications on the efficiency of the economy. Assuming that there is free entry into and out of occupations, individuals will be attracted, ceteris paribus, to those occupations which provide the greatest after-tax wages (i.e. those which provide the opportunity to conceal income). Even though pre-tax wages will adjust to equalize after-tax wages, too many resources, in terms of efficiency, will be allocated to occupations which provide the greatest opportunities for noncompliance. This result also suggests an obvious strategy for the development of an effective tax audit system. One which targets the self-employed in construction and service

occupations is likely to be more effective.

The estimates provided in Table 4 by age category indicate clearly that the degree of noncompliance by entrepreneurs decreases with age. The estimates increase monotonically with age– from a high of 1.16-1.29 in households headed by an individual aged 25-34, to a low of essentially full compliance (0.95-1.06) among those whose head is aged 55-64. This is again consistent with estimates for the US using TCMP data which indicates that the degree and incidence of noncompliance is lower among households headed by older individuals (US General Accounting Office 1990).

I am unaware of any study which looks at non-compliance by the number and sex of the self-employed members of a household. Do households in which there are two members earning self-employment income conceal a greater fraction of income than households with only a single self-employed member? If there is only one adult member self-employed, does it matter if it is the husband or wife that is self-employed? The estimates provided in the lower section of Table 4 suggest that households in which both the head and spouse earn self-employment income conceal a lower fraction of income than households in which only one adult member is self-employed. One possible explanation for this result is the likelihood that increases in the number of self-employed individuals in a household will yield marginally fewer opportunities to conceal income. For instance, if one individual in a household is self-employed and illegitimately claiming a deduction for the use of the family car as a business vehicle, a second self-employed household member will not be able to overstate the same deduction. In response to the second question posed above, it appears that it does not matter whether it is the husband or the wife that is self-employed when only one adult earns self-employment income. Estimates of non-compliance in households where only

the head is self-employed are not significantly different from those where only the wife is self-employed. The estimated range is 13-26 percent among self-employed heads and 15-28 percent among self-employed wives.

(6) Conclusions

The rise in self-employment experienced recently in Canada has brought to the forefront some of the problems associated with this form of employment. In particular, the fact that the self-employed are often able to conceal income from tax authorities makes understanding the determinants of non-compliance by this sector of the labour market increasingly important to those studying labour markets and to tax policy makers. To this point, however, estimates of the degree of income tax noncompliance by the self-employed have revealed little about the factors which influence the decision to conceal income. This paper employs an approach developed by Pissarides and Weber (1989) to obtain estimates of the degree of noncompliance across various dimensions, focusing primarily on individual and household characteristics commonly reported on income tax returns.

Estimates of the extent of noncompliance by households obtaining 30 percent or more of household income from self-employment yield an average lower bound estimate of 11 percent of household income over the six years examined and an average upper bound estimate of 23 percent. The estimate of noncompliance for the 1990 calendar year found here to be between 12 and 24 percent is somewhat higher than the 10 to 15 percent found in Mirus and Smith (1996) for the same year. The difference in the estimates may be attributable to the fact that the sample contained in this paper includes only full-time full-year workers.

The results when disaggregated across years, occupation, and demographic characteristics

provide important insights into the determinants of noncompliance among the self-employed and suggest actions that may be beneficial in designing tax enforcement policy. First, with regard to the overall determinants of noncompliance among the self-employed in Canada, I find little evidence that the degree of under-reporting has increased between 1969 and 1992. Further, there appears to be no clear relationship between the degree of noncompliance by a household and the household head's age. Second, for policy makers involved in developing taxation enforcement policies including audit schemes the results offer the following guidance. Given the limited financial resources available to enforce compliance, a tax audit scheme which targets occupations which provide self-employed individuals with the greatest opportunity to conceal income, such as the construction and service occupations, is likely to be more effective. In addition, targeting younger entrepreneurs and households with just one self-employed member for tax audits might also be an effective strategy in enforcing tax compliance. Clearly, such a policy, if continued for a period of time, would lead to the misrepresentation of occupation and other characteristics by self-employed tax filers to avoid being audited. However, if the information from these audits is used effectively, policies can be designed to reduce noncompliance. For instance, the results of such audits could be used to identify the source of noncompliance– overstated expenses or understated income. If, for example, it is revealed that in any given occupational or demographic group most of the noncompliance is achieved through overstated expenses of one form or another, policy makers could then alter which types of expenses to allow in order to reduce noncompliance. Finally, in light of the relative magnitude of the estimates of noncompliance among the self-employed provided above, the ease with which individuals are able to be classified as self-employed is of concern. The "relabelling" of wage and salary earners as contractors raises significant concerns in light of these

findings. To address this issue, additional restrictions might be placed on the requirements necessary for one to be classified as self-employed**.**

Tables

Table 1 Sample Means		
Employment Status	**Self-Employed***	**Wage and Salary**
Sample Size	682	7781
Heads Age	40.61 (9.24)	39.35 (9.73)
Young Children	0.56 (0.79)	0.56 (0.78)
Middle Children	0.67 (0.83)	0.58 (0.84)
Older Children	0.38 (0.71)	0.32 (0.66)
Head Less Than 9 Years Education	0.13 (0.34)	0.11 (0.31)
Head Some Secondary Education or Degree	0.39 (0.49)	0.42 (0.49)
Head Some Post-Secondary Education	0.09 (0.29)	0.08 (0.27)
Head Post-Secondary Degree	0.15 (0.36)	0.19 (0.39)
Head University Degree	0.24 (0.42)	0.21 (0.40)
Value of House if Owned (In 000's of 1986 real dollars)	83.01 (127. 99)	57.68 (87.30)
Rent Indicator	0.22 (0.41)	0.30 (0.46)
Atlantic Region	0.11 (0.32)	0.16 (0.37)
Quebec	0.20 (0.40)	0.22 (0.41)
Table 1 Continued		

Ontario	0.23 (0.42)	0.23 (0.42)
Prairies Region	0.31 (0.46)	0.29 (0.45)
British Columbia	0.15 (0.35)	0.10 (0.29)
ln Income	9.66 (1.32)	9.49 (1.34)
ln Food Expenditure	7.93 (1.19)	7.71 (1.21)
Values in parentheses are standard deviations. * Self-employed households are defined as households deriving 30 percent or more of household income from self-employment.		

Table 2
Key Parameter Estimates and Estimated Degree of Under-Reporting: by Year

Year	$\hat{\beta}$	$\hat{\delta}$ [a]	$\bar{k}_l$	$\bar{k}_u$

1969	0.389 (0.026)	0.070 (0.031)	1.14 [1.17]	1.26 [1.31]
1974	*	0.044 (0.032)	1.06 [1.07]	1.18 [1.22]
1984	*	0.055 (0.037)	1.09 [1.13]	1.21 [1.30]
1986	*	0.060 (0.039)	1.11 [1.15]	1.23 [1.32]
1990	*	0.063 (0.038)	1.12 [1.17]	1.24 [1.35]
1992	*	0.076 (0.039)	1.15 [1.21]	1.28 [1.40]
N	8463			
R-Squared	0.934			

Values in round parentheses are standard errors.
Upper and lower bound estimates in square parentheses are estimates weighted by the average fraction of household income derived from self-employment

For all estimates $\sigma^2_{\lambda SE} = 0.173$ and $\sigma^2_{\lambda EE} = 0.068$.

a: Estimates are the implied values for each year from the pooled regression.
*: Constrained to be the same in each year.
R-Squared taken from end stage regression.

Table 3
Key Parameter Estimates and Estimated Degree of Under-Reporting: by Head's Education Level

Head's Education	$\hat{\delta}$	$\bar{k}_l$	$\bar{k}_u$
Less Than 9 Years	0.049 (0.039)	1.08 [1.10]	1.19 [1.23]
Some Secondary or Secondary Degree	0.069 (0.023)	1.13 [1.17]	1.26 [1.33]
Some Post-Secondary	0.104 (0.047)	1.24 [1.35]	1.38 [1.55]
Post-Secondary Degree	0.017 (0.036)	0.99 [0.97]	1.10 [1.14]
University Degree	0.064 (0.030)	1.12 [1.16]	1.24 [1.32]
N	8463		
R-Squared	0.934		

Values in Parentheses are standard errors.
Upper and lower bound estimates in square parentheses are estimates weighted by the average fraction of household income derived from self-employment.

$\hat{\beta}$ = 0.389 with standard error 0.026

For all estimates $\sigma^2_{\lambda SE}$ = 0.172 and $\sigma^2_{\lambda EE}$ = 0.068.

R-Squared taken from end stage regression.

Table 4
Key Parameter Estimates and Estimated Degree of Under-Reporting:
by Characteristics Observed on Income Tax Returns

	$\hat{\delta}$	$\bar{k}_l$	$\bar{k}_u$
Occupation*			
Managerial and Administration	0.050 (0.042)	1.10 [1.14]	1.20 [1.28]
Professional and Technical	0.055 (0.041)	1.09 [1.12]	1.20 [1.27]
Sales	0.026 (0.047)	1.02 [1.03]	1.12 [1.17]
Services	0.154 (0.065)	1.38 [1.57]	1.52 [1.77]
Product Fabricating	0.0002 (0.077)	0.96 [0.94]	1.05 [1.07]
Construction	0.157 (0.049)	1.39 [1.54]	1.53 [1.73]
Other Occupations	0.029 (0.058)	1.02 [1.03]	1.12 [1.18]
N	3901		
R-Squared	0.42		
Head's Age Category			
Aged 25-34	0.079 (0.027)	1.16 [1.21]	1.29 [1.38]
Aged 35-44	0.078 (0.022)	1.16 [1.21]	1.29 [1.39]
Aged 45-54	0.025 (0.034)	1.01 [1.01]	1.13 [1.08]
Aged 55-64	0.001 (0.043)	0.95 [0.94]	1.06 [1.08]
N	8463		
R-Squared	0.934		
Table 4 Continued			
Adult Members Self-Employed			

Head Only	0.069 (0.016)	1.13 [1.17]	1.26 [1.34]
Wife Only	0.077 (0.066)	1.15 [1.36]	1.28 [1.67]
Both Head and Wife	0.010 (0.036)	0.97 [0.97]	1.08 [1.09]
N	8463		
R-Squared	0.934		

Values in Parentheses are standard errors.
* because of data limitations the sample for the analysis estimating across occupations is restricted to the years 1984, 1986, 1990 and 1992

For the analysis looking across occupations $\hat{\beta}$ = 0.415 with a standard error of 0.041 and the estimated variances $\sigma^2_{\lambda SE}$ = 0.161 and $\sigma^2_{\lambda EE}$ = 0.068

For the other two analyses $\hat{\beta}$ = 0.390 with standard error 0.026, $\sigma^2_{\lambda SE}$ = 0.175 and $\sigma^2_{\lambda EE}$ = 0.068.

R-Squareds are taken from end stage regression.

References

Andreoni, J., Erard, B., and J. Feinstein (1998) "Tax Compliance," Journal of Economic Literature, Vol. 36, pp. 818-860.

Auerbach, A., and Martin Feldstein eds. (1998) Handbook of Public Economics.

Baker, P., (1993) "Taxpayer Compliance of the Self-employed: Estimates from Household Spending Data," Working Paper No. W93/14, The Institute for Fiscal Studies.

Erard, B., (1997) "A Critical Review of the Empirical Research on Canadian Tax Compliance," Working Paper 97-6, Business Income Tax Division: Dept. of Finance Canada.

Ethier, M. (1985) "The Underground Economy: A Review of the Economic Literature and New Estimates for Canada," in Income Distribution and Economic Security in Canada, ed. F. Vallaincourt, Toronto, University of Toronto Press, pp. 77-109.

Hill, R. and M. Kabir (1996) "Tax Rates, the Tax Mix and Growth of the Underground Economy in Canada: What Can We Infer?," Canadian Tax Journal (44:6) 1552-1583.

Kuhn, P., and Herb J. Schuetze (2001) "Self-Employment Dynamics and Self-Employment Trends: A Study of Canadian Men and Women, 1982-1995," Canadian Journal of Economics, Vol. 34, No. 2. pp. 760-781.

Lin, Zhengxi, Janice Yates and Garnett Picot (1998), "Rising Self-employment in the Midst of High Unemployment: An Empirical Analysis of Recent Developments in Canada," Statistics Canada Analytical Studies Branch Research Paper.

Mirus, R., and Roger S. Smith (1981) "Canada's Irregular Economy," Canadian Public Policy (7:3), pp. 444-453.

Mirus, R., and Roger S. Smith (1996) "Self-Employment, Tax Evasion and the Underground Economy: Micro-Based Estimates for Canada," Working Paper 1002, International Tax Program, Harvard Law School.

Mirus, R., R.S. Smith and V. Karoleff (1994) "Canada's Underground Economy Revisited: Update and Critique," Canadian Public Policy (20:3) 235-52.

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

Pouftis, A. (1993) "Estimating the Size of the Underground Economy in Canada," M.A. Dissertation Essay, Queen's University.

Schuetze, H.J., (2000) "Taxes, Economic Conditions and Recent Trends in Male Self-Employment: A Canada-U.S. Comparison", Labour Economics Vol.7, number 5, pp. 507-544.

Slemrod, J. ed. (1992) Why People Pay Taxes: Tax Compliance and Enforcement University of Michigan Press.

Slemrod, J. and S. Yitzhaki (1997) "Tax Avoidance, Evasion and Administration," in Alan Auerbach and Martin Feldstein, eds., Handbook of Public Economics.

Smith, S., (1986) Britain's Shadow Economy, Oxford University Press, Oxford.

Spiro, P.S. (1997) "Taxes, Deficits, and the Underground Economy," in Owen Lippert and Michael Walker, eds., The Underground Economy: Global Evidence of its Size and Impact, The Fraser Institute, Vancouver, 37-52.

U.S. Government Accounting Office (1990) Tax Administration: Profiles of Major Components of the Tax Gap, Document GGD-90-53BR, Washington, DC.

Appendix A: The Model

Part I:

Assuming that p_i and k_i are log normal, one can write them as deviations from their means:

Appendix A: The Model

$$\ln p_i = \mu_p + u_i \qquad \textbf{(A1)}$$

$$\ln k_i = \mu_k + v_i \qquad \textbf{(A2)}$$

Where u_i and v_i have zero means and constant variances σ_u^2 and σ_v^2.

Substituting equations (A1)-(A2) into (4) yields:

$$\ln Xf_i = \alpha Z_i + \beta \ln Y_i^R - \beta(\mu_p - \mu_k) - \beta(u_i - v_i) + e_i \qquad \textbf{(A3)}$$

For wage and salary earners k=1, therefore, μ_k=0. For the self-employed k>1. Further, P&W show that, because the variance of p_i is greater for the self-employed than for wage earners, this implies that μ_p (the mean of the log of p_i) is smaller for the self-employed than for wage and salary earners. Thus, equation (A3) will differ between the self-employed and wage and salary earners because μ_p-μ_k differs between the two groups.

Part II:

The income equation can be written as follows:

$$\ln Y_i^R = \delta Z_i + \pi X_i + \lambda_i \qquad \textbf{(A4)}$$

where Z_i has the same interpretation as in equation (1) and X_i is a set of instruments. Under the assumption that unexplained variation in permanent income have the same variances for the self- and wage-employed (they have the same σ_u^2), P&W show that the difference between the residual variances of reported income from equation (A4) for the two groups can be written:

$$\sigma^2_{\lambda SE} - \sigma^2_{\lambda EE} = \sigma^2_{uSE} + \sigma^2_{vSE} - 2\,\mathrm{cov}(uv)_{SE} - \sigma^2_{uEE} \qquad \textbf{(A5)}$$

Assuming that noncompliance and aggregate events are uncorrelated ($\mathrm{cov}(uv)_{SE}$=0), equation (7) along with estimates of the difference between the residual variances from equation (A4) can be used to obtain upper and lower bounds on noncompliance.

Because σ^2_{vSE} and σ^2_{uSE} are inversely related in equation (7), a lower bound estimate of noncompliance can be obtained from the residual variances in (A4) when σ^2_{vSE} takes on its lowest value (it enters positively in equation 7) and an upper bound when σ^2_{uSE} takes its lowest value. The lowest possible value for σ^2_{vSE} is zero. Therefore, substituting this value in (A5) and (7) it can be shown that the natural log of the upper bound $\bar{k}_u$ can be estimated by substituting estimates of the parameters (β and δ) and of the differences in the residual variances($\sigma^2_{\lambda SE} - \sigma^2_{\lambda EE}$) in:

$$\ln \bar{k}_u = \frac{\delta}{\beta} + \frac{1}{2}(\sigma^2_{\lambda SE} - \sigma^2_{\lambda EE}) \qquad \textbf{(equation 9 above)}$$

Because it is assumed that the variance of self-employed earnings is at least as large as those of wage and salary earners, the lowest possible value for σ^2_{uSE} is when $\sigma^2_{uSE}=\sigma^2_{uEE}$. Substituting this value again in (A5) and (8) and replacing parameters with estimates yields an estimate of the lower bound ($\bar{k}_l$) which satisfies:

$$\ln \bar{k}_l = \frac{\delta}{\beta} - \frac{1}{2}(\sigma^2_{\lambda SE} - \sigma^2_{\lambda EE}) \qquad \textbf{(equation 10 above)}$$

Appendix Table

Appendix Table 1
Typical End Stage Regression: Dependent Variable ln Food Expenditure

Coefficient	Base Model (Self-Employment Indicator Crossed with Year)
ln(income)	0.389 (0.026)
sed (self-employment indicator)	0.070 (0.031)
sed cross 1974	-0.026 (0.044)
sed cross 1984	-0.015 (0.048)
sed cross 1986	-0.010 (0.049)
sed cross 1990	-0.007 (0.049)
sed cross 1992	0.006 (0.05)
Quebec	-0.039 (0.014)
Ontario	0.083 (0.013)
Prairies	0.019 (0.013)
British Columbia	-0.045 (0.013)
head's age	0.009 (0.001)
head's age squared X 10^3	-0.224 (0.039)
# young children	0.067 (0.005)
# mid aged children	0.090 (0.005)
# older children	0.123 (0.006)
year2 (1974)	0.397 (0.02)
year3 (1984)	1.332 (0.062)
year4 (1986)	1.424 (0.065)
year5 (1990)	1.585 (0.073)
year6 (1992)	1.608 (0.075)
rent	0.039 (0.009)

Appendix Table 1
Continued

value of house X 10^6	0.139 (0.061)
# of rooms	0.008 (0.003)
wife's education: Some Secondary	-0.042 (0.013)
wife's education: Some Post-Second	-0.055 (0.018)
wife's education: Post-Second Degree	-0.057 (0.016)
wife's education: University Degree	-0.074 (0.02)
Constant	2.695 (0.189)
Sample Size	8463
R - square	0.934
Values in parentheses are standard errors.	

Endnotes

1.The source for these data are taken from Statistics Canada published tabulations. The definition of self-employed can be found in Statistics Canada's "Guide to the Labour Force Survey". The definition and sample are different than those in the main body of the paper and are, therefore, not comparable. For a more in depth look at the trends in self-employment in Canada see Lin, Yates and Picot (1998) or Kuhn and Schuetze (2001).

2.Taken from "Income Statistics 2000: 1998 Tax Year," Canada Customs and Revenue Agency publication.

3.Erard (1997) page 2.

4.The results without pooling across years tend to be very noisy. In addition, the effect of the business cycle also resulted in a great deal of variance in the estimates from year to year. In the present paper this was remedied by restricting the sample to full-year full-time workers (a restriction Mirus and Smith did not make).

5.This assumption is not critical to the analysis. If, for instance, individuals are more willing to report income correctly in a survey because of assurances of confidentiality, these estimates can be interpreted as a lower bound estimate of income under-reporting on tax returns. Further, in looking across time and other characteristics the pattern is not likely to be biased because of violation of this assumption.

6.This assumes that tax expenses and deductions for the self-employed are allowed by the tax authority to achieve horizontal equity. Otherwise, a real differential in the tax treatment of the two groups would result and bias the estimates. What is important to this analysis is the intent of these tax laws and not the interpretation of such laws by the self-employed.

7.See the appendix, part I, for a description.

8. The income variance of errors is compiled of three types of errors: unexplained variations in permanent income, and the two types of errors arising from deviations from permanent income discussed above.

9.Equations 9 and 10 are derived in the appendix, part II.

10. The selection of a sample of family units as opposed to single adult households was made because there are insufficient observations to study the latter.

11.The ratio of restaurant bought to store bought food for wage and salary and self-employed households is 43 and 38 percent respectively. However, mean reported income among self-employed households was approximately 14 percent higher than wage employed households.

12.These results are available from the author.

13.Income after taxes is calculated as the difference between "income before taxes" (which includes wages and salaries, income from self-employment, investment income, government

transfers and retirement pensions and other money incomes) and "personal taxes" (individuals are asked to report personal income tax paid in the survey year).

14.In fact, an informal test of this assumption confirms that it cannot be rejected. In regressions for each year separately the marginal propensities to consume food were not statistically significantly different from one another at standard levels.

15.Several other alternatives to the 30% cutoff were tried. The results of these estimates were not significantly different. These results are available from the author.

16.The self-employment rate with this definition is not comparable to published estimates. The definition in this paper is based on the family's major source of income while published data are based on the individual's major source of income. Despite these differences the general upward trend found in published data over this period is also observed using this definition.

17.It is likely that the self-employment indicator, as defined, is also subject to some measurement error. Because some self-employment income is reported incorrectly by assumption, so is the fraction of income from self-employment. I experimented with a model treating the self-employment indicator as endogenous, using several different instruments (including many very similar to those used by P&W) which led to highly unstable results. Tests of the overidentifying restrictions suggest that these results may be caused by poor instruments. Hence, self-employment is treated as exogenous. Households which earn a large fraction of income from self-employment but conceal much of this income will be mis-classified as wage earners. This will likely lead to a greater than expected estimate of food consumption, controlling for income, among wage earners and a lower than expected estimate for the self-employed (i.e. a downward bias of the coefficient on self-employment status).

18.I experimented with several different specifications and estimated using various combinations of instruments. Unlike P&W, I found that measures such as the number of rooms and the value of the house belonged in both the food expenditure and income equations.

19.One might be concerned that these do not vary over time. Though not provided here, estimates obtained through generalized 2SLS for each year are consistent with this restriction.

20.For example, a household receiving 70 percent of income from self-employment may have an estimate of $\overline{k}$ equal to 1.5. This implies that 50 percent of total household income is concealed. However, this means that approximately 71 percent (50/70) of self-employment income is concealed.

21.For example, households may split self-employment income among the members of the household in order to reduce the tax burden. If this income is redistributed among the members of the household in the form of wages and salaries this will result in a lower than expected fraction of income from self-employment.

22.The one exception worth noting involves the estimates of noncompliance across the number of adult household members self-employed. The average fraction of household income from

self-employment among households in which only the wife is self-employed is low (42%) relative to head only (76%) and both head and wife (85%). This results in much higher estimates among wife only self-employed households. This may result from the fact that only 31 households in my sample contained only a self-employed wife.