International Studies Program

International Studies Program
Working Paper 03-19
December 2003

Mapping the Compliance Continuum: From Pathologically Honest to Flagrantly Defiant

Brian Erard
Chih-Chin Ho

Georgia State University
Andrew Young School of Policy Studies

International Studies Program
Working Paper 03-19

Mapping the Compliance Continuum: From Pathologically Honest to Flagrantly Defiant

Brian Erard
Chih-Chin Ho

December 2003

International Studies Program
Andrew Young School of Policy Studies
Georgia State University
Atlanta, Georgia 30303
United States of America

Phone: (404) 651-1144
Fax: (404) 651-4449
Email: ispaysps@gsu.edu
Internet: http://isp-aysps.gsu.edu

International Studies Program
Andrew Young School of Policy Studies

The Andrew Young School of Policy Studies was established at Georgia State University with the objective of promoting excellence in the design, implementation, and evaluation of public policy. In addition to two academic departments (economics and public administration), the Andrew Young School houses seven leading research centers and policy programs, including the International Studies Program.

The mission of the International Studies Program is to provide academic and professional training, applied research, and technical assistance in support of sound public policy and sustainable economic growth in developing and transitional economies.

The International Studies Program at the Andrew Young School of Policy Studies is recognized worldwide for its efforts in support of economic and public policy reforms through technical assistance and training around the world. This reputation has been built serving a diverse client base, including the World Bank, the U.S. Agency for International Development (USAID), the United Nations Development Programme (UNDP), finance ministries, government organizations, legislative bodies and private sector institutions.

The success of the International Studies Program reflects the breadth and depth of the in-house technical expertise that the International Studies Program can draw upon. The Andrew Young School's faculty are leading experts in economics and public policy and have authored books, published in major academic and technical journals, and have extensive experience in designing and implementing technical assistance and training programs. Andrew Young School faculty have been active in policy reform in over 40countries around the world. Our technical assistance strategy is not to merely provide technical prescriptions for policy reform, but to engage in a collaborative effort with the host government and donor agency to identify and analyze the issues at hand, arrive at policy solutions and implement reforms.

The International Studies Program specializes in four broad policy areas:

- Fiscal policy, including tax reforms, public expenditure reviews, tax administration reform
- Fiscal decentralization, including fiscal decentralization reforms, design of intergovernmental transfer systems, urban government finance
- Budgeting and fiscal management, including local government budgeting, performance-based budgeting, capital budgeting, multi-year budgeting
- Economic analysis and revenue forecasting, including micro-simulation, time series forecasting,

For more information about our technical assistance activities and training programs, please visit our website at http://isp-aysps.gsu.edu or contact us by email at ispaysps@gsu.edu.

Mapping the Compliance Continuum:
From Pathologically Honest to Flagrantly Defiant

Brian Erard
B. Erard & Associates
2350 Swaps Court
Reston, VA 20191
BEandAssoc@Aol.com
(703) 390-9368

and

Chih-Chin Ho
Internal Revenue Service
1111 Constitution Ave, NW
Washington, DC 20001
Chih-Chin.Ho@IRS.gov
(202) 874-0509

Prepared for conference on "The Hard to Tax: International Perspective," Andrew Young School of Policy Studies, International Tax Program, Georgia State University, Stone Mountain Park, May 15-16, 2003.

Note: Any opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Internal Revenue Service.

Mapping the Compliance Continuum:
From Pathologically Honest to Flagrantly Defiant

Brian Erard and Chih-Chin Ho

1. Introduction

There are by now vast academic literatures on both tax compliance and the underground economy. These literatures provide estimates of the overall degrees of non-compliance and shadow (or hidden) activity as well as numerous insights into their causes and consequences.[1] However, we believe it is fair to say that they provide an incomplete perspective on the characteristics of the individuals or groups who engage in such behaviors. Methods for evaluating shadow activity are frequently based on indirect measures, such as discrepancies in national account or labor force statistics, trends in the demand for currency or in monetary transactions, or variations over time in national

A more promising approach in the tax compliance literature has been the analysis of tax reporting behavior based on samples of audited tax returns.[2] Arguably, the best such samples have been generated by the U.S. Internal Revenue Service (IRS) through its Taxpayer Compliance Measurement Program (TCMP). Under this program, which operated through tax year 1988, a stratified random sample of filers was selected once every few years and subjected to intensive line-by-line audits of their federal individual income tax returns by experienced tax examiners. The line item return information originally reported by the taxpayer as well as the amounts that, in the examiner's judgment, should have been reported were both recorded in the TCMP data file. Research based on these TCMP samples has provided important insights into the characteristics of filers who underreport their taxes. In some instances, this research has even attempted to account for forms of non-compliance that have escaped detection during the audits.[3] Yet, while these studies have cast considerable light on the reporting compliance problem in the U.S., they do not reveal a complete picture of the compliance landscape, because many households who fail to pay their taxes also fail to file a tax return. Indeed, much of focus of the extensive underground economy literature is on shadow activities that are not captured in official records, such as national income accounts and tax returns.

Recently, Erard and Ho (2001) have provided some evidence on the characteristics of nonfilers and the determinants of the decision whether to file a tax return based on their analysis of a special TCMP study of U.S. households that failed to file a federal income tax return in tax year 1988. In this paper, we report on a micro-simulation data base we have developed using information from that TCMP study, a comparable TCMP study of filers in the same tax year, and supplementary information on tip earners and "informal suppliers". Although this data base would benefit from further refinement, particularly with respect to the imputation and allocation of certain forms of income, we believe it has the potential to provide a more complete depiction of compliance than has heretofore been available—one which accounts both for individuals who file returns but understate their taxes and individuals who neither file a return nor pay all of the taxes that they owe. As an illustration, we use our data base to develop a preliminary map of where members of 34 distinct occupational groups fall along a tax compliance continuum that ranges from "pathologically" honest to flagrantly defiant. We then explore some possible explanations for the variations in compliance behavior across occupations.

The remainder of our paper is organized as follows. Section 2 summarizes the key data elements underlying our micro-simulation data base. Section 3 describes our methodology for imputing certain forms of income to individual filers and nonfilers to account for income which has gone undetected during examinations. Section 4 lays out our preliminary map of the compliance continuum that is based on our micro-simulation model, and Section 5 presents some evidence on possible reasons for variations in compliance behavior by occupational group. Section 6 concludes.

2. Data sources

The core elements of our micro-simulation data base are derived from two separate TCMP studies that were conducted for tax year 1988, one for filers and another for nonfilers. Although these data are now some 15 years old, they have the advantage of providing detailed compliance information about both filers and nonfilers for a common tax year. Moreover, although the magnitude and composition of tax noncompliance are likely to have changed since these data were compiled, in response to changes in tax rules, economic conditions, and social factors, we believe that the data remain informative about the fundamental nature of the compliance decision and the factors associated with non-compliance.

2.1 TCMP filer data

The data for filers of 1988 federal income tax returns are taken from the IRS TCMP Phase III Survey. This survey contains the results of intensive line-by-line audits of a stratified random sample of approximately 54,000 individual income tax returns for tax year 1988. For most line items both the amount that was reported by the filer and the amount that the examiner determined should have been reported are available. For income items, changes assessed by the examiner to the amount originally reported by the taxpayer are broken down according to whether the change was based on a review information return documents or if it was based on other information. As discussed below in Section 3, this distinction is useful for purposes of imputing additional non-detected income to taxpayer returns. Information is also recorded about the prior filing history of the household, and a code is available for the primary filer's occupational category. The occupation code has been recorded by the IRS examiner, based on his assessment of the filer's main line of work. A set of sample weights is included to make the data representative of the national return population.[4]

2.2 TCMP nonfiler data

Our data on nonfilers comes from the examination-based segment of the IRS TCMP Phase IX Nonfiler Survey. The special TCMP study began with a stratified random sample of 23,283 potential nonfilers from a population of 83 million individuals for whom their was no record of a 1988 individual income tax return being filed.[5] Revenue officers set out to locate each of the individuals in this sample to determine whether they should have filed an individual income tax return for tax year 1988.[6] A total of 18,689 of the 23,283 potential nonfilers were successfully located through the search process. The revenue officers had access to information documents and past filing records. Armed with this information they conducted interviews or field visits to determine whether a successfully located individual was required to file a return; i.e., whether the potential nonfiler was a "true nonfiler". Tax returns were secured from 3,546 individuals who were deemed to have been in violation of their tax filing requirements, and a random sample of 2,195 of these returns were subjected to intensive line-by-line audits, comparable to the audits performed for the TCMP Phase III study of individual return filers. It is the details from these 2,195 examined returns that we include in our micro-simulation data base.

Since not all potential nonfilers in the original sample of 23,283 were located, it is highly likely that a number of true nonfilers went unidentified.[7] We have therefore modified the sample weights for our sample of 2,195 located true nonfilers to make these individuals broadly representative of all true nonfilers. To do this, we followed the same approach we used previously in our development of the official IRS estimate of the nonfiling tax gap from these data (Internal Revenue Service, 1996).[8] The first step is to perform a probit analysis of the likelihood that a potential nonfiler can be located. The probit equation takes the form

$$L^* = \beta_L^{'} X_L + \varepsilon_L , \quad (1)$$

where X_L represents a vector of regressors based on the information that was available to the revenue officer who attempted to locate the individual. Depending on the individual, information may have been available about the individual's age, whether a return had been filed for previous tax years, details concerning the individual's spouse, and details from information return documents. The parameter vector β_L represents the coefficients to be estimated, and ε_L is a standard normal disturbance term. The coefficients are estimated by the method of maximum likelihood and used to predict the probability that each individual can be located. To make located true nonfilers broadly representative of all true nonfilers, their original sample weights are divided by their predicted chance of being located.[9] The interested reader is referred to Internal Revenue Service (1996, Appendix A) for further details on this approach.

2.3 Combined sample

To develop our core data base, we merged together the detailed information (both per return and per exam) from the TCMP filer and non-filer data files. When weighted, the combined sample represents an estimated population of 112.3 million households, including 104.3 million filers and 9 million nonfilers. Our data base includes an imputed variable meant to approximate the burden associated preparing and filing a tax return for each of the households in our sample. Our measure is based on an IRS formula for the average time burden, in hours, for an individual whose return contains a particular set of forms and schedules.[10]

3. Imputation of undetected noncompliance

Even intensive examinations such as those conducted under the TCMP can fail to uncover significant amounts of noncompliance. To account for undetected noncompliance, we follow a procedure similar to that employed by the IRS to generate its official estimates of the individual income tax gap – the difference between the amount of income that households owe and the amount they voluntarily pay in a timely manner.

3.1 General approach

Several key income items, such as wages and interest, are largely subject to third party information reporting. With the aid of information returns, examiners have relatively little difficulty uncovering amounts of these items that have gone undeclared. However, for those items not subject to information reporting, non-detection can be a serious problem. Based on an earlier TCMP study, the IRS determined that examiners were typically able to identify only slightly less than one-third of undeclared income amounts when they did not have access to information returns. We follow the IRS in assuming that for every dollar of undeclared income detected on most line items without the aid of an information return, there is another $2.28 that has gone undetected by the examiner. For filers, our data base includes a breakdown of the portions of undeclared income on each line item detected with and without the aid of information returns, making it straightforward to apply this procedure. However, for the returns in our data base that were secured from nonfilers, this breakdown is not available. We therefore assume that the percentage of undeclared income detected with the aid of information returns on a given line item is the same for non-filers as it is for filers. Like the IRS, we assume that all undeclared income subject to information reporting has been fully detected. Two key exceptions to this general approach for imputing undeclared income are the treatment of undeclared tip income and undeclared "informal supplier" income. The treatment of these items is discussed below.

3.2 Undeclared tip income

The IRS commissioned the Bureau of Economic Analysis (BEA) to conduct a special study of unreported tip income earned by filers and nonfilers of federal income tax returns. For tax year 1988, the BEA estimated that filers reported $5.906 billion in tips on their returns, understating their true tip income by $11.617 billion. In addition, the BEA estimated that non-filers received $532 million in tips.[11]

Unfortunately, the information in our data base does not include a separate line item tip income. Rather, tip income is merged with wages and salaries. It was therefore necessary to identify households who appeared likely to receive tip income on the basis of their occupation codes and allocate a portion of their combined wages, salaries and tips to tip income.[12] We assumed that those employed in occupations involving food and beverage preparation and service, personal services (barbers, hairdressers, guides, ushers, porters, bellhops, shoe shiners, etc.), or certain forms of transportation (taxicab, bus, or limousine) were all tip earners. The weighted number of filers in these occupations totaled 4.38 million. To allocate reported tips among these households, we assumed that reported tips represented a fixed proportion of overall reported wages, salaries, and tips.[13] The proportion (15.54 percent) was chosen so that the total amount of reported tips would be equal to the BEA's estimate of $5.906 billion.

Essentially all undeclared wage, salary, and tip income detected with the aid of information returns is attributable to undeclared wages and salaries. We therefore allocated all such income to detected wages and salaries. However, based on a prior TCMP study, only approximately 35 percent of all undeclared wages, salaries, and tips detected without the aid of information returns is attributable to undeclared wages and salaries. We therefore allocated 35 percent of such income to detected wages and salaries and assumed the remainder represented undeclared tip income identified by the examiner during the audit. The portion of detected income assigned to undeclared wages and salaries was multiplied by 3.28 to account for the general assumption that income not declared on third party information returns is only partially detected during the examination. On the other hand, the portion of detected income assigned to undeclared tips for each household was "topped off" so that the aggregate expanded amount was equal to the BEA's estimate for undeclared tips of $11.617 billion. With little information to guide the allocation of additional tip income, we simply assumed that each of the 4.38 million filers had an additional $2,654 ($11.617 billion divided by 4.38 million) of undeclared tip income that escaped detection during the audit. We employed a comparable procedure for imputing reported and undeclared tip income based on the secured returns of the nonfilers in our database.

3.3 Undeclared informal supplier income

The 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." (Internal Revenue Service, 1996, p. 43)

Examples include self-employed domestic workers, street-side vendors, and moonlighting tradesmen. Conceptually, the informal economy within such individuals operate includes all types of market economic activity that are potentially under-measured in the National Accounts owing to the vendors' informal business style (sales in cash, lack of adequate records of sales and purchases, etc.) Since the detection of noncompliance among such individuals is likely to be especially difficult, the IRS commissioned the Survey Research Center of University of Michigan to conduct some special studies during the 1980s to derive estimates the gross sales revenue earned by informal suppliers.[14]

3.3.1 *University of Michigan study*

It would be exceedingly difficult to derive estimates of the size of the informal economy by surveying informal suppliers about their transactions for two main reasons. First, since informal suppliers do not always comply with licensing, registration, permit, and tax filing requirements, there is no straightforward way to design a probability sample from which national estimates could be derived. Second, given that informal suppliers are not always in compliance with federal, state, and local requirements, it is doubtful that they would be completely forthcoming on a survey about their transactions in the informal economy.

Rather than attempt to interview the suppliers of goods and services in the informal economy, the University of Michigan researchers therefore elected to interview the purchasers. Specifically, they relied on telephone surveys of nationally representative samples of households in 1981, 1985, and 1986 to estimate the gross value of purchases made by consumers in the informal economy. Although the responses of such samples of consumers are likely to be both reasonably candid and statistically representative, it remained a challenge to distinguish between purchases that were made in the formal and informal economies. As detailed below in Table 1, the University of Michigan study focused on 14 broad classes of goods and services that were believed to be sold in the informal economy. Since many of these goods and services are also provided by established businesses that operate in the formal economy, supplementary information about the nature of the transaction and the characteristics of the vendor was used as a guide to infer whether the transaction took place in the formal or informal sectors. For example, child care services were assigned to the informal economy only if they were provided in the home of the family buying such care. Similarly, housekeeping services were classified as informal transactions only if the provider was not engaged or employed through a commercial cleaning firm. In some cases, it was especially difficult to determine whether a transaction took place in the formal or informal economy. For instance, the University of Michigan reported that the classification of automobile repair services was problematic owing to the lack of adequate information about the vendors' characteristics.

3.3.2 *From gross sales to net underreported income*

Based on the University of Michigan survey results, the IRS was able to develop estimates both of the aggregate overall purchases (or equivalently, sales) in the informal economy as well as the amount spent within each of the broad good and service categories. These figures were modified to exclude earnings of domestic employees (which, at least theoretically, were already captured in IRS' unreported wage estimates) as well as earnings of friends and relatives from lawn work or babysitting, who were assumed to have income below the tax filing threshold.

Like businesses in the formal sector, informal suppliers have legitimate expenses and deductions that must be taken into account when estimating the tax gap. Based on an analysis of tax year 1981 data, the IRS determined that reported net income amounted to approximately 51 percent of reported gross receipts on returns that appeared to have informal business income. The IRS applied this percentage to its estimate of gross sales based on the University of Michigan surveys to arrive at an aggregate net income figure of $62.15 billion for informal suppliers.

Some informal suppliers do report at least a portion of their net income from sales on their tax returns. To estimate the amount that was reported, the IRS developed criteria for identifying likely informal suppliers based on tax return information. Specifically, a taxpayer was designated as an informal supplier if (s)he: (1) filed a Schedule C return; (2) reported a principal industrial activity (PIA) that was closely aligned with one of the 14 categories of goods and services listed in Table 1; and (3) made no claim for certain types of business expenses (taxes, rent, insurance, etc.) that informal suppliers are not believed to typically incur.

Based on this approach, we designated a subsample of households in our data base representing 2.74 million filers and 711,566 nonfilers as informal suppliers. We divided the estimated $62.15 billion in true net informal supplier income among filers and nonfilers according to their shares of the overall informal supplier population. We assumed that all of the Schedule C (self-employment) net income reported by these households ($9.5 billion by filers and $9.7 by nonfilers) on their tax returns was attributable to informal activities. The aggregate difference between our measures of true and reported informal supplier income for each group represented our estimate of total undeclared income. For simplicity, we imputed an equal share of this estimated total to each member of the group.

3.4 Expanded estimate of tax noncompliance

Our imputations result in additional net taxable income for many households beyond that detected during the examination. We applied a simplified tax calculator to translate this additional income into additional tax liability.[15] A more elaborate algorithm was required to estimate the additional self-employment tax associated with our imputations of additional self-employment income to returns.[16]

For filers, we computed our overall measure of tax noncompliance as the difference between our expanded measure of total tax after credits (inclusive of the Earned Income Tax Credit) and the amount originally reported on the return.[17] In the case of nonfilers, our measure was the difference between our expanded measure of total tax after credits (inclusive of the Earned Income Tax Credit) and the total amount of tax that was prepaid (for instance, through withholding and estimated tax payments).

4. Mapping the compliance continuum

Within an economy, tax compliance behavior falls along a continuum. At one extreme are almost "pathologically" honest individuals, who fully report and pay their tax obligations despite any opportunities or incentives to cheat. At the other extreme are flagrantly defiant households who undertake considerable efforts to conceal their income and repudiate their tax responsibilities. Using our micro-simulation data base, we develop a preliminary map of where members of 34 distinct occupational groups fall along the U.S. compliance continuum. The members of these groups were allocated on the basis of the per exam value of the occupational classification code contained in our data base.

Tables 2 and 3 present our estimates of noncompliance by occupational category. The results presented in Table 2 are sorted by the estimated average level of noncompliance, whereas the results presented in Table 3 are sorted by the estimated ratio of noncompliance to true tax liability (referenced under the heading "% of total taxes not paid"). A comparison of the two tables indicates that certain occupational groups are responsible for large dollar values of noncompliance, but relatively small values in relation to their overall true tax liabilities. For instance, lawyers and judges rank fourth highest in terms of the average level of noncompliance, underpaying taxes by an average of $2,079 per return. However, this represents only about 8.2 percent of their overall tax liability, compared to a 14.1 percent underpayment for all occupations as a whole. Similarly, doctors and dentists rank sixth highest in terms of average dollars of noncompliance ($1,689), but third lowest in terms of the share of their overall liability that goes unpaid (5.5 percent).

Conversely, certain occupational groups rank relatively low in terms of average dollars of noncompliance, but quite high in terms of the share of tax liability that goes unpaid. For instance, individuals employed in service occupations other than those associated with tip earners, informal suppliers, or protective services understate their taxes by an estimated $359 – well below the mean of $615 for the population as a whole. However, this represents some 32.4 percent of their estimated overall tax liability, which is very large relative to the average underpayment rate of 14.1 percent. Similarly, helpers and handlers are estimated to understate taxes by the relatively low amount of $401 on average, but this represents 23.5 percent of their estimated overall tax liability.

A number of occupational groups rank consistently high or low regardless of which way compliance is measured. For instance, the vehicle sales group ranks highest both in terms of estimated average level of noncompliance ($6,278) and estimated share of overall taxes not paid (50.5 percent). Other occupational groups that rank consistently high in terms of being noncompliant are: informal suppliers; farm and agriculture-related workers; tip earners; real estate, financial, and insurance; construction and extraction; and forestry, logging, fishing, hunting, and trapping. Occupational groups that rank consistently low include: military; administrative support; retired or disabled; production and manufacturing; protective services; accountants, auditors, and tax preparers; postsecondary teachers; and other teachers, counselors, and librarians.

Table 4 provides separate tabulations of the estimated average level of noncompliance within each occupation group for filers and nonfilers. It appears from this table that nonfiling is rather heavily concentrated within certain occupational groups. In particular, individuals employed in the helpers and handlers, other service, and informal suppliers categories account for over 60 percent of the overall nonfiler population. In contrast, they account only for 11 percent of the filer population. Across all occupations, the average level of noncompliance is over twice as large among nonfilers as it is among filers ($1,200 compared to $564). This is consistent with Erard and Ho (2001), who found that the aggregate share of noncompliance attributable to nonfilers was large in relation to their representation in the population.

5. Some possible explanations for the findings

In this section we briefly investigate a few of the many potential explanations for the pattern of results presented in Section 4. Assuming decreasing relative risk aversion, the standard expected utility theory of noncompliance suggests that noncompliance will tend to increase with income, all other factors held equal. So, one possible story why some occupational groups are less compliant is that their members enjoy higher earnings. Table 5 presents our estimates of noncompliance and average true adjusted gross income (AGI) by occupational category. The top 5 occupations in terms of the average estimated level of noncompliance all have high average levels of AGI relative to the overall population, and the bottom 5 occupations have relatively low average levels of AGI.

This suggests that income may play some role in explaining variations in compliance by occupation. However, it clearly cannot explain all of the variation. In particular, there are some occupations with relatively low levels of AGI, such as tip earners and farm and agriculture-related workers, who nonetheless rank high in terms of the average level of noncompliance. Further, there are some occupations with relatively high levels of income, such as technologists and technicians (other than health) and accountants, auditors, and tax preparers, who rank low in terms of average noncompliance.

The classical expected utility theory of noncompliance predicts that noncompliance will tend to be decreasing in marginal tax rates; however, alternative models can be developed which imply the opposite relationship. In any case, as is clear from the results in Table 5, the marginal tax rate is fairly highly correlated with AGI, making it difficult to distinguish whether occupational groups are responsive to the marginal tax rate independently of the level of income.

Most compliance theories predict that noncompliance will be more prevalent among households with better opportunities for noncompliance, as reflected in a lower relatively likelihood of audit, detection, or penalty. We are working on developing an index reflecting the differential opportunities across occupation groups, but do not yet have any results to report. However, a casual review of the occupations ranked high in terms of noncompliance suggests that many of these occupations are likely to be associated with relatively high opportunities for underreporting taxes.

The results in Table 5 also indicate that nonfiling is relatively more common among certain occupations. One possible explanation for this finding is that nonfiling is relatively more common when the burden of filing a return is high. Although the results indicate that burden is high for some occupations with relatively high nonfiling rates (informal suppliers, vehicle sales, and real estate, financial, and insurance), there are some occupations with relatively high nonfiling rates and low estimated filing burdens (helpers and handlers and other services). The last column of Table 5 indicates the percentage of households in each occupational category that filed a federal income tax return in tax year 1987. Consistent with the findings of Erard and Ho (2001), nonfiling appears to be a persistent phenomenon in that those occupations with a low filing rate in 1987 tend to have a relatively high incidence of nonfiling in 1988.

6. Conclusion

In this paper, we have described a micro-simulation data base for understanding compliance that we have been developing. We believe it has the potential to broaden our understanding of who is hard to tax and why. As an illustration of the data base's potential, we have used it to derive a preliminary map where different occupational groups in the U.S. fall along the compliance continuum, and we explored a few of the many possible reasons for the wide divergence in reporting and filing compliance that we have identified. Clearly this analysis is very preliminary and more work needs to be done.

In terms of understanding the variation in compliance by occupation, we plan in future research to develop an index of opportunity to determine the extent to which differences in the likelihood of detection and penalty are responsible for the variation in compliance behavior across occupations. We are also exploring possible econometric specifications for more rigorously measuring the contributions of different factors to observed compliance outcomes.

In terms of our data base, we are considering ways to improve the imputation of undetected noncompliance to individuals. One possibility is to employ a detection controlled econometric specification to estimate the level of undetected noncompliance on various line items of each return. We are also considering ways to implement a hot decking procedure to allocate external aggregate estimates for items such as tips and informal supplier income among the returns in our sample. Given that the data are some 15 years old, we are also exploring the feasibility of aging our database to make it more representative of current conditions.

Finally, the model is not limited to analyzing noncompliance by occupational category. In future work, we hope to apply it to examine other interesting questions as well.

References

Alm, James. (1991) "On the experimental analysis of taxpayer reporting," *The Accounting Review,* 66:3, 577–593.

Alm, James, Roy Bahl, and Matthew N. Murray. (1991) "Tax Base Erosion in Developing Countries," *Economic Development and Cultural Change*, 39:4, 849–872.

Erard, Brian and Chih-Chin Ho. (2001) "Searching for ghosts: who are the nonfilers and how much tax do they owe?," *Journal of Public Economics,* 81:1, 25–50.

Erard, Brian. (1997) "Self-selection with measurement errors: a microeconometric analysis of the decision to seek tax assistance and its implication for tax compliance," *Journal of Econometrics,* 81: 2 , 319–356.

Feinstein, Jonathan S. (1991) "An econometric analysis of income tax evasion and its detection," *RAND Journal of Economics*, 22:1, 14–35.

Internal Revenue Service. (1996) "Federal tax compliance research: individual income tax gap estimates for 1985, 1988, and 1992," Publication 1415 (Revised 4-96), Washington, DC.

Schneider, Friedrich and Dominik H. Enste. (2000) "Shadow economies: size, causes, and consequences," *Journal of Economic Literature,* 38:1, 77-114.

Smith, J and T. Adams. (1987) "The measurement of selected income flows in informal markets 1981 and 1985-1986," Survey Research Center, University of Michigan, April.

Table 1: Broad categories of informal economy goods and services used in the University of Michigan Studies

1. Food
2. Home Repairs
3. Vehicle Repairs
4. Appliance Repairs
5. Personal Care
6. Housekeeping
7. Lawn & Garden
8. Clothing Repairs
9. Flea Market Goods
10. Fuel
11. Lessons
12. Cosmetic Services
13. Catering
14. Sidewalk Vendor

Table 2: Distribution of noncompliance by occupation, sorted by estimated average level of noncompliance

Occupation	Avg. level of noncompliance	% of total taxes not paid	Group's share of population	Group's share of total tax gap
Vehicle sales	$6,278	50.5%	0.1%	0.51%
Investors	$4,178	14.3%	0.2%	1.39%
Informal suppliers	$3,928	43.6%	3.0%	19.46%
Lawyers and judges	$2,079	8.2%	0.5%	1.69%
Real estate, financial, insurance	$1,968	19.4%	1.4%	4.48%
Doctors and dentists	$1,689	5.5%	0.5%	1.47%
Farm and agriculture related	$1,321	30.8%	2.0%	4.31%
Tip earners	$1,000	49.5%	4.0%	6.48%
Construction & extraction	$992	21.5%	4.5%	7.23%
Non-govt. officials & administrators	$983	5.3%	3.4%	5.49%
Forestry, logging, fishing, hunting, trapping	$919	22.5%	0.3%	0.49%
Other sales occupations	$908	18.0%	6.7%	9.89%
Writers, performing artists, editors, announcers	$768	12.9%	1.0%	1.26%
Social and religious workers	$701	20.8%	0.7%	0.76%
Athletes and related workers	$678	9.3%	0.1%	0.13%
Managers, consultants, public relations	$642	9.3%	2.2%	2.28%
Social scientists	$616	5.9%	0.1%	0.07%
Mechanics & repairers	$575	15.4%	3.5%	3.32%
Transportation & material Moving	$544	14.1%	2.8%	2.51%
Mathematicians, engineers, computer & natural scientists, architects	$505	5.9%	2.6%	2.13%
Govt. officials & administrators	$424	7.4%	0.7%	0.51%

Occupation	Avg. level of noncompliance	% of total taxes not paid	Group's share of population	Group's share of total tax gap
Post-secondary teachers	$406	5.9%	0.3%	0.18%
Helpers and handlers	$401	23.5%	7.1%	4.62%
Other teachers, counselors, librarians	$397	9.7%	2.1%	1.34%
Other services	$359	32.4%	4.8%	2.80%
Accountants, auditors, tax preparers	$355	5.0%	1.1%	0.64%
Other health workers	$343	9.5%	3.1%	1.71%
Technologists & technicians (other than health)	$324	6.3%	2.1%	1.10%
Protective services	$265	6.9%	1.6%	0.68%
Production/ manufacturing	$265	8.4%	7.0%	3.02%
Retired or disabled	$261	8.8%	11.8%	5.03%
Administrative support	$171	7.7%	7.8%	2.18%
Military	$125	7.2%	1.4%	0.28%
Other	$37	6.6%	9.4%	0.56%
All occupations combined	$615	14.1%	100.0%	100.0%

Table 3: Distribution of noncompliance by occupation, sorted by estimated % of taxes underpaid

Occupation	Avg. level of noncompliance	% of total taxes not paid	Group's share of population	Group's share of total tax gap
Vehicle sales	$6,278	50.5%	0.1%	0.51%
Tip earners	$1,000	49.5%	4.0%	6.48%
Informal suppliers	$3,928	43.6%	3.0%	19.46%
Other services	$359	32.4%	4.8%	2.80%
Farm and agriculture related	$1,321	30.8%	2.0%	4.31%
Helpers and handlers	$401	23.5%	7.1%	4.62%
Forestry, logging, fishing, hunting, trapping	$919	22.5%	0.3%	0.49%
Construction & extraction	$992	21.5%	4.5%	7.23%
Social and religious workers	$701	20.8%	0.7%	0.76%
Real estate, financial, insurance	$1,968	19.4%	1.4%	4.48%
Other sales occupations	$908	18.0%	6.7%	9.89%
Mechanics & repairers	$575	15.4%	3.5%	3.32%
Investors	$4,178	14.3%	0.2%	1.39%
Transportation & material Moving	$544	14.1%	2.8%	2.51%
Writers, performing artists, editors, announcers	$768	12.9%	1.0%	1.26%
Other teachers, counselors, librarians	$397	9.7%	2.1%	1.34%
Other health workers	$343	9.5%	3.1%	1.71%
Managers, consultants, public relations	$642	9.3%	2.2%	2.28%
Athletes and related workers	$678	9.3%	0.1%	0.13%
Retired or disabled	$261	8.8%	11.8%	5.03%
Production/ manufacturing	$265	8.4%	7.0%	3.02%
Lawyers and judges	$2,079	8.2%	0.5%	1.69%

Occupation	Avg. level of noncompliance	% of total taxes not paid	Group's share of population	Group's share of total tax gap
Administrative support	$171	7.7%	7.8%	2.18%
Govt. officials & administrators	$424	7.4%	0.7%	0.51%
Military	$125	7.2%	1.4%	0.28%
Protective services	$265	6.9%	1.6%	0.68%
Other	$37	6.6%	9.4%	0.56%
Technologists & technicians (other than health)	$324	6.3%	2.1%	1.10%
Mathematicians, engineers, computer & natural scientists, architects	$505	5.9%	2.6%	2.13%
Post-secondary teachers	$406	5.9%	0.3%	0.18%
Social scientists	$616	5.9%	0.1%	0.07%
Doctors and dentists	$1,689	5.5%	0.5%	1.47%
Non-govt. officials & administrators	$983	5.3%	3.4%	5.49%
Accountants, auditors, tax preparers	$355	5.0%	1.1%	0.64%
All occupations combined	$615	14.1%	100.0%	100.0%

Table 4: Distribution of noncompliance by occupation and whether a return was filed

Occupation	Filers		Nonfilers		Filers & nonfilers combined	
	Avg. level of non-compliance	% of filer popn.	Avg. level of non-compliance	% of nonfiler popn.	Avg. level of non-compliance	% of overall popn.
Vehicle sales	$6,507	0.1%	$2,647	0.04%	$6,278	0.1%
Investors	$3,998	0.2%	$8,701	0.10%	$4,178	0.2%
Informal suppliers	$3,437	2.6%	$5,824	7.88%	$3,928	3.0%
Lawyers and judges	$1,784	0.5%	$5,434	0.51%	$2,079	0.5%
Doctors and dentists	$1,661	0.6%	$10,923	0.02%	$1,689	0.5%
Real estate, financial, insurance	$1,643	1.4%	$5,044	1.68%	$1,968	1.4%
Farm and agriculture related	$1,266	2.1%	$3,375	0.65%	$1,321	2.0%
Tip earners	$971	4.2%	$1,728	1.89%	$1,000	4.0%
Non-govt. officials & administrators	$951	3.6%	$1,970	1.33%	$983	3.4%
Construction & extraction	$948	4.7%	$1,967	2.43%	$992	4.5%
Forestry, logging, fishing, hunting, trapping	$863	0.3%	$2,423	0.15%	$919	0.3%
Other sales occupations	$819	6.8%	$2,204	5.40%	$908	6.7%
Social and religious workers	$723	0.7%	$146	0.33%	$701	0.7%
Writers, performing artists, editors, announcers	$683	1.0%	$2,177	0.72%	$768	1.0%
Athletes and related workers	$670	0.1%	$1,130	0.03%	$678	0.1%
Managers, consultants, public relations	$620	2.3%	$1,050	1.40%	$642	2.2%
Social scientists	$616	0.1%	$0	0.00%	$616	0.1%

Occupation	Filers		Nonfilers		Filers & nonfilers combined	
	Avg. level of non-compliance	% of filer popn.	Avg. level of non-compliance	% of nonfiler popn.	Avg. level of non-compliance	% of overall popn.
Other services	$550	2.8%	$142	28.23%	$359	4.8%
Transportation & material moving	$504	3.0%	$2,752	0.63%	$544	2.8%
Mathematicians, engineers, computer & natural scientists, architects	$487	2.8%	$1,546	0.55%	$505	2.6%
Mechanics & repairers	$460	3.8%	$5,373	1.04%	$575	3.5%
Post-secondary teachers	$421	0.3%	$17	0.12%	$406	0.3%
Govt. officials & administrators	$412	0.8%	$718	0.36%	$424	0.7%
Other teachers, counselors, librarians	$401	2.2%	$99	0.35%	$397	2.1%
Other health workers	$350	3.3%	$0	0.76%	$343	3.1%
Accountants, auditors, tax preparers	$336	1.2%	$1,029	0.37%	$355	1.1%
Technologists & technicians (other than health)	$303	2.2%	$828	1.04%	$324	2.1%
Helpers and handlers	$295	5.6%	$678	24.57%	$401	7.1%
`Production/ manufacturing	$266	7.6%	$67	0.55%	$265	7.0%
Protective services	$258	1.7%	$723	0.29%	$265	1.6%
Retired or disabled	$257	12.4%	$391	4.72%	$261	11.8%
Administrative support	$171	8.3%	$172	2.68%	$171	7.8%
Military	$137	1.4%	-$307	0.46%	$125	1.4%
Other	$33	9.5%	$87	8.73%	$37	9.4%
All occupations combined	$564	100.0%	$1,200	100.00%	$615	100.0%

Table 5: Distribution of noncompliance and selected potential determinants by occupation

Occupation	Avg. level of non-compliance	% of total taxes not paid	% of nonfilers in group	Avg. true AGI	Avg. marg. tax rate	Avg. filing burden (hours)	% of prior year filers in group
Vehicle sales	$6,278	50.5%	5.9%	$55,551	20.5%	29.5	89.6%
Investors	$4,178	14.3%	3.8%	$77,775	23.0%	25.9	91.8%
Informal suppliers	$3,928	43.6%	20.6%	$45,194	21.2%	24.2	77.0%
Lawyers and judges	$2,079	8.2%	8.1%	$114,747	27.2%	25.4	92.6%
Real estate, financial, insurance	$1,968	19.4%	9.6%	$54,122	22.0%	22.5	92.2%
Doctors and dentists	$1,689	5.5%	0.3%	$138,079	26.8%	27.8	98.6%
Farm and agriculture related	$1,321	30.8%	2.6%	$22,912	17.4%	22.8	86.4%
Tip earners	$1,000	49.5%	3.8%	$16,109	16.1%	8.2	79.6%
Construction & extraction	$992	21.5%	4.3%	$31,198	18.8%	14.9	92.5%
Non-govt. officials & administrators	$983	5.3%	3.1%	$95,959	23.9%	18.6	95.0%
Forestry, logging, fishing, hunting, trapping	$919	22.5%	3.6%	$28,209	17.7%	15.3	98.1%
Other sales occupations	$908	18.0%	6.4%	$32,378	19.2%	14.2	89.3%
Writers, performing artists, editors, announcers	$768	12.9%	5.7%	$35,872	20.2%	17.4	91.2%
Social and religious workers	$701	20.8%	3.9%	$23,953	17.1%	16.8	90.9%
Athletes and related workers	$678	9.3%	1.7%	$39,550	18.1%	15.7	76.9%
Managers, consultants, public relations	$642	9.3%	5.1%	$44,173	21.8%	15.3	92.9%
Social scientists	$616	5.9%	0.0%	$61,288	24.0%	19.8	100.0%
Mechanics & repairers	$575	15.4%	2.3%	$30,195	18.9%	13.3	93.1%
Transportation & material Moving	$544	14.1%	1.8%	$31,011	18.6%	13.6	95.3%
Mathematicians, engineers, computer & natural scientists, architects	$505	5.9%	1.7%	$54,835	24.4%	17.5	97.7%
Govt. officials & administrators	$424	7.4%	3.9%	$42,850	21.8%	15.7	96.3%
Post-secondary teachers	$406	5.9%	3.5%	$47,139	23.7%	17.8	96.1%
Helpers and handlers	$401	23.5%	27.6%	$13,082	16.4%	10.8	66.6%
Other teachers, counselors, librarians	$397	9.7%	1.3%	$31,778	20.6%	14.0	95.2%
Other services	$359	32.4%	46.8%	$12,188	15.7%	10.0	45.4%

Occupation	Avg. level of non-compliance	% of total taxes not paid	% of nonfilers in group	Avg. true AGI	Avg. marg. tax rate	Avg. filing burden (hours)	% of prior year filers in group
Accountants, auditors, tax preparers	$355	5.0%	2.7%	$45,240	22.3%	16.4	94.8%
Other health workers	$343	9.5%	2.0%	$26,328	19.4%	11.3	91.6%
Technologists & technicians (other than health)	$324	6.3%	4.0%	$37,923	21.8%	13.4	96.4%
Protective services	$265	6.9%	1.5%	$32,479	19.6%	13.3	95.0%
Production/ manufacturing	$265	8.4%	0.6%	$27,167	18.5%	10.7	93.5%
Retired or disabled	$261	8.8%	3.2%	$22,845	17.3%	14.5	94.4%
Administrative support	$171	7.7%	2.7%	$20,082	17.4%	9.1	90.2%
Military	$125	7.2%	2.7%	$20,745	16.1%	9.2	95.6%
Other	$37	6.6%	7.4%	$5,211	15.2%	6.6	66.0%
All occupations combined	$615	14.1%	8.0%	$28,836	18.5%	13.1	85.3%

Endnotes

[1] Refer to Andreoni, Erard, and Feinstein (1998) and Schneider and Enste (2000), respectively, for surveys of these literatures.

[2] Numerous experimental studies of tax compliance have also been undertaken. See Alm (1991) for a survey. Although such studies have provided valuable evidence on how individuals respond to incentives to cheat or comply, the results provide only a very rough guide to the extent to which different individuals or groups in society might likely to participate in tax noncompliance.

[3] Examples include Feinstein (1991) and Erard (1997).

[4] The TCMP population excludes returns that were filed late as well as returns filed by non-resident taxpayers.

[5] Non-residents and individuals without valid social security numbers were excluded from the analysis.

[6] In the U.S., households with income below a specified filing threshold that varies according to age. marital, and dependency status are not required to file a federal income tax return.

[7] Unlocated individuals in the sample tended to have much larger sample weights as a consequence of the way the sample was stratified. The sample weights for the 4,594 individuals in the sample aggregate to approximately 43 percent of the potential nonfiler population.

[8] Our approach includes an enhancement to the original IRS approach in that we adjust the weights separately by sampling stratum to make the 2,195 returns broadly representative of all nonfilers who were located during the search process. For the 1996 tax gap report, the IRS adjusted the sample weights for all 2,195 returns by the same factor.

[9] The intuition behind this approach is as follows. Suppose a true nonfiler with given characteristics has a probability of, say, one half of being located. This suggests that for every one true nonfiler with these characteristics who has been located, there is another true nonfiler with the same characteristics who has not. This is analogous to drawing a 50 percent random subsample of all true nonfilers with such characteristics. To make the located true nonfilers with these characteristics representative of all true nonfilers with these characteristics, the original sample weight of the located individuals is therefore divided by the implied sampling probability -- in this case by one half. As a further adjustment, we have divided the sample weights for the secured delinquent returns of married joint nonfilers by a factor of two. All else equal a delinquent married couple's return has approximately twice the chance of being included in our sample as a delinquent single individual's return. This is because it would be sufficient for either member of the couple to be included in the sample of located nonfilers for their joint return to be secured.

[10] We employ the IRS measure of filing burden originally developed by Arthur D. Little, Inc., which is computed by aggregating the estimated average completion times associated with each form and schedule used by the taxpayer. Thus, in essence, the measure reflects a weighted number of forms and schedules, where the weights are the estimated completion times.

[11] This estimate represents “true nonfilers”; individuals with no legal filing requirement were separately estimated to have received $93 million in tips.

[12] In the case of households that reported self-employment income, we also relied on the per exam codes for their principal industrial activity.

[13] To avoid assigning too large a share of tips to individuals reporting unusually large amounts of income, we set a ceiling for reported tips. This ceiling represented the amount that would be allocated to a household reporting $20,000 in wages, salaries, and tips.

[14] See Smith and Adams (1987).

[15] Our calculator ignores issues such as the Alternative Minimum Tax, but does take into account the phase-out of personal exemptions that applies to taxpayers with high levels of income.

[16] The principle difficulty was computing the additional self-employment tax for married joint filers. For such households, it was not possible using our data to determine what shares of additional self-employment and wage and salary income were attributable to each spouse. Nor was it possible to determine which households were entitled to use the optional method for computing self-employment taxes. Details on the algorithm used to compute the change in self-employment tax are available from the authors.

[17] This measure includes not only income taxes, but also the items classified as "additional taxes" (taxes on distributions from trusts) and "other taxes" (self-employment tax, alternative minimum tax, recapture tax, social security tax on tip income not reported to employer, etc.).