Estimating the Informal Supplier Tax Gap[1]

James Alm, Georgia State University, and
Brian Erard, B. Erard and Associates

The Internal Revenue Service (IRS) defines informal suppliers as "individuals who provide products or services through informal arrangements which frequently involve cash-related transactions or `off the books' accounting practice" (IRS, 1996, p. 43). Examples include self-employed domestic workers, street-side vendors, and moonlighting tradesmen. Conceptually, the informal economy within which such individuals operate includes all types of market economic activity that are potentially undermeasured in the National Income Accounts owing to the vendors' informal business styles (e.g., sales in cash, lack of adequate records of sales and purchases). More relevant for our purposes, though, is the extent to which taxable self-employment earnings from informal market activities are—or are not— reported on individual income tax returns. This paper presents a new method by which the informal supplier "tax gap" (or the amount of true income tax liability of informal suppliers that is not reported on their income tax returns) can be estimated.

Owing in large part to the lack of a paper trail, tax noncompliance among informal suppliers can be especially difficult to uncover through examinations, even relatively intensive ones such as those performed under the National Research Program (NRP) or its predecessor, the Taxpayer Compliance Measurement Program (TCMP). In past tax gap reports, the IRS attempted to address this potentially severe nondetection problem with the aid of supplementary information from a special survey it periodically commissioned of consumer purchases in the informal sector. The first step was to use the survey results to develop an estimate of the aggregate gross receipts of informal suppliers. This estimate included reported consumer purchases of informally supplied goods and services as well as a rough approximation of business purchases from informal suppliers. The second step was to convert this estimate of gross receipts into an estimate of net self-employment income using an assumption about the ratio of net earnings to gross receipts. In the third step, IRS researchers attempted to identify informal suppliers on the

basis of the limited information available on the TCMP and to assess how much of this net income was actually reported on their tax returns. The difference between the estimated amount that was earned by informal suppliers (based on the survey) and the amount that was reported (based on the TCMP) served as the IRS measure of noncompliance.[2]

This previous survey-based approach provided useful information. However, the survey used in the development of past IRS estimates of the informal supplier tax gap has not been updated since the 1985-1986 period. Even if a more recent survey was available, we would still have reservations about using this methodology. Among our more serious concerns is that the accuracy of the approach depends critically on one's ability to distinguish formal from informal transactions, not only on the survey but also on tax returns. The kinds of information collected from these sources, especially from examinations of tax returns, do not seem adequate for this purpose.

Rather than rely on a dubious distinction between formal and informal transactions, we propose an alternative approach that examines the combined level of underreporting by formal and informal suppliers within those industry categories where informal activities are concentrated. More specifically, our methodology for estimating noncompliance involves comparing national survey results on self-employment earnings within selected industry categories to NRP statistics on the amounts actually reported for tax purposes. By focusing on a carefully chosen set of industry categories, we believe that the resulting estimate should encompass the vast majority of all noncompliance committed by informal suppliers, as well as any noncompliance among formal suppliers within these categories.

It should be noted that our analysis is restricted to informal vendors who receive money income for the goods and services they provide. We do not address the question of how to estimate noncompliance relating to barter income, which has been included in past IRS estimates of the informal supplier tax gap.[3]

Below, we lay out our methodology, present our results, and discuss the relative merits of our approach. We begin in the next section by identifying a reasonably comprehensive set of industry categories within which informal suppliers are likely to operate. We then describe the detailed "crosswalks" we have developed that link the relevant industry and occupation codes for these categories in the NRP database to the corresponding codes in our primary survey data source, the Current Population Survey. The Current Population Survey (CPS) data samples that we have drawn together for our analysis are summarized in the next section. In the following section, we present our raw estimates of self-employment earnings in 11 of our 12 selected industry categories based on the earnings reported by individuals in the CPS. These

estimates are adjusted in the next section to account for self-employment income that has been misreported as wages using two alternative approaches. We then present our methodology for estimating self-employment earnings within our remaining industry category (caterers and roadside vendors) based on the Consumer Expenditure Survey (CES). We rely on the CES to develop estimates for this industry category because the industry and occupation codes in the CPS for this category are unacceptably broad. We conclude with a comparison of our results to those of the previous survey-based approach, and a discussion of the relative merits of our methodology.

Selection of Industry and Occupation Categories

Based on our review of the existing literature on goods and services provided by informal suppliers, we have identified those industries that we believe account for the vast majority of informal supplier activities. As indicated in Table 1, our list contains 12 broad industry categories.

Each of the broad industry categories listed in Table 1 is associated with a detailed listing of specific industries and occupations. To implement our methodology, it was necessary to develop "crosswalks" that link the detailed NRP industry and occupation codes associated with our 12 broad categories to the comparable industry and occupation codes used in our primary survey data source (the CPS).

For each Schedule C (self-employment) return, the NRP identifies the industry category for the business using the North American Industry Classification System (NAICS). We carefully reviewed the NAICS codes and identified those that were relevant to each of the 12 industry categories described in Table 1. Our primary source of survey information on earnings is the Current Population Survey. Unfortunately, the 2002 CPS (which corresponds to Calendar Year 2001, or the NRP year) does not rely on NAICS.[4] Rather, a different industry coding system is used. We have therefore developed a crosswalk between the relevant NAICS codes and the corresponding 2002 CPS industry codes.

Our focus is on the discrepancy between the earned and reported amounts of self-employment income that fall into our 12 broad industry categories. However, past research has indicated that some individuals misreport their self-employment income as wages. It is therefore important to be able to identify reported wages that are attributable to our selected industry categories. For each primary taxpayer who reports wage and salary earnings, the NRP contains a detailed occupation code that describes the taxpayer's main occupation. Unfortunately, no occupation code is available for the secondary taxpayer on a joint return. A later section describes how we address this

issue. Our 2002 CPS sample also contains a detailed occupation code for wage earners; however, it is based on a different coding system. We have therefore developed a crosswalk between the relevant NRP and CPS occupation codes for each of our 12 industry categories.

To facilitate a comparison of our methodology for estimating the informal supplier tax gap with the approach used in earlier tax gap reports, we have also developed crosswalks between the 1986 and 2002 CPS occupation and industry codes. In addition, we have developed crosswalks between the 2002 and 2003 CPS occupation and industry codes to aid in the future development of estimates for Tax Year 2002.

Table 1. Key Industry/Occupation Categories for Informal Suppliers

1. Food Catering and Roadside Stands
2. Direct Sales
3. Building Maintenance/Landscaping
4. Forestry, Fishing, Hunting, and Trapping
5. Arts and Entertainment
6. Construction
7. Teaching/Lessons
8. Care of Children and Elderly (Including Home Health Services)
9. Personal Services
10. Auto Repair and Maintenance
11. Other Repair and Maintenance
12. Transportation and Moving

Description of CPS Data Sources

We employ data from the Current Population Survey (CPS) to estimate earnings within all but one of our 12 broad industry categories. The exception is the food catering and roadside stand category. We have determined that the industry and occupation codes associated with this particular category are unacceptably broad for the purposes of our analysis. For this reason, we rely on the Consumer Expenditure Survey (CES) to estimate the earnings of caterers and roadside vendors. Our approach for this category is described later.

CPS Annual Demographic File

Our primary CPS data source is the March 2002 Annual Demographic File (ADF). This file contains detailed microlevel demographic, employment, and income information for some 217,000 individuals belonging to a stratified random sample of approximately 78,000 households from across the U.S.

The file includes codes describing the industry and occupation of the individual's current job (as of approximately March 2002), as well as of his or her longest job in 2001. It also contains a detailed breakdown of annual 2001 earnings by source (e.g., wages and salaries, nonfarm self-employment, or farm self-employment). Separate earnings figures are provided for the individual's longest job and for all other jobs he or she held in 2001.[5] A code on the file identifies whether the individual was an unincorporated sole proprietor. In addition, researchers at the U.S. Census Bureau have imputed Tax Year 2001 Federal filing status and other tax information onto the file using the comprehensive income and demographic information contained in the survey. Sample weights are available to make statistics computed from the survey representative of the general U.S. population in 2001 so that they can be compared with figures from the NRP, which also covers the 2001 period.[6]

Supplemental CPS Data Sources

Using the ADF, it is possible to identify individuals with self-employment earnings from a longest job in one of our selected industry categories. What is more challenging, however, is to identify individuals who held a second job in one of these categories. A respondent may have held a second job either because he or she changed jobs during the year or because he or she "moonlighted" (held down more than one job at the same time).

To identify job changers, we compare the industry code for the job reported at the time of the interview (around March 2002) to the code for the longest job held in 2001. To identify moonlighters, we rely on supplemental information from relevant monthly CPS surveys. Of particular interest, the March 2002 CPS file contains supplementary information for each member of a large subsample of the March 2002 ADF, specifically, for 156,821 individuals from 55,498 households.[7] Further, when properly weighted, this subsample of the ADF is representative of the overall U.S. population, just like the entire ADF.

Individuals in the monthly CPS files are sampled for four consecutive months before rotating out of the sample. For the outgoing rotation group, which constitutes roughly 25 percent of the overall sample, the monthly file identifies the industry and occupation codes not only for the individual's main job, but also for his or her second job (if any). The coding system is the same as that used for the ADF. Therefore, the March file contains the desired supplementary information about an individual's second job (if any) for roughly one-fourth of the 156,821 individuals in the representative ADF subsample. In principle, supplementary information for another fourth of the ADF subsample should be available in each of the monthly CPS files from April to June, so that

details on the second job (if any) can be obtained for essentially all members of the ADF subsample.

In practice, however, we were only able to match information for 130,558 individuals, or 83.3 percent of the ADF subsample. This is largely due to sample attrition, whereby certain individuals dropped out of the sample prematurely (for example, because they changed their residences). The monthly CPS files also do not contain a unique code that can be used to definitively link individuals to their records in the ADF. Rather, a set of variables common to the monthly files and the ADF were used to match individual records.[8] Although our matching procedure works well, it is not perfect, so that some individuals who are present on both the monthly file and the ADF may not be successfully matched. It was therefore necessary to adjust the sample weights to make our matched ADF subsample of 130,558 individuals broadly representative of the overall U.S. population.[9]

CPS Measure of Reported Self-Employment Income

We summarize below our CPS-based methodology for developing a raw estimate of aggregate net 2001 self-employment income among unincorporated sole proprietors in our 11 broad industry categories. Later, we adjust this estimate to account for self-employment earnings that were erroneously reported on the CPS as wages. Our analysis is restricted to individuals who, on the basis of their reported information, appear to have had a Federal income tax filing obligation for Tax Year 2001.[10]

Using the ADF, we are able to identify cases where an individual reports self-employment earnings in 2001 from a longest job that falls into one of our 11 selected industry categories. We are also able to determine whether an individual reports self-employment income from a second job in 2001. Unfortunately, however, no details are available about the industry or occupation associated with the second job; so, we cannot determine whether it belongs to one of our selected industry categories. As discussed earlier, we instead rely on industry codes for second jobs that were held at a somewhat later date (e.g., at the time of an interview conducted between March and June 2002).

To estimate 2001 net self-employment income in the case of a longest job, we rely directly on the earnings reported in the ADF. In the case of a second job, however, it is necessary to impute earnings. Among all ADF respondents who report self-employment earnings from a second job in 2001, we find that the ratio of self-employment earnings from the second job to earnings from the longest job is 26.5 percent. To impute self-employment earnings from a second job in one of our selected industry categories (among those individuals who reported having a second job at the time of the 2002

monthly interview), we apply this ratio to the earnings reported for their longest jobs held during 2001.

The results of our analysis are summarized in Table 2. Aggregate reported net self-employment earnings from a longest job in our 11 selected industry categories amounted to $97.40 billion. Estimated self-employment earnings from a second job (for job changers and moonlighters combined) contributed an additional $8.60 billion, for a total of $106.00 billion.

Table 2. Aggregate Net Self-employment Earnings for 11 Industry Categories (Excluding Food Catering and Roadside Stands)

Source of Earnings	**Amount ($ billions)**
Longest Job	97.40
Second Job:	
Job Changers	2.85
Moonlighters	5.75
Total	106.00
Total Earnings Misclassified--Approach 1	7.54
Total Earnings Misclassified--Approach 2	46.34
Aggregate Net Self-employment Earnings for 11 Industry Categories--Approach 1	1,253.11[a]
Aggregate Net Self-employment Earnings for 11 Industry Categories--Approach 2	152.34

[a] Note that the Aggregate Net Self-Employment Earnings for 11 Industry Categories--Approach 1 includes estimated wages on returns where the primary filer's occupation falls with the 11 industry categories, an amount equal to $1,139.57 billion.

Accounting for Misclassified Earnings

Roemer (2002) presents evidence that some individuals in the CPS misreport their self-employment earnings as wages. In particular, he identifies a non-trivial number of cases where a CPS respondent reports earning wages from his longest job but where matched administrative records from the Social Security Administration show only self-employment income for the respondent. Further, he finds that the problem is especially pronounced among many of the occupations within our selected industry categories. Averaging his tabulated results over the relevant occupations, we estimate that 4.1 per-

cent of CPS-reported wages in our selected industry categories are actually misclassified self-employment earnings.

We use two alternative approaches to account for misclassified self-employment income within our selected industry categories. For each approach, we restrict our attention to filers in the CPS who appear to have had a legal filing obligation for Tax Year 2001.

First Approach

Our first approach to account for self-employment income from our 11 selected industry categories that has been misclassified as wages involves distinct treatments depending on whether the primary taxpayer in a household reports wage earnings in the CPS from a relevant longest job.

To identify whether an individual in the CPS is a primary filer, we begin by examining the tax filing status that has been assigned. If the individual is assigned a status other than married filing a joint return, we treat the individual as a primary filer. In cases where a couple is assigned married joint filing status, we treat the male member of the couple as the primary filer so long as he reports nonzero earnings for 2001. Conversely, if the male reports zero earnings and his spouse reports nonzero earnings, we treat his spouse as the primary filer.

In cases where a primary filer in the CPS reports wage earnings from a longest job in one of our 11 selected industry categories, we tabulate the total amount of wages reported from all jobs. If the individual has been assigned married joint filing status, this tabulation includes any wages earned by the secondary filer. Using this approach, we obtain an aggregate estimate of $1,139.57 billion in reported wages. This CPS-based wage figure is intended for comparison with the aggregate amount of wages reported in the NRP on returns with a primary taxpayer occupation code that falls into one of our 11 selected industry categories. By extending our analysis to consider the discrepancy between the CPS and NRP measures of both self-employment income and wage income for these cases, the measure of noncompliance is unaffected by any misclassification of self-employment earnings as wages.

Since the NRP contains only an occupation code for the primary filer's main job, it is not possible to perform a similar comparison of CPS and NRP wage measures in cases where reported wages from a relevant industry category are solely attributable to a second job held by a primary filer. It is also not possible to perform a comparison when the wages are attributable to work performed by a secondary filer. In each of these situations, we develop an explicit estimate of the aggregate amount of self-employment income that has been misreported as wages in the CPS for our 11 selected industry categories.

We begin by developing an estimate of the total reported wages within these categories that are attributable to the following sources:

- second jobs held by primary filers (both job changers and moonlighters);
- the longest job held by secondary filers; and
- second jobs held by secondary filers (both job changers and moonlighters).

As discussed previously, we estimate that 4.1 percent of the reported wages in each of these categories is actually misclassified self-employment earnings. Excluding cases where the primary filer reports wages from a relevant longest job, we estimate that $7.54 billion in net self-employment earnings from our 11 selected industry categories were misclassified as wages in the CPS (Table 2).

We then use this estimate of misclassified earnings ($7.54 billion) to develop an adjusted CPS-based estimate of aggregate net self-employment income for our 11 selected industry categories. This estimate is $113.54 billion (or $106.00 billion plus $7.54 billion). We add to this total our estimate of wage earnings for cases in which the primary filer reports a longest job in one of these categories ($1,139.57 billion). Our combined estimate equals $1,253.11 billion. A comparison of this figure to the corresponding amount of self-employment and wage income actually reported in the NRP would yield our proposed measure of noncompliance. In the case of self-employment earnings, the relevant amount to include is the total net income reported on Schedule C returns for our 11 selected industry categories (including any reported wage income that was reclassified by the examiner as self-employment income). Classification into these industry categories should be based on the NAICS codes identified in Alm and Erard (2004, Appendix B). In the case of wage earnings, the relevant amount to include is the total value of wages reported on returns for which the primary taxpayer occupation code matches one of the various codes listed in Alm and Erard (2004, Appendix B) for our 11 selected industry categories.

We then apply our estimate that 4.1 percent of the reported wages in each of these categories is actually misclassified self-employment earnings. Including those cases in the CPS where the primary taxpayer reports wages from a longest job in one of our 11 selected industry categories, we estimate that $46.34 billion in net self-employment earnings were misclassified as wages (Table 2).

Applying this result, we derive an adjusted CPS-based estimate of aggregate net self-employment earnings for our 11 selected industry categories. Our estimate equals $106.00 billion plus $46.34 billion, or $152.34 billion. A

comparison of this figure to the corresponding amount of net self-employment earnings actually reported in the NRP would yield our proposed measure of noncompliance. The NRP measure should include the total net income reported on Schedule C returns within our 11 selected industry categories, including any reported wage earnings that have been reclassified by the examiner as self-employment income. Classification into these categories should be based on the NAICS codes identified in Alm and Erard (2004, Appendix B).

Relative Merits of the Two Approaches

Our two approaches to account for the misclassification of self-employment earnings as wages differ only in one respect: the treatment of returns on which primary filers report having wage income from a longest job in one of our 11 selected industry categories. In our first approach, we circumvent the misclassification problem on these returns by expanding our analysis to include the combined amount of wages and self-employment income that has been reported. In the second approach, we tackle the problem directly by estimating the share of reported wages on these returns that is actually misclassified self-employment earnings.

The inclusion of reported wages under our first approach dramatically increases the overall amount of income that is to be compared. Whereas the CPS-based estimate under our second approach yields $152.34 billion in self-employment earnings for comparison with the NRP, our first approach yields the substantially larger figure of $1,253.11 billion in combined wages and self-employment earnings for comparison.

It is difficult to predict which approach will yield a more accurate estimate of noncompliance. The first approach is likely to be more sensitive to any errors in the aggregate reporting of wages in the CPS. In contrast, the second approach relies much more heavily on the accuracy of our estimate of the share of wages that is actually misclassified self-employment earnings. Our recommendation is to implement both approaches and compare the results. If there is a significant discrepancy, the source of that difference can be investigated, and a decision can be reached as to which result is most plausible. Generally, however, a very high percentage of wages is accurately reported on tax returns. Therefore, it may turn out that the two approaches yield a fairly similar estimate of noncompliance.[11]

Methodology for Caterers and Roadside Vendors

As noted earlier, we are not able to use the CPS to estimate self-employment earnings for one industry category, food caterers and roadside stands, be-

cause the industry and occupation codes in the CPS for this category are unacceptably broad. Instead, we estimate the gross receipts of vendors within this category based on tabulations from the Consumer Expenditure Survey (CES). We then propose a methodology for converting this estimate of gross receipts into a measure of net self-employment income using information from the NRP.

The CES is conducted by the Bureau of Labor Statistics in the U.S. Department of Labor, and it provides detailed information on the expenditure patterns of American consumers (information that is also used to revise the Consumer Price Index). The survey consists of two separate components: a quarterly "Interview Survey" in which each consumer unit in the sample is interviewed every 3 months over a 15-month period, and a "Diary Survey" completed by a subsample of consumer units for two consecutive 1-week periods. We rely on the Interview Survey for our analysis. It has the advantage of including a much larger sample of respondents who report purchases from caterers and roadside vendors. The Bureau of Labor Statistics estimates that 90 percent to 95 percent of total consumer expenditures are covered in this survey.

Our estimate of consumer expenditures on food catering and roadside stands is based on the Detailed Expenditure Files (DEF) from the 2001 CES Interview Survey. Included under "Miscellaneous Expenses" in the DEF are expenditures made on "Catered Affairs." Similarly, under "Expense Patterns for Food, Beverages, and Other Selected Items", consumer expenditures on vegetable stands and farmers' markets are included.[12] On an annual basis, 2001 total expenditures were $4.13 billion for food caterers and $1.68 billion for roadside stands, for a combined total of $5.81 billion.

It is worth noting that the above estimates of expenditures on caterers and roadside stands are comparable to those from other, independent sources. For example, the National Restaurant Association (2004) estimated that 2001 expenditures on social and mobile caterers totaled $4.8 billion. Similarly, the U.S. Department of Agriculture (2004) estimated that the "value of agricultural products sold directly to individuals for human consumption" was $812 million in 2002.

Our combined CES estimate of $5.81 billion provides a measure of the gross receipts of food caterers and roadside vendors, not their net incomes. Therefore, it will be necessary to adjust this figure to account for legitimate expenses that could be claimed against these receipts. Under the earlier methodology for estimating the informal supplier tax gap, the IRS assumed that net informal supplier income represents 51 percent of gross receipts. However, as we have emphasized, we have reservations about this assumption. As an

alternative, we propose using relevant information from the NRP to convert our measure of gross receipts into an estimate of net income.

Under our proposed approach, one would begin by computing the aggregate expenses allowed by the NRP examiner on Schedule C returns falling under the relevant NAICS codes. We follow IRS practice here in assuming that the "per exam" figure for expenses is accurate. One would then reduce our estimate of $5.81 billion in gross receipts by an appropriate amount to account for the receipts attributable to nonfilers. The ratio of net income to gross receipts would then be computed as one minus the ratio of allowable expenses to our adjusted measure of gross receipts. This ratio would then be multiplied by $5.81 billion to arrive at an estimate of the net income that should have been reported by caterers and roadside vendors. The difference between this figure and the amount actually reported (based on the NRP) would represent our proposed measure of noncompliance.

We recommend adding the aggregate estimate of noncompliance among caterers and roadside vendors based on the above approach to the aggregate estimate of noncompliance within our other 11 selected industry categories to arrive at an overall measure of noncompliance for the 12 categories combined. A tax calculator can be employed to assess the amount of income tax owing on the underreported self-employment income.

Comparisons with Earlier Results

It is instructive to compare results from the previous University of Michigan survey for 1985-1986 with results from applying an abbreviated version of our proposed methodology to 1986 CPS data for nine fairly comparable industry categories. Our hope is that this provides at least a rough indication of how our methodology for estimating the informal supplier tax gap compares with the one used in past IRS tax gap reports.

Excluding the food and street vendor categories of goods and services, the estimated gross receipts for informal suppliers based on the University of Michigan consumer survey amounted to $65.9 billion in 1985, including an estimated $24.9 billion in business purchases from informal vendors. The IRS assumed that net self-employment earnings were 51 percent of gross receipts in this year. Applying this assumption yields an aggregate estimated $33.6 billion in net informal supplier income.

Based on results from the same survey, McCrohan, Smith, and Adams (1991) report that estimated gross receipts of formal vendors from sales of the identical types of goods and services amounted to $135.1 billion in that year. This figure does not include business purchases. If one assumes that

the ratio of business purchases to household purchases was the same for formal and informal vendors, accounting for business purchases raises the estimate to $186.1 billion. Applying the same IRS assumption that net self-employment earnings are 51 percent of gross receipts yields an aggregate estimated $94.9 billion in net formal supplier income for 1985.

Combining the above results, aggregate estimated net income among formal and informal suppliers within the selected goods and services categories based on the University of Michigan survey approach amounted to $128.5 billion in 1985.

We have applied an abbreviated version of our proposed CPS-based methodology to estimate net self-employment earnings within nine of our selected industry categories for 1985.[13] These industry categories correspond at least roughly to the goods and services covered in the University of Michigan analysis.

Based on the 1986 CPS, approximately $39 billion in net self-employment earnings were reported by individuals who had a longest job in one of the nine selected industry categories in 1985. To account for self-employment earnings from a second job in one of these categories, we assume that the ratio of total earnings to longest job earnings presented in Table 2 for 2001 also applies to 1985. To account for self-employment income that has been misclassified as wages in the CPS, we assume that the ratio of our adjusted estimate to our raw estimate for 2001 also applies to 1985. After allowing for these adjustments, our estimate of aggregate net self-employment income within the nine selected industry categories in 1985 amounts to $61 billion.

Our estimate of $61 billion in aggregate net self-employment earnings includes earnings from both formal and informal sole proprietors in the nine selected industry categories. While our estimate should therefore exceed the net earnings of informal suppliers in these categories, it should fall short of the combined net income of all formal and informal suppliers because many of the formal sales in these industry categories are presumably attributable to sales by partnerships and corporations, which are not accounted for in our estimate.

In fact, our estimate of $61 billion for 1985 substantially exceeds the estimated $33.6 billion in net informal supplier income based on the University of Michigan approach, while falling well short of the estimated $128.5 billion in combined formal and informal supplier net earnings based on the Michigan approach. It therefore appears that our estimate of net earnings within the selected industry categories for 1985 falls within a plausible range. Given the nature of the problem and the available data, this is perhaps the most assurance one can reasonably hope to attain prior to actually applying our methodology to data from the NRP.

Conclusions: Relative Merits of our Proposed Methodology

Our proposed methodology has a number of advantages over the methodology used in the development of previous estimates of the informal supplier tax gap. These include:

- the survey information used in our approach is publicly available, meaning that no special surveys need to be commissioned;
- the number of respondents to the surveys used in our approach is much larger than the number of respondents to the special surveys used in the earlier methodology;
- the thorny issue of distinguishing informal suppliers from formal suppliers is avoided;
- sales to both consumers and businesses are fully accounted for;
- with the exception of the catering and roadside vendor industry category, the approach provides a direct estimate of net earnings, thereby avoiding the need to rely on assumptions about the relationship between net income and gross receipts; and
- detailed crosswalks have been developed that provide a tight linkage between the coding used in the surveys for selected industries and occupations and the coding used in the NRP.

The chief disadvantage of our methodology is that it relies on the accuracy of income information reported in the CPS by individuals who operate businesses in our selected industry categories. As a group, these individuals may be more willing to provide an accurate accounting of their incomes on an independently administered and confidential survey than they would on their tax returns. Nevertheless, it is entirely possible that the amounts reported on the CPS fall somewhat short of true earnings, in which case our methodology would tend to underestimate the informal supplier tax gap.

The evidence on this issue is reasonably encouraging. In particular, Roemer (2002) examines an exact match between earnings reported by respondents in the CPS and their Detailed Earnings Records (DER) from the Social Security Administration. He finds evidence that many respondents, particularly those employed in informal occupations, report earnings on the

CPS that are not recorded in the DER. He interprets this as evidence that the CPS measure of income includes underground earnings. Indeed, he concludes that the CPS picks up more of these earnings than another commonly used survey, the Survey of Income and Program Participation.

Endnotes

[1] This paper is an abridged version of our full report, "Development of a Methodology for Estimating the Informal Supplier Tax Gap," Internal Revenue Service Order Number TIRNO-03-P-00651, September 23, 2004. The full report contains more detailed information on our methodology and its application.

[2] This served as a measure of unreported self-employment income, and an IRS tax calculator converted this measure into an estimate of unreported taxes.

[3] For instance, estimated barter income represented $7.3 billion of the estimated $62.1 billion in net informal supplier income in Tax Year 1988 (Internal Revenue Service, 1996, p. 45.)

[4] As of 2003, the CPS has adopted NAICS for industry classification.

[5] The earnings information in the CPS is top-coded for individuals who have high levels of income. We assume that the earnings of most informal suppliers fall below the relevant threshold. Therefore, we make no adjustment to account for top-coding in our analysis.

[6] More specifically, the ADF universe is the civilian noninstitutional population of the United States living in housing units and members of the Armed Forces living in civilian housing units on a military base or in a household not on a military base.

[7] In addition to the March monthly CPS sample of households, the 2001 ADF contains supplemental samples to improve the accuracy of statistics on Hispanics as well as State-level estimates of children's health insurance coverage. By applying the appropriate sample weights, statistics from either the March monthly sample or the full ADF can be made representative of the overall U.S. population.

[8] The variables used for matching include the household identification number, the person line number, gender, and age. When matching the March monthly sample to the ADF, we also compared the recorded values of the current industry codes.

[9] It was possible to match essentially all of the outgoing rotation groups from the March monthly sample to the ADF. For each subsequent rotation group from April to June, the group sample weights were proportionally adjusted upwards to account for members of the group who were not successfully matched, either because of attrition or imperfections in the matching criteria. There was a small discrepancy in the aggregate weighted populations between the ADF and the March monthly file (282.1 million compared to 278.1 million). Therefore, a small final proportional adjustment was applied to all matched monthly records (multiplication by 1.014) to make the weighted population total equal to the corresponding ADF total.

[10] All households coded on the CPS with a filing status other than "nonfiler" were assumed to have a Federal income tax filing obligation. This included, among others, all individuals with more than $400 in reported self-employment income.

[11] Under our second approach, noncompliance with respect to the reporting of wages would be accounted for separately through a comparison of the per exam and per return amounts of reported wages (after accounting for wages that have been misclassified self-employment earnings) on the relevant returns. Since unreported wages are relatively easy to detect, our expectation is that the combined value of this estimate for underreported wages and the estimate for underreported self-employment earnings from our second approach should be fairly similar to the estimate of noncompliance from our first approach.

[12] The exact description of the expenditure variable is "Quarterly expenditure for food or nonalcoholic beverages from places other than grocery stores, such as home delivery, specialty stores, bakeries, convenience stores, dairy stores, vegetable stands, or farmers' markets."

[13] Our CPS analysis excludes the transportation and moving and arts and entertainment categories, as well as the caterers' and roadside vendors' category that is addressed in our methodology using results from the CES.

References

Alm, James and Erard, Brian (2004), "Development of a Methodology for Estimating the Informal Supplier Tax Gap," Internal Revenue Service Order Number TIRNO-03-P-00651 (September 23, 2004).

Internal Revenue Service (1996), "Federal Tax Compliance Research: Individual Income Tax Gap Estimates for 1985, 1988, and 1992," Internal Revenue Service Publication 1415, Washington, D.C.

Internal Revenue Service (1988), "Income Tax Compliance Research, Supporting Appendices to Publication 7285," Internal Revenue Service Publication 1415, Washington, D.C.

McCrohan, K. F.; Smith, J. D.; and Adams, T. K. (1991), "The Cyclical Nature of Household Purchases in Informal Markets," *European Journal of Marketing* 25 (7), pp. 22-40.

National Restaurant Association (2004), "2004 Restaurant Industry Forecast, Executive Summary," Washington, D.C. (www.restaurant.org).

Roemer, M. (2002), "Using Administrative Earnings Records To Assess Wage Data Quality in the March Current Population Survey and the Survey of Income and Program Participation" U.S. Census Bureau Staff paper, Washington, D.C.

Smith, J.D. and Adams, T. K. (1987), "The Measurement of Selected Income Flows in Informal Markets, 1981 and 1985-1986". Report prepared for the Internal Revenue Service, Contract Number TIR 84-0084, Washington, D.C.

United States Department of Agriculture (2004), "2002 Census of Agriculture," Washington, D.C. (www.usda.gov/nass/).