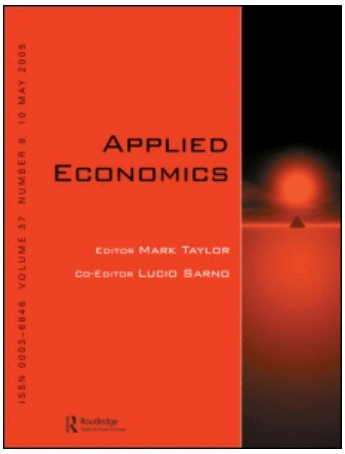
This article was downloaded by: *[Tedds, Lindsay]*
On: *15 April 2009*
Access details: *Access Details: [subscription number 910455535]*
Publisher *Routledge*
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Applied Economics
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713684000

Keeping it off the books: an empirical investigation of firms that engage in tax evasion

Lindsay M. Tedds [a]
[a] School of Public Administration, University of Victoria, Victoria, BC, V8W 2Y2, Canada

First Published on: 09 April 2009

To cite this Article Tedds, Lindsay M.(2009)'Keeping it off the books: an empirical investigation of firms that engage in tax evasion',Applied Economics,99999:1,
To link to this Article: DOI: 10.1080/00036840701858141
URL: http://dx.doi.org/10.1080/00036840701858141

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Applied Economics, 2009, 1–15, iFirst

Routledge
Taylor & Francis Group

Keeping it off the books: an empirical investigation of firms that engage in tax evasion

Lindsay M. Tedds

School of Public Administration, University of Victoria, Victoria, BC, V8W 2Y2, Canada
E-mail: ltedds@uvic.ca

Downloaded By: [Tedds, Lindsay] At: 15:49 15 April 2009

This article uses a unique dataset that contains detailed information on firms from around the world to investigate factors that affect under-reporting behaviour. The empirical strategy employed exploits the nature of the dependent variable, which is interval coded, and uses interval regression which provides an asymptotically efficient estimator provided that the classical linear model assumptions hold. These assumptions are investigated using standard diagnostic tests that have been modified for the interval regression model. Evidence is presented that shows that the firms in all regions engage in under-reporting. Regression results indicate that government corruption has the single largest causal effect on under-reporting, resulting in the percentage of sales not reported to the tax authority being 51.3% higher. Taxes have the second single largest causal effect on under-reporting, resulting in the percentage of sales not reported to the tax authority being 18.0% higher, followed by access to financing at 8.9% higher and organized crime at 7.6% higher. Inflation, political instability, exchange rates and the fairness of the legal system were found to have no effect on under-reporting. It is also found that there is a significant correlation between under-reporting and the legal organization of the business, size, age, ownership, competition and audit controls.

I. Introduction

It is generally accepted that taxes and tax evasion are intrinsically linked; one cannot exist without the other. As a result of a great deal of theoretical, experimental and empirical research conducted over the last 20 years, there exists an extensive knowledge base regarding tax evasion by individuals. However, research regarding tax evasion by businesses is, by comparison, surprisingly modest.[1] This is startling, given the importance of businesses and their decisions not only in economic models but also in tax systems and the economy as a whole.

[1] The focus of this article is on illegal tax evasion and not legal tax avoidance. Tax evasion or tax noncompliance refers to income tax that is legally owed but is not reported or paid whereas tax avoidance refers to legal actions taken to reduce tax liability. There exists a voluminous literature on corporate tax avoidance, notably work by James Hines at the University of Michigan, and interested readers are encouraged to consult this literature.

Applied Economics ISSN 0003–6846 print/ISSN 1466–4283 online © 2009 Taylor & Francis
http://www.informaworld.com
DOI: 10.1080/00036840701858141

Downloaded By: [Tedds, Lindsay] At: 15:49 15 April 2009

One of the only existing estimates on the distribution of share of legally generated income that goes unreported by firms is produced by the United States (Internal Revenue Service (IRS), 2004). These estimates, denoted in US dollars, indicate that in total, businesses evaded $174.6 billion in taxes in the 2001 tax year, which amounted to almost 10% of total taxes paid voluntarily. Given that there is some evidence which supports the notion that businesses engage in tax noncompliance, several questions arise and require investigation. First, do businesses around the world engage in tax evasion or is it confined to a few countries or regions? Second, does the legal status of the business (e.g. sole proprietorship, partnership, corporation, etc.) affect the incidence and/or intensity of tax evasion? If so, then it may be possible to effect changes in the legal system in order to increase tax compliance. Third, do businesses that engage in tax noncompliance share common and observable characteristics, or is there too wide a variety of shapes and sizes to permit a useful generalization about them? If it were possible to define a typical evader, the tax authority could target their auditing activities more accurately. Finally, while there is considerable agreement internationally about the factors that likely trigger tax noncompliance (e.g. the tax burden, the degree of regulation, the level of enforcement, confidence in government, labour force characteristics and morality), how do these features influence the intensity of noncompliance? With this information, policy makers could effect changes to increase the amount of tax revenues collected from businesses.

One of the main constraints to investigating and attempting to provide answers to these and related questions is the lack of data. To date, most of the existing empirical research on firm tax evasion uses data from tax audits. However, these sources are only available for a very small number of selected countries, the data is costly to collect, and access to the data is limited. In addition, tax audit data is subject to both selection and measurement error bias. The selection effect occurs because firms are generally not randomly audited. Rather, firms are usually selected for an audit because they have a particular characteristic (e.g. US businesses with more than $10 million in assets are audited annually) or because their tax filing documents raise a 'red-flag' with the tax authority. The measurement error bias occurs because not only are audits unlikely to completely and accurately measure tax noncompliance but also the final noncompliance report is usually based on negotiations between the firm and the tax authority as well as various legal appeals.

More recently, however, an alternative data source has become available that is conducive to investigating issues related to business tax noncompliance namely, the World Business Environment Survey (WBES) which was hosted by the World Bank's Enterprise Analysis Unit. The survey was administered to more than 10 000 firms in 80 countries in late 1999 and early 2000, and provides information regarding a firm's characteristics as well as responses to multiple questions on the investment climate and business environment. The survey also includes a question regarding the extent and intensity to which firms fail to report income to the tax authority. Rather than reporting the exact amount of income that went unreported to the tax authority, firms were provided with seven possible intervals of varying size that were grouped according to the percentage of sales that are reported for tax purposes. When a quantitative outcome is grouped into known intervals on a continuous scale the data is said to be interval coded. Asymptotically consistent estimates can be obtained using interval regression. The advantage of interval regression is that, provided certain key assumptions are met, the coefficients on the exogenous variables can be interpreted as if Ordinary Least Squares (OLS) were applied to a continuous dependent variable.

Overall, and not surprisingly, high taxes and government corruption results in lower compliance and the magnitude of these effects are quite large. It is found that government corruption has the single largest effect, resulting in the percentage of sales not reported to the tax authority being a whopping 51.3% higher, taxes have the second largest effect, causing the percentage of sales not reported to the tax authority to be 18.0% higher, access to financing has the third largest effect, causing the percentage of sales not reported to the tax authority to be 8.9% higher, and organized crime has rounds out the group, causing the percentage of sales not reported to the tax authority to be 7.6% higher. Access to capital, political instability, organized crime, inflation, exchange rates and the legal system are found to have no statistically significant effect on tax noncompliance. It is also found that there is a significant correlation between under-reporting and the legal organization of the business, size, age, ownership, competition and audit controls.

This article begins with a review of the relevant tax evasion literature. Section III describes the WBES data. Section IV outlines the empirical technique used in this article as well as a description of the key statistical assumption and associated tests. The results are then summarized and discussed in Section V. The article ends with some concluding comments.

Downloaded By: [Tedds, Lindsay] At: 15:49 15 April 2009

II. Literature Review

To date, research that seeks to measure and explain unreported activity has predominately employed macro methods which produce estimates of the total size of unreported activity.[2] For example, Hill and Kabir (2000), Giles *et al*. (2002), Giles and Tedds (2002) and Tedds (2005) reported estimates for Canada; Giles (1997a, 1999), Giles and Caragata (2001), Draeseke and Giles (2002) and Giles and Johnson (2002) reported estimates for New Zealand; Bajada (1999, 2005) reported estimates for Australia; Mathews (1984) reported estimates for the US; and Dell'Anno *et al*. (2007) reported estimates for France, Greece and Spain. While estimates such as these are valuable to understanding under-reporting behaviour, there is a great deal that these estimates do not tell us. In particular, they do not distinguish between hidden legal activity (e.g. under-reporting of profits or sales) and illegal or criminal activity (e.g. production and sale of illicit drugs, prostitution) and they are unable to distinguish between under-reporting by individuals and firms. This detailed information is important for policy makers since different policy tools are required to curb criminal behaviour than those required to curb hidden legal activity, the causal factors that drive under-reporting differ between individuals and firms, and firms may respond differently to policy changes aimed at curbing hidden activity than individuals.

In order to obtain this type of information, a method that employs microeconomic data is required. While microeconometric analysis of tax evasion is not extensive, there have been a small number of empirical studies that have explored tax evasion by firms, most of which focus on tax noncompliance by the self employed.[3] The self employed are commonly believed to have lower compliance rates than wage and salary earners, primarily due to a lack of third-party reporting and withholding. One approach, popularized by Pissarides and Weber (1989), used household income and expenditure data to estimate the degree of income under-reporting. This approach, along with various modifications, has been applied by Apel (1994) and Engström and Holmlund (2006) for Sweden; Besim and Jenkins (2005) for North Cyprus; Schuetze (2002) and Tedds (2007) for Canada; Lyssiotou *et al*. (2004) for Great Britain; and Feldman and Slemrod (2007) for the United States.

Several studies have investigated business tax noncompliance using data from tax audits. Probably the most comprehensive tax audit dataset in the world is the US Internal Revenue Service (IRS) Tax Compliance Measurement Program (TCMP).[4] US tax audit data has also been used to explore tax evasion by the self employed and these studies include Erard (1992), Christian (1994) and Joulfaian and Rider (1998). Joulfaian (2000), using TCMP data from 1987, found that noncompliant corporations are three time more likely to be managed by executives who have evaded personal taxes. Rice (1992) used TCMP data from 1980 to investigate tax compliance by small corporations (defined as corporations with assets between $1 and $10 million). He found that compliance is positively associated with publicly traded corporations and negatively associated with both highly profitable and under-performing corporations, marginal tax rates and firm size.

Hanlon *et al*. (2007) were the first to the use operational data from the Voluntary Compliance Baseline Measurement (VCBLM) programme compiled by the Large and Mid-Sized Business (LMSB) Research Division of the IRS to examine corporate tax noncompliance, as measured by deficiencies identified during audit. This data is more up to date, containing audit files up to and including 2002, than the TCMP, which ended in 1988. They found that business tax noncompliance relative to scale is a convex function, with medium-sized businesses (within the population of firms with assets exceeding $10 million) having the lowest rate of noncompliance. Private firms, multinationals and firms with incentivized executive compensation schemes are associated with higher noncompliance while foreign controlled

[2] Such methods include: the Currency-Ratio Approach (Gutmann, 1977); the Monetary-Transactions Method (Feige, 1979); Tanzi's Approach (Tanzi, 1980); National Accounts/Judgmental Methods and the Latent Variable/Multiple Indicator Multiple Indicator Cause (MIMIC) model (Frey and Weck-Hanneman, 1984).

[3] As previously noted, an extensive literature exists on tax evasion by individuals and the seminal contribution was provided by Allingham and Sandmo (1972). This model has been extended in a number of dimensions over the last 30 years and this literature is nicely surveyed by Andreoni *et al*. (1998). There is a much smaller pool of literature that addresses tax noncompliance by businesses and Cowell (2004) provided an excellent review of this literature. The theoretical findings clearly suggest that the type of business, whether it be owner managed or otherwise, along with other factors such as tax rates, audit rates and penalties greatly affects the tax evasion decision of firms.

[4] The TCMP features data from a random sample of individual and small corporate income tax returns filed in a given year that were subject to intensive audits by experienced examiners. For certain groups, such as nonfilers and proprietors who tend not to report a significant amount of their income, the results from special research studies are used to supplement TCMP data. Finally, data for large corporations are obtained from routine operational audits.

Downloaded By: [Tedds, Lindsay] At: 15:49 15 April 2009

firms are more compliant. In general, firms in the manufacturing industry, trade, transportation and warehousing industry, and education, healthcare and warehousing industry are less compliant. They also found no association between compliance and effective tax rates.

Only two studies exist that use tax audit data from countries other than the US. First, Giles (1997b) investigated factors that determine the probability of noncompliance amongst New Zealand businesses that were audited by the Inland Revenue Department in that country between 1993 and 1995. He found that in general, an aggressive use of legitimate tax minimization instruments (i.e. tax avoidance behaviour such as the deduction of interest and depreciation costs, and the writing-off of bad debts) tended to be associated with increased compliance. He also found that businesses in the construction, wholesale trade, retail trade, accommodation and cafes and restaurants sectors exhibited below average compliance rates as do relatively inefficient businesses. On the other hand, businesses that were registered off-shore were generally more compliant than their on-shore counterparts (a finding that is shared by Hanlon *et al*. (2007)) and an increase in the scale of the business, regardless of how this is measured, unambiguously raises the probability of compliance, once other characteristics are controlled for. Second, Chan and Mo (2000) analysed 583 tax audit cases, made available by the Chinese tax authorities, on corporate tax noncompliance by foreign investors. Their results indicate that the corporate taxpayers tax holiday position significantly affects noncompliance, notably: (i) companies in the pre-holiday position are least compliant; (ii) companies are most compliant in the tax exemption period that has a zero tax rate and a heavy penalty for evasion; (iii) domestic market-oriented companies have a higher rate of noncompliance than their export-oriented counterparts; and (iv) wholly foreign-owned and manufacturing-oriented companies have higher compliance than joint ventures and service-oriented companies.

Finally, a small number of studies use data on firm reported tax noncompliance rather than audit data. Using firm level data from a 1997 survey of private manufacturing firms in Poland, Romania, Russia, Slovakia and the Ukraine, Johnson *et al*. (2000) investigated the relationship between government corruption, criminal activities and firm tax compliance. The authors found a positive and significant relationship between under-reporting of sales and bribing of corrupt officials, but no relationship between under-reporting of sales and protection payments to the mafia, tax payments, or efficiency of the legal system. Finally, firm tax noncompliance was greater in Russia and the Ukraine than in the other countries included in the study.

Batra *et al*. (2003) were the first to use the WBES to investigate the determinants of under-reporting by the firms. The WBES asks each firm to provide an estimate of the percentage of sales revenues that firms like their own report to the tax authority. Rather than the answer being recorded on a continuous scale, seven answers were possible: 100, 90–99, 80–89, 70–79, 60–69, 50–59 and <50%. The authors defined seven binary variables based on these seven intervals and each binary variable takes a value of 1 if the firm selected the particular interval of interest and 0 otherwise. They then estimate seven different OLS regressions which include variables for firm characteristics, rule of law, business constraints and country fixed effects. Overall they found that: (i) '... small or medium-size firms that produce for the domestic market (nonexporters), lack foreign investment and are located in large cities (but not necessarily in the capital) tend to engage more in unofficial activity' (p. 76); (ii) the prevalence, though not the unpredictability, of corruption also significantly affects noncompliance; and (iii) '... a firm's age, sector or mode of ownership do not influence (a firm's) under-reporting of revenue.' (p. 78).

There are two key reasons to be concerned about these reported results. First, the number of observations in each of the seven regressions ranged from a low of 3802 to a high of 4781 from a total number of observations of over 10 000 available in the WBES. The authors do not discuss the reasons for the attrition but it is likely the result of missing observations and nonresponse for both the dependent and independent variable. As attrition of this magnitude can lead to biased results, it is important that it be investigated further. Second, Batra *et al*. (2003) did not exploit the nature of the dependent variable. The firm's response regarding under-reporting behaviour is grouped into categories. When a quantitative outcome is grouped into known intervals on a continuous scale, the data are said to be 'interval-coded'. However, Batra *et al*. (2003) defined the dependent variable as a binary outcome for each category and estimated the resulting equations by OLS. The considerable statistical limitations of such a linear probability model are well known.[5] There is an estimation technique that has

[5] In the WBES, over 40% of the firms surveyed indicate that 100% of their sales are reported to the tax authority and between 6 and 14% of firms appear in one of the other categories. This means that the linear probability models estimated by Batra *et al*. will be either 0 or 1 inflated. This results in a large number of predicted probabilities falling outside the unit interval.

been developed specifically for interval-coded data. This estimation procedure is known as 'interval regression' and is undertaken using maximum likelihood techniques.

This article addresses these shortcomings. First, it uses the interval regression technique and the extent to which various covariates affect the estimation results is also explored. Second, information from existing theoretical and empirical literature of firm tax noncompliance is used to build the empirical model explored in this article. Third, the issue of missing observations or nonresponse is explored. This article also builds on the findings of Johnson *et al*. by extending the number of countries included in the sample and examines firms in a variety of different industries.

III. Data

In 1998, the World Bank Group launched its WBES. The WBES used many of the same questions from the enterprise survey conducted for the *1997 World Development Report* (World Bank, 1997) but expanded the number of businesses and countries surveyed and the questions/issues covered.[6] The survey is a sample from the universe of registered businesses. The WBES uses a form of stratified random sampling along with over-sampling and nonrandom methods to correct for biases in the representation across firms with respect to industry characteristics that were not common.[7] To ensure adequate representation of firms by industry, size, ownership, export orientation and location, sampling targets were agreed on across all regions.

The survey was completed by over 10 000 firms across 80 countries, including the West Bank and Gaza over a 20-month period between the latter half of 1998 and mid-2000. The purpose of the survey was to assess and compare the business environment in a large number of countries. To achieve this goal, the survey gathered information regarding the firm's characteristics, such as size and ownership structure, as well as responses to multiple questions on the investment climate and the local business environment as shaped by domestic economic policy, governance, regulatory, infrastructural and financial impediments, as well as assessments of public service quality. A more detailed description of the survey can be found in Batra *et al*. (2003).

Of interest to this study, the survey included a question regarding the extent and intensity to which firms fail to report income to the tax authority which permits exploring the links between tax noncompliance and various firm, business, governance and political characteristics.[8] The main advantages of using the WBES to explore tax evasion over audit data include the fact that: (i) audit data are not widely available, unlike the WBES, which covers 80 countries; (ii) audit data only include firms that are selected (for diverse reasons) or caught by the tax authorities, while the WBES is a broad sample of firms; and (iii) the WBES includes additional information that is not included in tax audits. In particular, factors that firms perceive as business obstacles, such as taxes and regulations, which may affect a firm's decision to under-report, are included.

The WBES does not come without disadvantages. One disadvantage shared by both the WBES and audit data is that neither data source includes 'ghosts'; firms that operate solely in cash and avoid normal business obstacles and regulations. The effect of omitting these firms from the estimates will bias estimates of tax noncompliance downwards and this bias will be larger for those countries that have a larger percentage of unregistered firms. Further, the percentage of sales that the firm fails to report to the tax authority is not expressed in actual quantities, as with Johnson *et al*. (2000), but rather it is expressed in intervals. There were seven possible answers: 100, 90–99, 80–89, 70–79, 60–69, 50–59 and <50%. This represents something of a constraint but its key advantage is that the interval-coded nature of the responses is likely to reduce a respondent's reporting bias.

[6] Unfortunately, the *1997 World Development Report* dataset contains neither detailed information on characteristics of the firms nor the key variable of interest contained in the WBES dataset. As a result, it cannot be appended to the WBES dataset.
[7] Unfortunately, the survey does not contain survey weights. The lack of survey weights means that individual indicators should not be used for precise country rankings in any particular dimension measured by the survey.
[8] Similar to the data used by Johnson *et al*., the exact phrasing of the question was 'What percentage of total sales would you estimate the typical firm in your area of activity reports for tax purposes?'.

Downloaded By: [Tedds, Lindsay] At: 15:49 15 April 2009

Table 1. Univariate frequencies of percentage of sales reported to tax authorities, excluding missing observations

	<50%	50–59%	60–69%	70–79%	80–89%	90–99%	100%
Frequency	936	694	501	703	916	1096	3307
Percentage	11.48	8.51	6.14	8.62	11.24	13.44	40.56
Observations	8153						

Note: 1879 missing observations.

The distribution of answers to the sales reporting question is given in Table 1.[9] It shows that 60% of firms worldwide indicate that the typical firm fails to report their sales in full to the tax authority and, of those firms, over 19% of them fail to report more than half their sales. This shows that business tax compliance is a significant issue. Not surprisingly, there appears to be some differences in perceived tax noncompliance across regions, as is shown in Fig. 1. In particular, in Organization for Economic Cooperation and Development (OECD) countries only ~40% of firms' under-report their sales to the tax authority and of those, ~50% fail to report only up to 10% of their sales. Further, compared to other regions, significantly more firms in Latin America and Asia fail to report more than 50% of their sales. In addition, there are significant differences in perceived firm tax compliance across countries.

Downloaded By: [Tedds, Lindsay] At: 15:49 15 April 2009

IV. Empirical Framework

Interval regression model

The survey question that forms the basis for the dependent variable and is described above refers to categories. As a result, a firm's reporting behaviour is not directly observed. Rather, firms are categorized on the basis of the percentage of sales that go unreported. When a quantitative outcome is grouped into known intervals on a continuous scale, the data is said to be 'interval-coded'. The appropriate estimation procedure for this type of data is known as 'interval regression' and is undertaken using maximum likelihood techniques. Interval regression is preferred to OLS, as OLS on the grouped dependent variable model is inconsistent.

The following is the general setup of the model and is based on the discussion contained in Stewart (1983). The responses for the dependent variable are coded 1 through 7 to capture seven distinct sales under-reporting categories. Let y_i denote the observable ordinal variable coded in this way for the i-th firm and let y_i^* denote the underlying variable that captures the sales under-reporting of the i-th firm. This can be expressed as a linear function of a vector of explanatory variables x_i using the following relationship[10]:

$$y_i^* = x_i'\beta + u_i, \quad u_i \sim N(0, \sigma^2) \tag{1}$$

It is assumed that y_i^* is related to the observable ordinal variable y_i as follows:

$$\begin{aligned} y_i &= 1 \quad \text{if } y_i^* \leq 49\% \\ y_i &= 2 \quad \text{if } 49\% < y_i^* \leq 59\% \\ y_i &= 3 \quad \text{if } 59\% < y_i^* \leq 69\% \\ y_i &= 4 \quad \text{if } 69\% < y_i^* \leq 79\% \\ y_i &= 5 \quad \text{if } 79\% < y_i^* \leq 89\% \\ y_i &= 6 \quad \text{if } 89\% < y_i^* \leq 99\% \\ y_i &= 7 \quad \text{if } y_i^* > 99\% \end{aligned} \tag{2}$$

[9] Just over 18% of the firms in the sample did not respond to the question of interest. It is quite common in survey data to have missing data of this nature. If the data is missing for unknown reasons and the missing data is unrelated to the completeness of the other observations then it is reasonable to exclude the missing observations from the analysis, though efficiency is sacrificed. If, however, the missing data are systematically related to the phenomenon being modelled then there will be a sample selection problem. For example, if firms whose response is missing for the sales reporting question are more likely to be firms who under-report their income, then the missing observations represent more than just missing information. Unfortunately, there is insufficient evidence to definitively ascertain the reasons for the missing observations in this particular case. Rather, the relationship between these missing observations and firm characteristics that will be used as explanatory variables in our model were explored. If the missing observations are randomly distributed amongst these explanatory variables then this provides some evidence that the missing information may not cause systematic bias in the estimates. There appeared to be no significant systematic relationship between nonresponse to the sales reporting question and the explanatory variables, with one exception; no firm located in Albania responded to the question, indicating that survey was not carried out in a similar fashion to that in other countries rather than a response bias. As a result, Albania is excluded from the analysis.

[10] Wik *et al*. (2004) extended the interval regression model to include random effects.

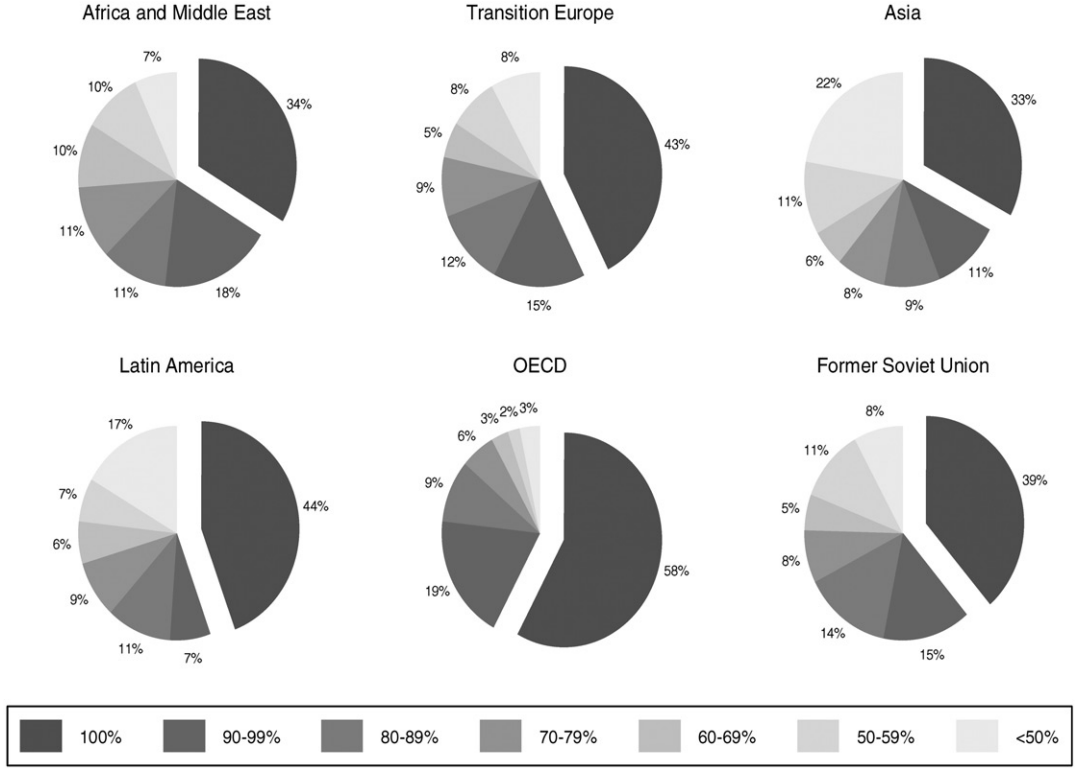


Fig. 1. Percentage of sales reported to tax authorties
Source: World Business Environment Survey.

Downloaded By: [Tedds, Lindsay] At: 15:49 15 April 2009

Based on the above, the general log likelihood function for the i-th individual is expressed as:

$$
\begin{aligned}
L_i = {} & I[y_i = 1] \times \log_e \left\{ \Phi\left(\frac{49 - x_i'\beta}{\sigma}\right) - \Phi\left(\frac{a_0 - x_i'\beta}{\sigma}\right) \right\} \\
& + I[y_i = 2] \times \log_e \left\{ \Phi\left(\frac{59 - x_i'\beta}{\sigma}\right) - \Phi\left(\frac{49 - x_i'\beta}{\sigma}\right) \right\} \\
& + I[y_i = 3] \times \log_e \left\{ \Phi\left(\frac{69 - x_i'\beta}{\sigma}\right) - \Phi\left(\frac{59 - x_i'\beta}{\sigma}\right) \right\} \\
& + I[y_i = 4] \times \log_e \left\{ \Phi\left(\frac{79 - x_i'\beta}{\sigma}\right) - \Phi\left(\frac{69 - x_i'\beta}{\sigma}\right) \right\} \\
& + I[y_i = 5] \times \log_e \left\{ \Phi\left(\frac{89 - x_i'\beta}{\sigma}\right) - \Phi\left(\frac{79 - x_i'\beta}{\sigma}\right) \right\} \\
& + I[y_i = 6] \times \log_e \left\{ \Phi\left(\frac{99 - x_i'\beta}{\sigma}\right) - \Phi\left(\frac{89 - x_i'\beta}{\sigma}\right) \right\} \\
& + I[y_i = 7] \times \log_e \left\{ \Phi\left(\frac{a_K - x_i'\beta}{\sigma}\right) - \Phi\left(\frac{99 - x_i'\beta}{\sigma}\right) \right\} \qquad (3)
\end{aligned}
$$

where Φ denotes the cumulative standard normal distribution function and $I[\cdot]$ is an indicator function that takes the value of 1 when the statement in the square brackets is true and 0 when it is false. The relevant part of the log-likelihood is then triggered by the indicator function for whether the individual falls within one of the seven intervals in question. The maximum likelihood procedure now involves the estimation of the β parameter vector and σ.

Diagnostic tests

Unlike the situation with the ordered probit estimation, the estimated coefficients from an interval regression are interpretable as if y_i^* is observed for each i and estimated $E(y^*|x) = x\beta$ by OLS. That is, the estimated coefficients can be interpreted as the marginal effects (i.e. the change in the percentage of sales reported given a change in the independent variable, holding all other constant). It should be noted that the estimated elements of the β parameter vector are only interpretable in this way due to the assumption that y^* given x satisfies the classical linear model assumptions. If these assumptions do

not hold then the interval regression estimator of β would be inconsistent. As a result, it is important to test the key assumptions of functional form, homoscedasticity and normality, all of which are used to derive the log-likelihood function in Equation 3. Unfortunately, these assumptions, despite their importance, are seldom tested in applied research involving such models.[11] Equally, these assumptions, particularly homoscedasticity and normality, are rarely met when using micro data.

Chesher and Irish (1987) outlined diagnostic tests for (pseudo) functional form, normality and homoscedasticity for the ordered probit model that are easily modified for the interval regression model. These tests are all score (or Lagrange Multiplier) tests for which the test statistics take the form:

$$\xi = 1'F(F'F)^{-1}F'1 \tag{4}$$

where 1 is an n-dimensional vector of ones, and F is a matrix with row order n where each row contains the score contributions for all the parameters of the model. ξ can be easily calculated as n times the noncentred R^2 from a regression of 1 on the columns of F.

The construction of the relevant F matrices are based on computations of the pseudo residuals. Usually, residuals are defined as the difference between the observed and estimated values of the dependent variable. However, the estimated values of the dependent variable obtained in the interval regression have no counterpart in the data. Rather, pseudo errors need to be computed. An expression for the pseudo errors is obtained by differentiating Equation 3 with respect to the intercept and for the i-th individual is denoted as:

$$u_i = \frac{\phi\left(\frac{a_{(j-1)i} - x_i'\beta}{\sigma}\right) - \phi\left(\frac{a_{ji} - x_i'\beta}{\sigma}\right)}{\left[\Phi\left(\frac{a_{ji} - x_i'\beta}{\sigma}\right) - \Phi\left(\frac{a_{(j-1)i} - x_i'\beta}{\sigma}\right)\right]} \tag{5}$$

where $\phi(\cdot)$ denotes the standard normal probability density function and a_{j-1} and a_j denote the known interval parameters for individual i. The pseudo residuals, e_i, are obtained by replacing the unknown parameters in Equation 5 with their maximum likelihood estimates.

For the homoscedasticity and normality tests, higher order moment residuals are required, specified as:

$$M_{\tau i} = \frac{\left(\frac{a_{(j-1)i} - x_i'\beta}{\sigma}\right)^{\tau}\phi\left(\frac{a_{(j-1)i} - x_i'\beta}{\sigma}\right) - \left(\frac{a_{ji} - x_i'\beta}{\sigma}\right)^{\tau}\phi\left(\frac{a_{ji} - x_i'\beta}{\sigma}\right)}{\sigma\left[\Phi\left(\frac{a_{ji} - x_i'\beta}{\sigma}\right) - \Phi\left(\frac{a_{(j-1)i} - x_i'\beta}{\sigma}\right)\right]} \tag{6}$$

The higher order moment residuals are obtained by replacing the unknown parameters in Equation 6 with their maximum likelihood estimates. The first four moment residuals are required for the desired tests and are defined as follows:

$$\begin{aligned} e_i^1 &= e_i \\ e_i^2 &= \hat{M}_{1i} \\ e_i^3 &= 2e_i^1 + \hat{M}_{2i} \\ e_1^4 &= 3e_i^2 + \hat{M}_{3i} \end{aligned} \tag{7}$$

The F matrix, or score contributions, is obtained by multiplying the pseudo residuals by the various auxiliary variables in question.

Pseudo functional form test. The (pseudo) functional form test is a modified version of the Ramsey Regression Equation Specification Error Test (RESET) (Ramsey, 1969). F is given as

$$F = \left(e^1 x, e^1 \hat{y}^{*2}, \ldots, e^1 \hat{y}^{*K}, e^2\right) \tag{8}$$

where x includes a column of 1s if the grouped model contains an intercept and $\hat{y}^{*K}$ is the K-th power of model's predicted standardized index, $\hat{y}^* = x'\hat{\beta}/\sigma$. That test statistic ξ is distributed as $\chi^2(K-1)$. If the null hypothesis is rejected, this is evidence of model misspecification that may be rectified by considering additional variables and/or alternative functional forms (such as a semi-log model).

Test for homoscedasticity. For the test of homoscedasticity, F is given as[12]:

$$F = \left(e^1 x, e^2 d_i\right) \tag{9}$$

The test statistic ξ is distributed as $\chi^2(q)$ where q is equal to the number of variables that are interacted with e^2. If the form of the heteroscedasticity is unknown, then $d_i = xx'$. This is simply White's test for heteroscedasticity of unknown form, modified for

[11] Caudill and Jackson (1993) discussed the implications of heteroscedasticity in the interval regression model and apply the correction outlined here but they neither formally test for heteroscedasticity nor discuss the functional form and normality assumptions and associated tests. Reilly *et al.* (2004) discussed and applied the tests for functional form, heteroscedasticity and normality but, while they found that their model suffers from heteroscedasticity and nonnormality, they did not make any modifications to their model based on these results.

[12] When an intercept is estimated so that x always contains a unit element, e^2 is redundant in the test for homoscedasticity.

Downloaded By: [Tedds, Lindsay] At: 15:49 15 April 2009

the interval regression model. The test is sensitive to model misspecification so if the test for (pseudo) functional form is rejected, it is also likely that the test for homoscedasticity will also be rejected.

As indicated previously, the estimated coefficients are inconsistent in the presence of uncorrected heteroscedasticity but it is possible to address this inconsistency by correcting for heteroscedasticity. Typically, the 'sandwich' estimator of the variance is used in place of the conventional maximum likelihood covariance estimator. This estimator is expressed as:

$$\mathrm{Var}(\hat{\beta}) = [I(\hat{\beta})]^{-1}\left(x_i' e_i^2 x_i\right)[I(\hat{\beta})]^{-1} \tag{10}$$

where $I(\hat{\beta})$ is the information matrix for the $\hat{\beta}$ vector, computed at the maximum likelihood estimates. However, unlike with linear regression, in an interval regression model it is not possible to correctly specify $E(y^*|x)$ but misspecify $\mathrm{Var}(y^*|x)$ which implies that the validity of this correction is highly questionable in the interval regression model and, hence, is not used in this article. In this situation, the preferred choice is to correct for heteroscedasticity using a variation of Harvey's (1976) multiplicative heteroscedasticity, denoted as $\mathrm{Var}(u_i) = \exp(d_i\gamma)^2$ where d_i is a matrix of independent variables, including a column of one's, that are the source of the heteroscedasticity and γ is a coefficient vector. The expression $\exp(d_i\gamma)$ replaces the σ noted in Equation 3.[13]

Normality test. Finally, F in the usual $\chi^2(2)$ test for zero skewness and/or excess kurtosis is given by:

$$F = \left(e^1 x, e^2, e^3, e^4\right) \tag{11}$$

If the hypothesis of normality is rejected, this implies that the resulting coefficient estimates will be biased and may be inconsistent. However, it is also well known that normality will often be rejected if heteroscedasticity is present in the model.

Explanatory and control variables

As was outlined above, previous empirical work using audit data found relationships between tax noncompliance and various economic, political and firm characteristics and these relationships will also be explored in this article. Generally anything that is perceived by the firm as affecting their ability to conduct business provides incentive for the firm to engage in tax noncompliance, in part to reduce costs and be more competitive, but it is generally acknowledged that factors such as high tax rates, corruption, extortion by organized crime and the legal environment are primary drivers of tax evasion. It is widely acknowledged high taxes, whether they be corporate, payroll and/or sales taxes, provide incentives for firms to hide output to reduce their tax burden. If corruption is common, then among other effects, it will increase the cost of business, reduce morality and reduce a firm's confidence in government; all of which are likely to have a negative relationship with tax compliance. The relationship between tax compliance and inconsistency in the application of laws and regulations, however, is more ambiguous. If a firm can individually garner the favour of the government(s) and/or courts in the interpretation and application of laws and regulations, then this may reduce tax noncompliance. On the other hand, if the firm does not benefit from this inconsistency, then it may resort to tax noncompliance.

Other drivers of firm tax noncompliance include: access to capital, inflation, exchange rate, political instability and regulatory burden. Andreoni (1992) argued that '... individuals facing binding borrowing constraints may use tax evasion to transfer resources from the future to the present. Even if a person finds tax evasion undesirable in the absence of borrowing constraints, it could become desirable if a borrowing constraint is binding. Tax evasion, therefore, may be a high-risk substitute for a loan.' (pp. 35–36). Price inflation may influence under-reporting in two ways. First, if tax rates are not indexed then the tax burden rises simply due to 'bracket creep'. Second, inflation generates uncertainty and businesses may hedge against this uncertainty by engaging in (more) underground activity. Exchange rates can be an obstacle to doing business if the exchange rate is unpredictable and can increase the input costs faced by businesses. Various governmental factors can also influence under-reporting. Political instability constrains a countries ability to collect taxes and also reduces the efficiency of collection, thereby increasing noncompliance. The degree of regulation is often cited as a factor that influences people to engage in underground activity as regulations reduce a firms' freedom of choice (e.g. Deregulation Commission, 1991; Schnieder and Enste, 2000).

Various characteristics of the firm will also be included as control variables. The previous literature makes it clear that tax noncompliance likely varies across governance models. The WBES provides an opportunity to explore the relationship between the legal structure of the business and tax noncompliance. Based on the literature, it is

[13] In STATA, the heteroscedastic corrected interval regression model is estimated using the 'het' option on the 'intreg' command.

Downloaded By: [Tedds, Lindsay] At: 15:49 15 April 2009

Table 2. Data summary

Variable	Obs.	Mean	SD	Minimum	Maximum
% of sales reported	8153	2.978	2.166	1	7
Causes of under-reporting					
Financing	7526	0.811	0.391	0	1
Political instability	7376	0.842	0.365	0	1
Organized crime	7196	0.631	0.482	0	1
High taxes and regulations	7660	0.892	0.311	0	1
Corruption	7485	0.586	0.497	0	1
Inflation	7425	0.846	0.361	0	1
Exchange rate	7259	0.741	0.438	0	1
Laws and regulations inconsistent	8029	0.478	0.493	0	1
Legal organization of company					
Sole proprietary	8153	0.180	0.384	0	1
Partnerships	8153	0.173	0.378	0	1
Private corporation	8153	0.280	0.449	0	1
Public corporation[a]	8153	0.130	0.336	0	1
Other business	8153	0.154	0.361	0	1
Industry sector					
Manufacturing[a]	7418	0.370	0.483	0	1
Service	7418	0.424	0.494	0	1
Other	7418	0.035	0.185		
Agriculture	7418	0.075	0.264	0	1
Construction	7418	0.096	0.294	0	1
Firm size					
Small	8139	0.393	0.488	0	1
Medium[a]	8139	0.412	0.492	0	1
Large	8139	0.194	0.396	0	1
Firm age					
<5	8153	0.257	0.437	0	1
5–15[a]	8153	0.354	0.478	0	1
>15	8153	0.389	0.487	0	1
Number of competitors					
No competitors	7907	0.111	0.314	0	1
1–3[a]	7907	0.399	0.490	0	1
>3	7907	0.490	0.500	0	1
Other					
Foreign owned	8153	0.183	0.387	0	1
Government owned	8153	0.120	0.325	0	1
Exporter	8153	0.339	0.473	0	1
Financial statements audited	8153	0.586	0.493	0	1

Note: [a] Denotes the omitted category in estimation.

Downloaded By: [Tedds, Lindsay] At: 15:49 15 April 2009

expected that publicly traded corporations will be the most compliant. In most countries, publicly traded corporations have a greater probability of being audited and are subjected to public disclosure requirements and independent financial auditing, which tends to expose any under-reporting behaviour to the authorities. As audit levels and detection probability are greater, compliance should be higher. Industry sector, firm size, number of competitors, age, ownership, export status and whether the firm subjects its financial statements to audits will also be included. Table 2 provides summary statistics for the key variables described in this section.

V. Results

The first task is to specify the values of the two end intervals, denoted above as a_0 and a_K

Downloaded By: [Tedds, Lindsay] At: 15:49 15 April 2009

in Equation 3.[14] It is difficult to argue that a firm would report <0% of their sales. Instead, it could be argued that the first interval should begin at either 1 or 0%: 1% since it can be argued that firms that report 0% of their sales would likely be operating completely as 'ghosts' and would not have been selected for an interview since there would be no formal record of the firm. The results are quite insensitive to the choice of a_0 hence the value of a_0 is set to 1%, which seems economically reasonable given the sampling design which excluded unregistered firms.

The values of ∞ and 100 are considered for the values of a_K. When the last interval is treated as open ended ($a_K = \infty$), the maximum predicted value is in excess of 140 and over 25% of the predicted values exceed 100%. These numbers seem economically implausible, particularly that over 25% of firms report in excess of 100% of their sales to the tax authority. Hence, the last interval is set to 100. With the bounds of the intervals set at 1 and 100, the maximum predicted value is ~105 and ~1% of the predicted values exceed 100%. These figures seem to be economically plausible and accord with the findings of Rice (1992) who found that 6% of firms do over-report to some extent.

It is also important to consider the necessary assumptions for consistent estimates and in particular, the assumption of normality. It is anticipated that the assumption of normality is likely violated in this model. Table 1 clearly shows that the distribution of sales reporting is negatively skewed. Normality may be more closely approximated by taking the natural logarithm of 100 minus the dependent variable, which is simply the natural logarithm of the percentage of sales not reported to the tax authority. Such a transformation requires that the dependent variable take only strictly positive values and care exercised in interpreting the resulting estimates. In particular, the coefficients in such a model are no longer marginal effects but 100 multiplied by the estimated coefficient represents the percentage change. Preliminary investigation supports this transformation and is the specification used to obtain the results reported here.

The results are presented in Table 3. The basic model is presented in the second column (Model 1a). Various goodness-of-fit measures and the results from the various diagnostic tests are reported near the end of the table. Larger (less negative) log-likelihood values are indicative of a better fit. The log-likelihood values can be compared across the models only if they have the same samples and dependent variable. The R^2, for technical reasons, cannot be computed in the same way in interval regressions as it is in OLS regression. Various pseudo-R^2 measures, however, have been proposed, but there is no generally accepted measure. Veall and Zimmermann (1996) recommended the measure of McElvey and Zavoina (1975), which is reported in Table 3.[15]

For Model 1a, the null of correct functional form is not rejected at all the usual significance levels, but the tests for homoscedasticity and normality are both rejected at the usual significance levels. Plots of the pseudo residuals from this models supports the finding of nonnormality and it is not clear if any further transformations of the dependent variable, beyond those already pursued, would result in normality. In addition, further transformations would make it difficult to interpret the estimated coefficients. It is, however, possible to correct for the heteroscedasticity, as noted above, if a form for it is specified and, as noted above, correcting for the heteroscedasticity can often achieve normality in the (pseudo) residuals. It is not unreasonable to assume that much of the heteroscedasticity results from countries and business type. Applications of the test for homoskedasticity of known form across subsets of the variables confirm this suspicion. Model 1a is re-estimated to account for this known form of the heteroscedasticity and these results are presented in Table 3 under the column heading Model 1b.[16] The first thing to note is that while correcting for heteroscedasticity is important on theoretical grounds, in this application the practical importance is not terribly striking as there are no sign changes in the coefficient estimates or changes in statistical significance.

These results for Model 1b will now be discussed in detail. With respect to the causal variables, firms that perceive high taxes and regulations, organized crime and government corruption as obstacles to doing businesses report significantly less of their sales to the

[14] It should be noted that the sensitivity of the results and predicted values to the treatment of the end intervals is often overlooked in the literature, yet can be of clear importance to the economic validity of the resulting estimates.
[15] The McElvey and Zavoina's (1975) R^2 is computed in STATA with the 'fitstat' command.
[16] The diagnostics tests for functional form, heteroscedasticity and normality that are outlined in this article are not performed on the heteroscedasticity corrected model. Not only are these tests much more difficult in the presence of heteroscedasticity (Gasser *et al.*, 1986) but the form, distribution and power of these tests have not been confirmed for the heteroscedasticity corrected model. The author is unaware of any existing work that addresses these issues and these theoretical issues constitute an area for future exploration.

Table 3. Estimation results for semi-log model (dependent variable: natural log of percentage of sales not reported to tax authority)

Variable	Model 1a		Model 1b Heteroscedasticity correlation	
Constant	0.185	(0.182)*	0.934	(0.155)*
Causes of under-reporting				
Obstacle: financing	0.096	(0.04)**	0.089	(0.040)**
Obstacle: political instability	0.012	(0.048)	−0.022	(0.043)
Obstacle: organized crime	0.092	(0.045)**	0.076	(0.043)*
Obstacle: taxes and regulations	0.212	(0.049)***	0.180	(0.046)***
Obstacle: corruption	0.527	(0.045)***	0.513	(0.044)***
Obstacle: inflation	−0.013	(0.051)	0.010	(0.047)
Obstacle: exchange rate	0.068	(0.047)	−0.045	(0.044)
Obstacle: laws and regulations inconsistent	−0.067	(0.04)*	−0.046	(0.038)
Control variables[a]				
Sole proprietary	0.211	(0.082)***	0.185	(0.075)***
Partnerships	0.031	(0.077)	0.024	(0.072)
Other	0.040	(0.086)	0.097	(0.083)
Private corporation	0.131	(0.066)**	0.133	(0.059)**
Other industry	−0.069	(0.136)	0.064	(0.127)
Service	−0.070	(0.046)	−0.052	(0.042)
Agriculture	−0.096	(0.083)	−0.081	(0.080)
Construction	0.029	(0.072)	0.053	(0.069)
Small: <50 employees	0.174	(0.050)***	0.186	(0.047)***
Large: >500 employees	−0.163	(0.055)***	−0.176	(0.052)***
<5 years of age	−0.041	(0.050)	−0.071	(0.044)
>15 years of age	−0.063	(0.050)	−0.065	(0.046)
No competitors	−0.064	(0.065)	−0.111	(0.062)*
>3 Competitors	0.148	(0.060)**	0.101	(0.0553)*
Foreign owned	−0.124	(0.053)**	−0.093	(0.049)*
Government owned	−0.150	(0.071)**	−0.176	(0.063)***
Export	−0.049	(0.046)	−0.036	(0.042)
Audits	−0.239	(0.048)***	−0.174	(0.045)***
Country dummies	Yes		Yes	
σ (SE)	1.347	(0.013)	–	
Pseudo log-likelihood value	−11788.379		−11613.352	
McKelvey and Zavoina pseudo-R^2	0.228	–		
LRT-OS/Wald Chi[b] [d.o.f.; p-value]	1357.86	[102; 0.000]	16652.51	[102; 0.000]
Pseudo functional form test [p-value]	0.001	[0.979]	–	
Homoscedasticity test of unknown form [p-value]	1111.411	[0.000]	–	
Normality test [p-value]	1699.6542	[0.000]	–	
Observations	5393		5393	

Notes: ***, ** and * denote statistical significance at the 1, 5 and 10% levels, respectively. d.o.f. denotes degrees of freedom. SE are noted in parentheses, where the SE reported for Model 1b are heteroscedasticity-consistent.

[a] Omitted categories are public corporations, manufacturing, medium, between 5 and 15 years of age, between 1 and 3 competitors and 'United States'.

[b] For Model 1a, the results for the Likelihood Ratio Test (LRT-OS) is reported. For Model 1b, the results for the Wald test are reported. STATA only reports the LRT-OS for the uncorrected model and only reports the Wald for the heteroscedasticity correct model. These two tests are asymptotically equivalent.

Downloaded By: [Tedds, Lindsay] At: 15:49 15 April 2009

tax authority. The results presented in Table 3 indicate that government corruption has the single largest effect, resulting in the percentage of sales not reported to the tax authority being a whopping 51.3% higher, this is followed by taxes at 18.0% higher, access to financing at 8.9% higher and organized crime at 7.6% higher. In comparison, Johnson *et al*. (2000) also found a positive relationship between noncompliance and government corruption and organized crime but failed to find a relationship between compliance and tax payments. All other variables are statistically insignificant.

The control variables also provide interesting results and these results are also in accordance with those obtained using tax audit data. Sole proprietorships and private corporations are the least compliant form business of all the business models included in this study. This result accords with those found by Rice (1992), who found that the public corporations were more compliant than other types of firms, and Hanlon *et al*. (2007) who found that private corporations were less compliant. Small firms and firms with more than three competitors are also found to be the least compliant while large firms and firms with no competitors are more compliant. Giles (1997b) reported a similar result with respect to size, whereas Rice (1992) found firm size and tax noncompliance were positively related and Hanlon *et al*. (2007) found that compliance was a convex function over firm size. Finally, being foreign, government-owned and having financial statements audited all lead to significantly higher compliance. Both Giles (1997b) and Chan and Mo (2000) found that foreign-owned firms are more compliant, and the results in Table 3 provide further support for this result. Chan and Mo (2000) reported that export-oriented firms are more compliant but this characteristic is found to be insignificant. Previous empirical studies all found a significant relationship between industry and tax noncompliance whereas industry is found to have no relationship with reporting behaviour, however, it must be kept in mind that the industry the categories included in the WBES dataset are likely too broadly defined to yield any informative results. The relationship between internal audit controls of the firm and tax compliance has not been investigated in previous empirical work.

VI. Conclusion

Very little is actually known about firm tax compliance due to a lack of detailed and readily available data. The purpose of this article was to use a unique and recently available dataset that contained information on firms from around the world to investigate some of the factors that effect business tax compliance. This is one of the first studies to examine firm tax compliance using worldwide data. The majority of previous empirical studies were confined to examining firms within a particular country using tax audit data.

The empirical strategy employed in this article exploits the nature of the dependent variable, which is interval coded, and uses interval regression. The estimated coefficients from an interval regression are interpretable as marginal effects provided that the model satisfies the assumptions of correct functional form, homoscedasticity and normality. These assumptions are investigated using standard diagnostic tests that have been modified for the interval regression model. The test results support the use of the semi-log model that is corrected for heteroscedasticity, the origin of which is found to result at the country level and from the business type.

Overall, evidence is presented that shows that the firms in all regions around the world engage in tax noncompliance, but that there is substantial variation within regions. The detailed regression results indicate that, not surprisingly, taxes, organized crime, government corruption, and access to financing are all positively related to tax noncompliance and the magnitude is quite large in each case. Inflation, political instability, exchange rates, and the fairness of the legal system were found to have no effect on under-reporting. There is also a large correlation between the legal organization of a business and under-reporting, with sole proprietorships and private corporations being the least compliant. Firm size is also correlated with tax compliance, with small firms reporting less and large firms more of their sales to the tax authority.

The findings do suggest a role for public policy, as well as actions to be considered by the tax authority and items that require further study. First, the findings suggest that administrations interested in reducing business tax noncompliance should consider reducing business taxes, minimizing the number of regulations, and reducing, if not eliminating, government corruption. Admittedly, taking action on these issues is complex and involves more than just the tax authority. Second, tax authorities should consider auditing sole proprietorships and private corporations at a higher rate and requiring all firms to have their financial statement audited by a third party. Finally, further exploration is required to understand the compliance rates of foreign owned firms, firms in a competitive industry, and private corporations.

Downloaded By: [Tedds, Lindsay] At: 15:49 15 April 2009

Investigation into these relationships appears to be a worthwhile venture.

In addition, firm level data has recently been made available that was collected following 2000. This data includes similar variables to those included in this study, including the dependent variable. It would be worthwhile to pool these new data sources together with the WBES to investigate firm under-reporting in more detail.

Acknowledgements

I would like to thank Giorgio Valente, an anonymous referee, Tom Crossley, Mike Veall, David Bjerk, Wayne Simpson, David Giles and seminar participants at the University of Victoria, the 2006 Canadian Economics Association and the 2005 Tax Policy: Perspectives from Law and Accounting conference for their helpful comments and invaluable guidance. I am grateful for the diligent research support provided by Hassan Azziz and Adian McFarlane. I would also like to gratefully acknowledge the financial support from the Social Sciences and Humanities Research Council (No. 752-2004-1096 and No. 501-2002-0107), the University of Manitoba and the University of Victoria. I am solely responsible for any remaining errors and omissions.

Downloaded By: [Tedds, Lindsay] At: 15:49 15 April 2009

References

Allingham, M. and Sandmo, A. (1972) Income tax evasion: a theoretical analysis, *Journal of Public Economics*, **1**, 323–38.

Andreoni, J. (1992) IRS as loan shark: tax compliance with borrowing constraints, *Journal of Public Economics*, **49**, 35–46.

Andreoni, J., Erard, B. and Feinstein, J. (1998) Tax compliance, *Journal of Economic Literature*, **36**, 818–60.

Apel, M. (1994) An expenditure-based estimate of tax evasion in Sweden, Working Paper No. 1, Department of Economics, Uppsala University, Sweden.

Bajada, C. (1999) Estimates of the underground economy in Australia, *The Economic Record*, **75**, 369–84.

Bajada, C. (2005) Unemployment and the underground economy in Australia, *Applied Economics*, **37**, 177–89.

Batra, G., Kaufmann, D. and Stone, A. H. W. (2003) *Investment Climate Around the World: Voices of the Firms from the World Business Environment Survey*, The World Bank, Washington, DC.

Besim, M. and Jenkins, G. P. (2005) Tax compliance: when do employees behave like the self-employed?, *Applied Economics*, **37**, 1201–8.

Caudill, S. B. and Jackson, J. D. (1993) Heteroscedasticity and grouped data regression, *Southern Economic Journal*, **60**, 128–35.

Chan, K. H. and Mo, P. L. L. (2000) Tax holidays and tax non-compliance: an empirical study of corporate tax audits in China's developing economy, *The Accounting Review*, **75**, 469–84.

Chesher, A. and Irish, M. (1987) Residual analysis in the grouped and censored normal regression model, *Journal of Econometrics*, **34**, 33–61.

Christian, C. W. (1994) Voluntary compliance with individual income tax: results from the 1988 TCMP study, *The IRS Research Bulletin Publication*, **9**, 35–42.

Cowell, F. (2004) Sticks and carrots, in *The Crises in Tax Administration* (Eds) H. Aaron and J. Slemrod, The Brookings Institution Press, Washington, DC, pp. 230–58.

Dell'Anno, R., Gómez-Antonio, M. and Alañon-Pardo, A. (2007) The shadow economy in three Mediterranean countries: France, Spain and Greece. A MIMIC approach, *Empirical Economics*, **33**, 51–84.

Deregulation Commission (1991) Opening of markets and competition, Report presented to German federal government, Bonn.

Draeseke, R. and Giles, D. E. A. (2002) A fuzzy logic approach to modelling the underground economy, *Mathematics and Computers in Simulation*, **59**, 115–23.

Engström, P. and Holmlund, B. (2006) Tax evasion and self-employment in a high-tax country: evidence from Sweden, Working Paper No. 12, Department of Economics, Uppsala University, Sweden.

Erard, B. (1992) The influence of tax auditing on reporting behaviour, in *Why People Pay Taxes* (Ed.) J. Slemrod, University of Michigan, Ann Arbor.

Feige, E. L. (1979) How big is the irregular economy?, *Challenge*, **22**, 5–13.

Feldman, N. and Slemrod, J. (2007) Estimating tax noncompliance with evidence from unaudited tax returns, *Economic Journal*, **117**, 327–52.

Frey, B. S. and Weck-Hanneman, H. (1984) The hidden economy as an unobserved variable, *European Economic Review*, **26**, 33–53.

Gasser, T., Sroka, L. and Jennen-Stinmetz, C. (1986) Residual variance and residual pattern in nonlinear regression, *Biometrika*, **73**, 625–33.

Giles, D. E. A. (1997a) Causality between the measured and underground economies in New Zealand, *Applied Economics Letters*, **4**, 63.

Giles, D. E. A. (1997b) The hidden economy and tax-evasion prosecutions in New Zealand, *Applied Economics Letters*, **4**, 281.

Giles, D. E. A. (1999) The rise and fall of the New Zealand underground economy: are the responses symmetric?, *Applied Economics Letters*, **6**, 185–9.

Giles, D. E. A. and Caragata, P. K. (2001) The learning path of the hidden economy: tax and growth effects in New Zealand, *Applied Economics*, **33**, 1857–67.

Giles, D. E. A. and Johnson, B. J. (2002) Taxes, risk-aversion and the underground economy: a nonparametric analysis of New Zealand data, *Pacific Economic Review*, **7**, 97–113.

Giles, D. E. A. and Tedds, L. M. (2002) *Taxes and the Canadian Hidden Economy*, Canada Tax Foundation, Toronto.

Giles, D. E. A., Tedds, L. M. and Werkneh, G. (2002) The Canadian underground and measured economies: Granger causality results, *Applied Economics*, **34**, 2347–52.

Gutmann, P. M. (1977) The subterranean economy, *Financial Analysts Journal*, **34**, 24–7.

Hanlon, M., Mills, L. and Slemrod, J. (2007) An empirical examination of corporate tax non-compliance, in *Taxing Corporate Income in the 21st Century* (Eds) A. Auerbach, J. R. Hines Jr and J. Slemrod, Cambridge University Press, Cambridge, pp. 171–210.

Harvey, A. (1976) Estimating regression models with multiplicative heteroscedasticity, *Econometrica*, **44**, 461–65.

Hill, R. and Kabir, M. (2000) Currency demand and the growth of the underground economy in Canada, 1991–1995, *Applied Economics*, **32**, 183–92.

Internal Revenue Service (2004) *Interactive Tax Gap Map*, National Headquarters Office of Research, USA.

Johnson, S., Kaufmann, D., McMillan, J. and Woodruff, C. (2000) Why do firms hide? Bribes and unofficial activity after communism, *Journal of Public Economics*, **76**, 495–520.

Joulfaian, D. (2000) Corporate tax evasion and managerial preferences, *Review of Economics and Statistics*, **82**, 698–701.

Joulfaian, D. and Rider, M. (1998) Differential taxation and tax evasion by small business, *National Tax Journal*, **51**, 676–87.

Lyssiotou, P., Pashardes, P. and Stengos, T. (2004) Estimates of the black economy based on consumer demand approaches, *Economic Journal*, **114**, 622–40.

Matthews, K. G. P. (1984) The GDP residual error and the black economy: a note, *Applied Economics*, **16**, 443.

McElvey, W. and Zavoina, R. (1975) A statistical model for the analysis of ordinal level dependent variables, *Journal of Mathematical Sociology*, **4**, 103–20.

Pissarides, C. A. and Weber, G. (1989) An expenditure-based estimate of Britain's black economy, *Journal of Public Economics*, **39**, 17–32.

Ramsey, J. B. (1969) Tests for specification errors in classical linear least-squares regression analysis, *Journal of the Royal Statistical Society*, **31**, 350–71.

Reilly, B., Krstic, G. and Cominetta, M. (2008) Purchasing from informal markets: evidence from Serbia, *Applied Economics Quarterlys*, **52**, 30–50.

Rice, E. M. (1992) The corporate tax gap: evidence on tax compliance by small corporations, in *Why People Pay Taxes* (Ed.) J. Slemrod, University of Michigan Press, Ann Arbor, pp. 125–66.

Schnieder, F. and Enste, D. (2000) Increasing shadow economies all over the world – fiction or reality? A survey of the global evidence of their size and their impact from 1970 to 1995, *Journal of Economic Literature*, **38**, 77–114.

Schuetze, H. J. (2002) Profiles of tax non-compliance among the self-employed in Canada, 1969–1992, *Canadian Public Policy*, **28**, 219–37.

Stewart, M. (1983) On least squares estimation when the dependent variable is grouped, *Review of Economic Studies*, **50**, 737–53.

Tanzi, V. (1980) The underground economy in the United States: estimates and implications, *Banco Nazionale del Lavoro*, **135**, 427–53.

Tedds, L. M. (2005) The underground economy in Canada, in *Size, Causes and Consequences of the Underground Economy* (Eds) C. Bajada and F. Schneider, Ashgate Publishing, London, pp. 157–176.

Tedds, L. M. (2007) Estimating the income reporting function for the self-employed, MPRA Paper No. 4213, University Library of Munich, Germany, pp. 157–76.

Veall, M. R. and Zimmermann, K. F. (1996) Pseudo-R^2 measures for some common limited dependent variable models, *Journal of Economic Surveys*, **10**, 241–59.

Wik, M., Kebede, T. A., Bergland, O. and Holden, S. T. (2004) On the measurement of risk aversion from experimental data, *Applied Economics*, **36**, 2443–51.

World Bank (1997) *World Development Report 1997: The State in a Changing World*, World Bank Group, Washington, DC.

Downloaded By: [Tedds, Lindsay] At: 15:49 15 April 2009