Forthcoming in Journal of Political Economy

Tax Rates and Tax Evasion: Evidence from "Missing Imports" in China

*Raymond Fisman**
Columbia Business School and NBER

*Shang-Jin Wei***
IMF, Brookings, NBER, and CEPR

First Draft: March 1, 2001
This Draft: October 29, 2003

Abstract

Tax evasion, by its very nature, is difficult to observe. We quantify the effects of tax rates on tax evasion, by examining the relationship in China between the tariff schedule and the 'evasion gap' which we define as the difference between Hong Kong's reported exports to China at the product level and China's reported imports from Hong Kong. Our results imply that a one percentage point increase in the tax rate is associated with a 3 percent increase in evasion. Furthermore, the evasion gap is negatively correlated with tax rates on closely related products, suggesting that evasion takes place partly through mis-classification of imports from higher-taxed categories to lower-taxed ones, in addition to under-reporting the value of imports.

JEL Classification: H2, F1
Key Words: Tax evasion, Laffer curve, corruption

---

* Meyer Feldberg Associate Professor, 823 Uris Hall, Graduate School of Business, Columbia University, New York, NY 10027, phone: (212) 854-9157; fax: (212) 854-9895; email:rf250@columbia.edu. ** Advisor and Head of Trade Unit, IMF, and Senior Fellow, Brookings Institution, 1775 Massachusetts Ave., Room 731, Washington, DC 20036, phone: 202/797-6023, fax: 202/797-6181, swei@brook.edu.

We thank Jerzy T. Rozanski, Zhi Wang, and Yi Wu for help with extracting the data, Bill Gale, Jim Hines, Nick Lardy, Sung Yung-Wing, and seminar participants at an NBER meeting, the University of Chicago, University of Maryland, World Bank, Harvard University, the University of Toronto, and the University of Southern California for helpful discussion and Hayden Smith for editorial assistance.

1. Introduction

This paper studies the responsiveness of tax evasion to tax rates. Much of the work in the theory and empirics of taxation has taken tax collection as given and often costlessly executed. This simplification is unlikely to be realistic: even within the United States, where tax collection is considered to be relatively efficient, about 17 percent of income taxes are estimated as unpaid (Slemrod and Yitzhaki, 2000). One particularly important issue is understanding the relationship between tax rates and tax evasion. A number of models have evolved to incorporate tax evasion, but these models fail to provide any prediction regarding the uniform impact of tax rates on evasion. In the pioneering work of Allingham and Sandmo (1972), the relationship between tax rates and evasion is positive, but this depends on particular assumptions of risk aversion and the punishment for evasion. A broader review of the literature reports that, more generally, theoretical predictions of the effect of tax rates on evasion are highly sensitive to modeling assumptions (Slemrod and Yitzhaki, 2000).[1] Furthermore, even if the effect of tax rates on evasion may be signed, there is still a need to assess the magnitude of this effect. Hence, empirically examining the effect of tax rates on evasion would be very useful from the perspectives of both theory and policy. This has proven to be a challenging task due to the difficulties in measuring evasion, which by definition is not directly observed.

---

[1] Some models even yield strong predictions that run counter to the conventional wisdom that higher tax rates increase evasion. For example, the model of Yitzhaki (1974) predicts that if the punishment for evasion is dependent on the value of taxes evaded (as is the case in China), tax rate increases will reduce evasion.

A number of indirect approaches have been used to infer the behavior of tax evasion from measurable quantities such as currency demand or the discrepancies between national income and product accounts (e.g., Gutmann, 1977, Feige, 1979, Tanzi, 1980). These approaches have been criticized by Slemrod and Yitzhaki in their survey paper on the subject, since "[n]one of these approaches is likely to be reliable…as their accuracy depends either on unverifiable assumptions or on how well the demand for currency is estimated" (Slemrod and Yitzhaki, 2000). Furthermore, these approaches do not naturally generate an estimate of the responsiveness of evasion to tax rates.

As a more direct approach to examining tax evasion, researchers have used data from the U.S. Taxpayer Compliance Measurement Program (TCMP) conducted by the U.S. Internal Revenue Service. Based on intensive audits of a random sample of tax returns, the data set gives information on reported taxable income and what auditors later conclude to be true taxable income. Using these data, Clotfelter (1983) estimated that tax evasion is positively associated with tax rates, with the elasticity ranging from 0.5 to 3. Feinstein (1991), using a short panel of two years of TCMP data (1982 and 1985), found that increasing the marginal tax rate has a negative effect on evasion, contradicting Clotfelter's conclusion. However, the main source of variation on tax rates in both of these studies comes from differential marginal tax rates across income levels, so it is not really possible to disentangle tax rate effects from income effects.

In this paper, we take a new approach in measuring the effect of tax rates on evasion that is less likely to be contaminated by such problems. Specifically, we examine evasion in China's imports from Hong Kong, at a very disaggregated level (e.g., four-door passenger car), by comparing Hong Kong's reported exports and China's reported imports of the same products. In the absence of evasion (and measurement error), China and Hong Kong-reported numbers should be the same. So far, the extent to which they differ has generally been taken to be measurement errors (see, for example, Feenstra and Hanson, 2000). However, when we match these data up with product-specific tax rates in China (tariff plus value-added tax rates), we find that this 'evasion gap' is highly correlated with Chinese tax rates: much more value is 'lost' for products with higher tax rates. Our methodology is related to that of Pritchett and Sethi (1994), who find that tax revenues divided by imports increase at a rate less than the official tax rate, in a sample of four developing countries. Note, however, that our analysis is at a much higher level of disaggregation; furthermore, they are unable to disentangle illegal tax *evasion* from legal tax *avoidance* (e.g., taking advantage of tax loopholes and special exemption). Tax avoidance, as it is legal, is more readily observed than evasion.

Another novel feature of our study is that we are able to differentiate three different aspects of tax evasion: under-reporting of unit value, under-reporting of taxable quantities, and the mislabeling of higher-taxed products as lower-taxed products. We find strong evidence of mislabeling and limited evidence of under-reporting of unit value; on the other hand, once shifting reported imports from a higher- to a lower-taxed category is controlled for, we do not find evidence of under-reporting of overall imported quantities. In looking at the effects of changes in tax rates between 1997 and 1998 on changes in evasion, we obtain similar results. Finally, when we use a flexible functional form, we find that tax evasion occurs mostly at higher tax rates. The rest of the paper is organized as follows: Section 2 describes the data on taxes and imports/exports. Section

3 provides the details of our empirical specification and the results. Finally, Section 4 concludes.

2. Data

The trade flow data in this paper are taken from the World Bank's World Integrated Trade Solution (WITS) database, which in turn is derived from the United Nations' Comtrade database. These data are collected by the United Nations Statistical Division from individual countries' trade records, and include information on imports and exports for each country, recorded according to the 6-digit Harmonized Commodity Description and Coding System (HS). The United Nations allows individual countries to have a classification system more detailed than the HS 6 digit levels. In the case of China, an 8-digit classification (a refined version of the HS 6-digit classification) is available. However, we choose to use the import data at the 6-digit level in order to be compatible with Hong Kong-reported numbers. The current HS classification system began in 1996, which is also the earliest year for which we have year-end data on tax rates.

For every product that China imports from Hong Kong, we define *Export_Value* as the value of exports reported by Hong Kong destined for China and *Import_Value* as the value of imports reported by China arriving from Hong Kong. The original sample contained 5113 products at the 6-digit classification level in 1998. However, there were missing observations for 2820 classifications for either imports or exports.[2] Of those

---

[2] These were almost exclusively missing observations on both imports and exports; where observations were available for imports and not for exports, it is almost certainly the result of misclassified re-exports

remaining, a further 250 did not have consistent tax rates at the 6-digit level, and were also omitted, leaving a sample of 2043.

Because of Hong Kong's proximity to China and its special status as a former colony, it also does a considerable amount of indirect trade on behalf of other economies (including Taiwan).[3] Hong Kong reports (in the Comtrade database) separate data on indirect as well as direct exports destined for China. China, on the other hand, only reports what it considers to be direct imports from Hong Kong. Indirect imports (say from the U.S. via Hong Kong) are aggregated, in principle, with direct imports from the source country and are reported as part of the imports from that source country. Thus, in theory, China-reported imports from Hong Kong should match up to Hong Kong-reported direct exports to China. However, the data suggests that China cannot always successfully separate indirect imports from direct imports. One likely source of confusion is Taiwan's indirect exports to China via Hong Kong. As the government of Taiwan does not allow its firms to have direct trade with China, Taiwan's exports to China often label Hong Kong as the destination. Sometimes shipping labels are modified while the goods are en route to Hong Kong or in a warehouse in Hong Kong. At other times, an intermediary in Hong Kong is used to record the transaction as an import by Hong Kong from Taiwan plus an export from Hong Kong to China. While the Hong

---

(see below). When all regressions were repeated, using all observations on exports and redefining *Gap_Value* = log(1+*Export_Value*) – log(1+*Import_Value*), the sample size increased by about 2 percent, and our results were virtually identical to those reported in Section 3 below.

[3] Hong Kong's reliance on re-export trade has created the impression that manufacturing activity is virtually nonexistent in Hong Kong itself. To counter this misconception, we refer to a recent study by the Chinese commercial law firm Johnson, Stokes, and Master (JSM), which described Hong Kong's manufactured exports as including, "a wide range of products including clothing, electronics, watches & clocks, jewelry, textiles and chemicals." Their complete report on Hong Kong may be downloaded from: http://www.hg.org/guide-hongkong.html. Note that if we limit our sample to the industry categories implied by JSM, the implied effect of the tax rate on evasion increases somewhat, and the fit of the regressions improves marginally.

Kong customs authorities may understand this as an indirect export to China, China might misclassify at least a portion of such transactions as being direct imports from Hong Kong.[4] For our main results, we report regressions excluding the most 'indirect export intensive' industries, where problems of misclassification are likely to loom largest. We note, however, that results for the full sample are virtually identical, both in terms of the magnitude of the coefficients reported, as well as their significance. This is to be expected: since the same tariff rate applies to both direct and indirect exports, we have no reason to expect misclassification by Chinese customs or importers to be systematically linked to a product's tariff rate.

We define the direct export ratio to be:

$$\textit{Direct Export Ratio} = \frac{\textit{Direct Exports}}{(\textit{Direct Exports} + \textit{Indirect Exports})}$$

We drop from the sample all products where the Direct Export Ratio is below 0.01, which eliminates approximately 18 percent of the sample. Generally, the Direct Export Ratio takes on relatively low values, as illustrated in Figure 1, which gives the frequency distribution of this variable. As this is a cause of some concern, we will consider various cut-off values, ranging from 0.00 to 0.10, in the robustness tests below.[5] Our standard cutoff of 0.01 leaves a final sample size of 1663 products; some regressions involve fewer observations due to missing observations on other regressors (details

---

[4] We thank Professor SUNG Yung-Wing at the Chinese University in Hong Kong for a helpful discussion on this issue.

[5] The low rate of direct exports is certainly a cause for concern. However, there is considerable evidence in the sample statistics that the import numbers are more closely tied to direct, rather than indirect, exports. As some indication of this, the *overall* level of reported imports in our sample is 4,616,684, which is within twenty percent of the total reported level of direct exports (5,344,158). By contrast, the level of indirect exports differs by a factor of five (29,243,699). Furthermore, at the 2-digit level, log(direct exports) and

provided in Section 3). The distribution of the direct export ratio for our final sample is shown in Figure 1.

Comtrade contains data on both the value and quantity of imports/exports; we will utilize both sets of data. In the case of quantities, we are also required to know the units of measurement (e.g., weight, number; area); in most cases, these values match up between the Chinese (import) and Hong Kong (export) data. Where they do not, it is primarily because China reports the weight of imports, while Hong Kong reports the number of units. These observations are not included in the quantity regressions. We define *Export_Qty* to be the total quantity of exports from Hong Kong destined for China as reported by Hong Kong and *Import_Qty* to be the total quantity of imports reported by Chinese customs from Hong Kong into China.

Our basic definition of the evasion gap is given by[6]:

$$Gap\_Value = \log(Export\_Value) - \log(Import\_Value)$$

Thus defined, a larger gap is an indication of greater evasion. We similarly define the gap in quantities reported as:

$$Gap\_Qty = \log(Export\_Qty) - \log(Import\_Qty)$$

---

log(imports) are highly correlated (ρ=0.91), suggesting that imports and direct exports, in aggregate, approximately line up.

[6] Using logs creates problems with those observations where imports are zero; where exports are zero, and imports are non-zero, this is clearly a reporting error resulting most often from mistaken recording of re-exports. To get around this problem, we also ran our analyses using *(Exports – Imports)/(Export + Import)* as the dependent variable. This generated coefficients that were of similar statistical significance and magnitude as those reported in the body of the text.

The data on Chinese tariffs and taxes were also taken from WITS, derived from the UNCTAD TRAINS (Trade Analysis and Information System) database, which gives tariff rates at the 8-digit HS level. Since our import/export data are at the 6-digit level, we need some way of aggregating up the tariff rates to the 6-digit level. Since there is relatively little variation in tax rates at the 8-digit level within a 6-digit category, so we are able to restrict ourselves to the sample for which there are uniform rates at this level of aggregation. In addition to tariffs, there is a value-added tax levied on most imports, which varies from 13 to 17 percent. Our measure of taxation, Tax_{1998}, is the sum of tariff and VAT tax.[7]

The earliest year for which we have detailed data on tariffs is 1996, and our data reflect year-end tariff rates. Unfortunately, for most years since 1996, there were changes in the tariff rates in the middle of the year. Since the import and export data are cumulated for the entire year, matching imports with the appropriate tax rates is challenging. However, in 1998, a single tariff schedule was used throughout the year. Therefore, in our benchmark regressions, we report results utilizing data from 1998, which also happens to be the most recent year for which data were available on both imports and exports. As an extension, we also implement regressions using two years of data (1997 and 1998). We will explain the construction of the 1997 tax rate later in the paper.

---

[7] We argue that the sum of tariffs and the VAT is the most appropriate measure: if an importer already under-reports the value of a particular product, she would also evade the associated VAT. If a car is smuggled into the country without paying the tariff, it is likely to be sold on the black market without paying the VAT. However, as a robustness check, we have also replicated the key regressions restricting the tax variable to include only the tariff rate at customs, but not the VAT. The results are virtually identical to those reported in the main text.

Figures 2(I) and 2(II) illustrate the extent and sources of variation in tariff rate. In Figure 2(I), we give the frequency distribution of tariff rates for all products used in our analyses. As this figure shows, there is considerable variation in tariff rates, in general. While much of this variation occurs across 4-digit categories, there remains substantial within 4-digit variation, as illustrated by Figure 2(II), which shows the difference between the maximum and minimum for each 4-digit category, conditional on this value exceeding zero. For these 936 observations, there is a high frequency of observations at values up to 10. There is thus scope to properly identify the effect of within-category variation with these data.

Part of evasion is thought to take place by mislabeling imported products from a higher-taxed to a lower-taxed type, which is easier for "similar" products.[8] Operationally, two products are considered "similar" if they are in the same 4-digit category. We define *Avg(Tax_o)* to be the average level of *Tax* for all other products in a good's 4-digit class, weighted by *Export_Value*.[9] There are several reasons to use this level of aggregation for examining misclassification. Using a finer division produces almost no variation, so misclassification across 5-digit categories would not provide any benefit to importers. A 3-digit aggregation is overly coarse (see furniture example in the following paragraph).

The phenomenon of misclassification may be illustrated by the anecdotal example of furniture imports, which was reportedly a problem for customs authorities in China in

---

[8] To draw on an example in the context of a more developed economy, the U.S. steel industry provides an interesting case study. There is considerable dispersion of tariff rates within many relatively narrow classes of steel products, and importers in recent years have been caught mislabeling closely related products (e.g., low grade wire as high grade wire; reinforced steel rods as flat rolled steel; see Matthews, 2001).

[9] Since this calculation is based solely on Hong Kong reported values, where we do not expect re-exports to pollute the data, we utilize the full sample. If, in the calculation of *Avg(Tax_O)*, we utilize only the

the mid-1990s (U.S. Department of Commerce, 2000). Furniture are in a 3-digit category (940), which covers a highly heterogeneous range of products (mattresses vs. chairs; wooden vs. metal). At this level, clear differences across products exist, making mislabeling difficult. However, at the 4-digit level, products have a higher degree of similarity (e.g., 9401 – seats versus seat parts), and there is still a considerable variation in tariff rates within each 4-digit category: In 1996, the average difference between the minimum and maximum tax rates within 4-digit furniture categories was 20 percentage points. In that year, the average evasion gap was 64 percent higher for products with the maximum tax rates than those with the minimum tax rates. In 1997, when the Chinese government brought in uniform furniture tax rates, the evasion gaps became more similar across different types of furniture, with the difference between the maximum and minimum gaps narrowing to approximately 3 percent.

There are a number of categories of similar products that maintained their dispersion through to 1998, the year which is used for the bulk of our analyses. As one example, we list in Table 1 the tariff rates for machine tools (category 8459), where we see that the tariff rate depends mainly on whether the machinery is numerically controlled.

For a full list of variables, definitions, and sources, see the Appendix.

Summary statistics for our variables are contained in Table 2(I). One point to note is that the evasion gap has a negative mean. This appears to be due to Chinese customs misattributing some portion of indirect imports as direct imports, as discussed above. This is suggested by the fact that when we exclude observations for which the

---

subsample of observations with Direct Export Ratio > 0.01, our results suggest a marginally stronger effect of relabeling.

Direct Export Ratio is below the median, the evasion gap rises above zero (see Table 2(II)). Also note that the evasion gap is higher when measured in values rather than quantities; this is suggestive that some evasion takes the form of under-reporting of per unit values.

By examining some basic cross-tabulations in the data, we may informally get a sense of the extent of evasion through product misclassification and/or under-reporting. In Table 3(I), we consider the evasion gap for goods in higher-taxed 4-digit categories (i.e., those with tax rates above median at the 4-digit level) as compared to those in lower taxed categories (i.e., those with tax rates below the median). The median evasion gap is about 40% higher for the higher taxed products, suggesting that an increase in the tax rate is associated with a rise in evasion. In Table 3(II), we perform a different tabulation, comparing the evasion gaps at the minimum and maximum tax rates within each 4-digit classification. Column (1) shows that for the full sample, the evasion gap is marginally higher (by about 8%) for goods that have the highest tax rate within their respective 4-digit classes, relative to the gap for the lowest taxed goods within the same 4-digit category. The full sample includes, however, many products with very little dispersion within their 4-digit categories, and hence little difference between minimum and maximum tax rates. In Column (2), we limit the sample to those 4-digit categories for which the spread between the maximum and minimum tax rates at a 6-digit level within the 4-digit category exceeds 5 percentage points (the median spread). This subsample yields a much larger difference between the evasion gaps of the minima and maxima, of 0.40.

In Table 4, we cut the sample in the two dimensions previously described simultaneously: between minimum vs. maximum tax rates within 4-digit categories and between high vs. low tax rates across 4-digit categories. We focus on those 4-digit categories for which the difference between the maximum and minimum tax rates at the 6-digit level in the same 4-digit category exceeds 5 percentage points. The evasion gap is larger for higher taxed products. This is true when we look across industries at the 4 digit level (Column 3 in Table 4); it is also true *within* 4 digit industries when we compare the gaps of minima and maxima at the 6-digit level within a common 4-digit categories (Row C in Table 4). This suggests that the patterns in Tables 3(I) and 3(II) are both separately present in the data; to further confirm this, we proceed to regression analyses.

3. Empirical Analysis

Benchmark specification

We begin by defining the following: $Export_k$ = Hong Kong reported direct exports of good k to China (which we take as the true import of good k by China from Hong Kong). $Import_k$ = Direct imports of good k by China from Hong Kong as reported to Chinese Customs.

The prediction that we will be examining in the empirical test is that the difference between reported exports and imports is increasing in the tax rate, due to evasion. That is:

$$(1) \quad \log(Export_k) - \log(Import_k) = \alpha + \beta \, Tax_k + \varepsilon_k$$

Tax-induced evasion implies that $\beta>0$. Unfortunately, because China cannot always accurately separate indirect imports (e.g., U.S. exports to China passing through Hong Kong) from genuine direct imports from Hong Kong, we do not observe $Import_k$ directly. Instead, China's recorded imports from Hong Kong, $Import^*_k$, contains part of indirect imports misclassified as direct imports. In other words, we actually observe the following:

$$(2)\ Import^*_k = Import_k + Misclassified\ Indirect\ Import_k\ .$$

It is crucial to note that the same tax rate is applicable to both direct and indirect imports. Therefore, the magnitude of the misclassified indirect imports for a particular product should be uncorrelated with the tax rate for that product, since there is no tax advantage from misclassification. Rather, it is plausibly proportional to the magnitude of the import of that product (subject to some random error). We will assume that

$$(3) \qquad Misclassified\ Indirect\ Import_k = \theta\ \eta_k\ Import_k$$

Where θ is a constant, and η_k is an independent and identically distributed random variable, with distributional assumptions to be made later. Thus,

$$\begin{aligned}(4) \qquad Import^*_k &= Import_k + Misclassified\ Indirect\ Import_k \\ &= (1+ \theta\ \eta_k\ )\ Import_k\end{aligned}$$

Combining these four equations, we obtain

$$(5) \qquad log(Export_k) - log(Import^*_k) = \alpha^* + \beta\ Tax_k + e_k$$

or equivalently:

$$Gap\_Value_k = \alpha^* + \beta\, Tax_k + e_k$$

where α^* is a (new) constant, and e_k a composite error term that is assumed to be iid and normal with a mean of zero and a constant variance. To be more precise, if we denote the mean of $\varepsilon_k - \log(1+ \theta\, \eta_k)$ by α_0, then,

$$(6) \qquad \alpha^* \equiv \alpha - \alpha_0$$

and

$$(7) \qquad e_k \equiv \varepsilon_k - \log(1+ \theta\, \eta_k) - \alpha_0 \sim \text{Normal}(0, \sigma^2_e)$$

Equation (5) will be the benchmark for our regression specifications. The results are reported in Table 5. In the first column, we have the basic estimate of the sensitivity of evasion to tax rates, which is 2.93. This implies that if the tax rate increases by one percentage point, the gap between reported exports and imports increases by about three percent. When observations with the highest and lowest one percent of values of *Gap_Value* are excluded, the coefficient is virtually unchanged, as seen in the second column. In order to make direct comparisons with other results to be reported later in this paper, we also repeat this basic specification using the sub-sample of products with non-missing observations on taxes on similar products, and also limiting the sample to products with non-missing observations on quantities. These sub-samples may also be viewed as a sensitivity check for the main results. Columns 3 to 7 in Table 5 show that changing the sample in these ways has very little impact on the reported coefficient: One cannot reject the null hypothesis that the coefficient is 3 in all the cases in this Table.

Unfortunately, because of noise (possibly introduced by misclassified indirect imports), the fit of the regressions might be considered to be poor. A common method of dealing with noisy data is aggregation. We follow this approach, using as the outcome variable the mean value of *Gap_Value* for each tax rate. There are 42 distinct tax rates, thereby yielding a total of 42 observations. The basic relationship between the mean value of *Gap_Value* and tax rates is illustrated in Figure 3, and is suggestive of a positive correlation. The regression results weighted by number of observations per tax bracket, as well as the unweighted regression, are listed in Table 6, columns 1 and 2. The coefficients on *Tax*, 2.90 and 2.41, respectively, are similar to the baseline regressions from Table 5, significantly different from zero at the one percent level, but insignificantly different from 3.[10] The R^2 in the two regressions increase to 0.23 and 0.28, respectively.

Taking means as in Columns 1 and 2 of Table 6 is a linear operation, which allows for the same interpretation of the slope coefficients as in Table 5. As an alternative way to reduce the noise in the data, we also use the median of the evasion gap, *Gap_Value*, for each tax bracket, as the outcome variable. This approach has the advantage of further limiting the effect of outliers in the data, though the interpretation of the coefficient on *Tax* is not as straightforward. The regression results, reported in columns 3 and 4 of Table 6, are similar to those obtained from the mean regressions. Again, the point estimates on the tax rate are approximately 3, and the indicators of

---

[10] The observations at the far right of Figure 3 involving high tax rates could be outliers. When these three observations are removed from the sample, the unweighted regression yields a coefficient of 1.9, significant at the 5 percent level; the results for the weighted regression are virtually the same as those reported in Table 6.

goodness-of-fit or the adjusted R-squared in the two regressions increase to 0.30 and 0.29, respectively.

So far, we have concentrated our discussion on the statistical properties of the estimation. It is also useful to consider the economic implications of the point estimate in terms of revenue collection. From (1), we can infer how reported imports may be affected by an increase in the tax rate:

$$(8) \qquad \frac{(dImports/dTax)}{Imports} = \frac{(dExports/dTax)}{Exports} - \beta$$

Hence, the effect of a tax increase may reduce reported imports through two channels: by reducing the true imports (i.e., Hong Kong's exports, the first term in (8)); and by reducing the fraction of true imports that is reported to Chinese Customs (-β, the second term in equation (8)). While we do not have a direct estimate of the first term in Equation (8), it is reasonable to assume that it is negative. Therefore, an estimate of β equal to 3 percent implies that, for any product whose tax rate exceeds 33.3 percent, a one-percentage point increase in the tax rate would lead to a more than one-percentage point reduction in reported imports. The average tax rate on imports (tariff plus VAT) in China is 36 percent (see Table 2). These calculations suggest that the average tax rate is already on the wrong side of the Laffer curve: at the average rate, an increase in the tax rate will produce a reduction in tax revenue.

On a related point, we observe that we have assumed that tax rates are exogenously set by the government. However, if the government tries to protect tax revenue by setting tax rates systematically in inverse proportion to importers' ability to evade them, then the estimated responsiveness of tax evasion to tax rate reported here may underestimate the true degree of responsiveness.

Evasion by misclassification

In addition to underreporting the value of imports, evasion may take the form of misclassification – reporting a higher-taxed product as a lower-taxed variety. To investigate the existence of this type of evasion, we add the average tax rate of similar goods as a regressor. For a particular good k, its "similar products" are defined as all other products in the same 4-digit category. Define Avg(Tax_o) to be the average tax rate for product k's similar products, weighted by the value of their Hong Kong-reported exports. We implement a regression of the following form:

$$(9) \qquad Gap\_Value_k = \alpha + \beta_1 * Tax_k + \beta_2 * Avg(Tax\_o_k) + \upsilon_k$$

If the mislabeling of goods is prevalent, we expect $\beta_2 < 0$, so that holding a product's own tax rate constant, the lower the tax rate on product k's similar varieties, the greater the incentive to misreport the import of k as other similar products.

Table 7 reports results with *Avg(Tax_o)* included as a regressor. Consistent with the mislabeling interpretation, we find that the coefficient on *Avg(Tax_o)* is negative and significant at the 5 percent level, taking on values between -3 and -4.6. Furthermore, the inclusion of the average tax rate of similar goods as a regressor results in a substantial increase in the coefficient on *Tax*, which takes on values of 5.3 to 8.3.[11]

---

[11] Note that the difference in the magnitude of the coefficients on *Tax* and *Avg(Tax_o)* further suggests the presence of both mislabeling and underpricing. As an extension, we tried to examine the extent that underpricing increases as the potential for mislabeling declines. To implement this, we added the

We have also employed specifications that replace the average tax rate with the minimum tax rate among all "similar" products. A possible rationale is that the tax-evader may want to mislabel the import as the lowest-taxed "similar" product rather than any "similar" product. Empirically, this change in regressor has very little impact on the estimated coefficients. When both the minimum and average tariff rates are included in the regressions, the coefficient on the minimum tariff rate loses its significance, while the coefficient on average tariff rates is relatively unaffected (results not reported to save space).[12]

Quantity versus unit value

Thus far, we have not separated the under-reporting of unit values versus the under-reporting of quantities; both will result in a positive coefficient on *Tax* in the value regressions. We now turn to regressions that use the gap in imported quantity as the dependent variable. Specifically, we examine the following regressions:

$$(10) \quad Gap\_Qty_k = \alpha + \beta_1 * Tax_k + \upsilon_k$$

$$(11) \quad Gap\_Qty_k = \alpha + \beta_1 * Tax_k + \beta_2 * Avg(Tax\_o_k) + \upsilon_k$$

---

interaction of tariff rate levels with the standard deviation within 4-digit categories; this approach did not yield any statistically significant results.

[12] A further check on our mislabeling interpretation comes from the observation that, if a product is a relatively small fraction of total imports in a 4-digit class, we may expect a larger *proportional* effect from mislabeling. Define Proportion to be the ratio of the import of a particular product to total imports in that good's 4-digit class. Then (Tax Rate)*(Proportion) should take a negative coefficient, while the coefficient on (Taxes on Similar Products)*Proportion should be positive. We do find this to be the case (interaction terms significant at the 5 percent level). Results available from the authors.

If the under-reporting of quantities is prevalent, we expect to find a positive coefficient on Tax_k in the quantity regression, $\beta_1 > 0$. If there is mislabeling of imports from a higher-taxed category to a lower-taxed one, we expect to find $\beta_2 < 0$.

Results parallel to those of Table 5 and 7 using *Gap_Qty* as the dependent variable, are listed in Table 8. When *Avg(Tax_o)* is excluded from the regression, the coefficient on *Tax* is insignificantly different from zero. However, when *Avg(Tax_o)* is included, we find that the coefficient on *Tax* becomes significant, positive, and approximately equal to the coefficient on *Avg(Tax_o)* in absolute value. Thus, the data suggest that under-reporting of the total value of imports and mislabeling the type of goods are prevalent, while under-reporting of total quantities imported across all tax brackets is not significant.

Tariff exemptions

Many of Hong Kong's imports into China are exempt from import tariffs. These exemptions may impact incentives for evasion, and could be correlated with tariff rates, since high tax rates may increase incentives for exemption seeking.[13] More precisely, for products where exemptions are common, evasion may be less sensitive to tax rates than for products where exemptions are rare, since exemptions provide a legal means of avoiding tax payments.

We now examine whether taking into account exemptions has any material effect our main result. To do this, we first describe the calculation of the fraction of imports that is exempt from taxation for each of the 6-digit products. We obtained data at the 8-

---

[13] We have checked for this possibility, and found that there is a very low correlation between *Tax* and *Exemption*.

digit level on the exemption status of imports from the *Chinese Customs Statistics 1998* (Economic Information Agency, 2001). These data also include the value of China-reported imports from Hong Kong at the 8-digit level, which were used to calculate a weighted average of the proportion of imports exempt from tariffs for each 6-digit product, i.e.,

$$Exemption_{HS6} = \frac{\sum_{HS8 \in HS6} Import\_Value_{HS8} * Exempt_{HS8}}{\sum_{HS8 \in HS6} Import\_Value_{HS8}}$$

where Exempt is an indicator variable denoting whether a product is exempt from import tariffs, HS6 denotes products at the 6-digit level of aggregation, and HS8 denotes products at the 8-digit level of aggregation. Aggregating in this way is necessary in order to match to the Hong Kong-reported export data (which is at the 6-digit level).

Table 9 shows the results of the following specification, and its variations:

$$(12) \quad Gap\_Value_k = \alpha + \beta_1 * Tax_k + \beta_2 * Exemption_k + \beta_3 * Tax_k * Exemption_k + \beta_4 * Avg(Tax\_o_k) + \beta_5 * Avg(Tax\_o_k) * Exemption_k + \upsilon_k$$

Consistent with higher exemption rates lowering the incentives for evasion, the coefficient on *Exemption* is consistently negative and significant (See columns (1) to (4)). When interacted with *Tax*, the coefficient is negative, highly significant, and approximately the same size as the coefficient on *Tax*. This implies that for a product with complete exemption in 1998 (i.e., *Exemption*=1), there is no effect on evasion from tax increases. By contrast, for industries with no exemptions, the implied elasticity is about 16.[14] In looking at the effect of exemptions on incentives to re-label goods, we find that the coefficient on *Avg(Tax_o)*Exemption* is positive, though not significant at conventional levels (t-statistic of 1.5), implying less re-labeling for goods with high exemption levels. In columns (5) to (8), we report the results of the same regressions, using *Gap_Qty* as the outcome variable, and obtain similar results.

First differences in tax rates

Our primary results suggest a strong effect of taxes on evasion acting through both under-pricing and product mislabeling to lower-taxed categories. There may be concerns, however, that certain features of different products that are not directly measured and not included in the regressions may be driving the results. While we do not have specific factors in mind that may bias our result, as a further robustness check, use two years of data and adopt a first-difference specification that can net out time-invariant and product-specific determinants of tax evasion. We estimate the following:

$$(12) \quad DGap\_Value_k = \alpha + \beta_1 * DTax_k + \beta_2 * DAvg(Tax\_o_k) + \upsilon_k$$

where a prefix *D*- denotes the change between 1997 and 1998.

Since China underwent significant tariff reforms in 1997, largely driven by its goal of entry into the WTO, there is substantial variation in tariff rates across the two years. For the industries in our sample, tax rates dropped by an average of 5.6 percentage

---

[14] This figure is very large; if we omit outliers of Gap_Value, the implied elasticity of evasion with respect to taxes drops by about a third, while the significance of all coefficients in this regression are virtually unchanged. Also, there may be concerns that, because most of the exemption ratios are relatively high (the 25th percentile of *Exemption* is 0.68), making inferences about the effects of evasion at *Exemption* = 0 is too far out of sample. However, when we run regressions comparable to those reported in Table 7, limiting the sample to observations with *Exemption* < 0.5, we obtain a coefficient on *Tax* of 22. This suggests that functional form is not driving our results on the interaction of *Tax* and *Exemption*.

points. The dispersion in tariff rates within each 4-digit category also declined: the standard deviation of tariff rates within each 4-digit classification declined from an average of 1.9 to an average of 1.4. These simultaneous changes in both tariff rates and within-4-digit dispersion allows for the identification of the effects of changes in both *Tax* and *Avg(Tax_o)*.

Determining the appropriate tax rate for 1997 is not straightforward. While there was virtually no change in the tariff structure during 1998, a large-scale tariff reform occurred on October 1, 1997. Since our import data for 1997 are cumulated for the year, we use the weighted average of tax rates that prevailed before and after the tariff reform. However, for a potential importer, the knowledge of a tariff reform in the near future could affect the timing of the imports. We do not have any means of correcting for this, and assume that the effective tax rate for 1997 is given by:

$$\text{Tax}_{1997} = 0.75*(\text{Year-end tax rate for 1996}) + 0.25*(\text{Year-end tax rate for 1997})$$

We then define $DTax = Tax_{1998} - Tax_{1997}$.

Table 10 shows the estimation results. We note that the R-squared values are very low in these regressions; this suggests that we have differenced out much of the information in the data. In columns (1) and (2), with the change in *Gap_Value* as the dependent variable, the coefficient on *DTax* is significant at the 5 percent level, though marginally smaller than that obtained in our level regressions. The coefficient on *Davg(Tax_o)* is significant only at the 10 percent level, but is of the same sign as in the level regressions. We obtain similar results with the change in *Gap_Qty* as the dependent variable (see columns (3) and (4)).

Flexible functional form

We now allow the marginal effect of a tax increase on evasion to differ across different tax rates. Following, for example, Chamberlain (1997), we allow for the slope to differ across quartiles, with knots at tax rates of 29, 34, and 42 percent. The results, in Table 11, suggest that there is little effect on evasion at relatively low tax rates. However, as tax rates rise above the median level of 34 percent, the extent of evasion rises markedly. The marginal effect then tapers off at higher levels. As before, the effect of tax rate increases is larger when we control for average tax levels at the 4-digit level.

This pattern of non-linearity is consistent with the existence of a fixed cost in undertaking evasion activity. For example, if there is some fixity in the punishment for evasion, there may be a threshold tax level above which evasion becomes worthwhile.[15] Alternatively, it is also consistent with a probability of detection that is invariant to the tax rate, so that the benefit of evasion increases more rapidly than the cost as tax rates increase.

Robustness to Alternative Specifications of Indirect exports

As noted in the data section of this paper, the Chinese customs-reported import figures are likely to include part of the indirect imports that are misclassified as direct imports. We have argued that the amount of such misclassified imports should not be correlated with the tax rate on the product since the same tax rate is applied to both direct

---

[15] Unfortunately, the punishment code for customs evasion in China is sufficiently vague as to give little guidance on this question. While the punishment includes the confiscation of the goods involved, it also may incorporate a fine of an unspecified amount. Furthermore, in recent years, some importers and customs officials have been executed for tariff evasion, thereby highlighting the unpredictable nature of punishment.

and indirect imports. As a further precaution, we restricted our sample to those products where Hong Kong's direct exports to China are not trivial. Our choice of 0.01 as a cutoff value for *Direct Export Ratio = (Direct Export)/(Indirect Export + Direct Export)* is arbitrary, however. As a robustness check, we repeat the regressions of Tables 4 and 5, utilizing a range of different cutoff values. These regressions, reported in Table 12, show that our results are qualitatively unaffected by the choice of cutoff value. Similarly, Table 13 shows that the quantity results are also robust to differing cutoff values.[16]

4. Conclusions

In this paper, we take a new approach in measuring the effect of tax rates on tax evasion, by looking at the reporting gap in China's imports from Hong Kong, as a function of Chinese tax rates (tariff plus VAT rates). We find that this 'evasion gap' is highly correlated with tax rates: much more value is 'lost' for products with higher tax rates. The point estimates suggest that the Chinese average tax rate on imports is already on the wrong side of the Laffer curve: any increase in tax rate is likely to produce a reduction rather than an increase in tax revenue.

By comparing the evasion gap in quantities and in values, we conclude that there are widespread practices of under-reporting the unit values of imports, and mislabeling higher-taxed products as lower-taxed varieties.

As a broader contribution to the literature, we believe that our approach may also be applied to other countries. In addition to providing more information on the

---

[16] An alternative approach, since we are estimating a very reduced form of the evasion equation, is to include the value of indirect exports as a regressor. The results from this approach are virtually identical to those previously reported.

behavioral response of tax evasion to tax rates, the generalized multiple-country study could provide a more objective measure of the laxity of rule of law across countries – in contrast to the subjective perception based measures of corruption and rule of law now popular in empirical studies. We leave this, and other extensions, for future work.

Appendix: Variable Definitions

Import_Value: Value of imports from Hong Kong into China, as reported by the Chinese customs, for 1998, at the 6-digit HS level. Measured in U.S. dollars. Source: World Bank's World Integrated Trade Solutions (WITS), derived from the United Nations' Comtrade Database.

Export_Value: Value of exports by Hong Kong destined for China, as reported by Hong Kong customs, for 1998, at the 6-digit HS level. Measured in U.S. dollars. Source: WITS

Indirect Export_Value: Value of Indirect Exports from Hong Kong to China, originating from third-party countries. Reported by the Hong Kong customs, for 1998, at the 6-digit HS level. Measured in U.S. dollars. Source: WITS

Gap_Value = log(*Import_Value*) – log(*Export_Value*)

Import_Qty: Quantity of imports from Hong Kong into China, as reported by the Chinese customs, for 1998, at the 6-digit HS level. Measured in U.S. dollars. Source: WITS

Export_Qty: Quantity of exports from Hong Kong destined for China, as reported by Hong Kong customs, for 1998, at the 6-digit HS level. Measured in U.S. dollars. Source: WITS

Indirect Export_Quantity: Quantity of indirect exports, originating from third-party countries. Reported by the Hong Kong customs, for 1998, at the 6-digit HS level. Measured in U.S. dollars. Source: WITS

Gap_Qty = log(*Import_Qty*) – log(*Export_Qty*)

Tax: Total taxes rate levied on incoming goods by the Chinese Authorities, equal to the sum of tariffs and commercial taxes, for 1998. Source: WITS

Dtax: Difference in the tax rates in 1998 and 1997. See the text for details.

Exemption: At the 6-digit level, the proportion of goods exempt from import tariffs. Source: China Customs Statistics, purchased from the EIA CCS Information Service Center in Hong Kong.

Avg(Tax_o): Average of the level of *Tax* for all other goods in a product's 4-digit HS class, weighted by *Export_Value*.

Direct Export Ratio: Ratio of Direct Exports to Direct Exports + Indirect Exports

References

Allingham, Michael G. and Agnar Sandmo, "Income Tax Evasion: A Theoretical Analysis." *Journal of Public Economic*s 1 (1972): 323-338.

Andreoni, James, Brian Erard, and Jonathan Feinstein, "Tax Compliance." *Journal of Economic Literature* 36 (June, 1998): 818-860.

Bernard, Jean-Thomas, and Robert Weiner, "Multinational Corporations, Transfer Prices, and Taxes: Evidence from the U.S. Petroleum Industry," in *Taxation in the Global Economy* (A. Razin and J. Slemrod, eds.), Chicago: University of Chicago Press, 1990: 123-154.

Chamberlain, Gary, *Econometrics*, unpublished manuscript, 1997.

Clausing, Kimberly, "The Impact of Transfer Pricing on Intrafirm Trade." NBER Working Paper No.W6688, August 1998.

Economic Information Agency, *Chinese Customs Statistics 1998 (CCS)*, 2001.

Clotfelter, Charles T, "Tax Evasion and Tax Rates: An Analysis of Individual Returns." *Review of Economics and Statistic*s 65 (August, 1983): 363-373.

Feenstra, Robert, and Gordon Hanson, "Aggregation Bias in the Factor Content of Trade: Evidence from U.S. Manufacturing," *American Economic Review* 90 (May 2000):155-60.

Feige, Edgar L, "How Big is the Irregular Economy?" *Challeng*e 22, (November-December, 1979): 5-13.

Feinstein, Jonathan, "An Econometric Analysis of Income Tax Evasion and its Detection." *RAND Journal of Economic*s 22 (Spring, 1991):14-35.

Gutmann, Peter M., "The Subterranean Economy." *Financial Analysts Journal* 33 (November-December, 1977): 26-34.

Matthews, Guy, "Steel Smugglers Pull Wool Over the Eyes Of Customs Agents to Enter U.S.," *Wall Street Journal*, Nov. 1, 2001, p. A1.

Productivity Commission by the Department of Foreign Affairs and Trade, *Review of Australia's General Tariff Arrangements*, 2002.

Slemrod, Joel and Shlomo Yitzhaki, " Tax Avoidance, Evasion, and Administration," NBER Working Paper No. W7473, January 2000.

Tanzi, Vito, "The Underground Economy in the United States: Estimates and Implications," *Banco Nazionale del Lavoro Quarterly Revie*w, No. 135 (December, 1980): 427-53.

World Bank, *World Integrated Trade Solution (WITS)*, 2000.

Yang, Dean, “How Easily do Lawbreakers Adapt to Increased Enforcement? Philippine Smugglers' Responses to a Common Customs Reform,” mimeo, Harvard University, 2003.

Yitzhaki, Shlomo, "A Note on ‘Income Tax Evasion: A Theoretical Analysis’," *Journal of Public Economic*s 3 (May, 1974): 201-202.

Table 1 – Tariff Rates at the 6-digit level for machine tools (industry 8459)

HS-code	Industry Description	Tariff Rate (1998)
845910	Way-type unit head machines	35
845921	Other drilling machines :-- Numerically controlled	26.7
845929	Other drilling machines :-- Other	35
845931	Other boring-milling machines :-- Numerically controlled	26.7
845939	Other boring-milling machines :-- Other	37
845940	Other boring machines	31.9
845951	Milling machines, knee-type :-- Numerically controlled	26.7
845959	Milling machines, knee-type :-- Other	35
845961	Other milling machines :-- Numerically controlled	26.7
845969	Other milling machines :-- Other	37
845970	Other threading or tapping machines	35

Table 2(I) – Summary Statistics, Full Sample

	Mean	Std. Dev.	Min.	Max	Obs.
log(Value_Export)	5.35	2.41	-4.51	12.72	1663
log(Value_Import)	5.47	2.50	-5.12	12.20	1663
Gap_Value	-0.12	2.03	-7.95	9.64	1663
log(Qty_Export)	6.78	4.42	-2.30	20.17	1199
log(Qty_Import)	6.42	4.16	-2.30	21.99	1603
Gap_Qty	-0.63	2.33	-10.16	11.74	1102
Tax Rate (Tariff+VAT) (at the 6-digit level)	0.36	0.09	0.13	1.04	1663
Avg(Tax_o) (at the 4-digit level)	0.36	0.09	0.13	0.88	1470
Exemption Ratio	0.80	0.29	0.00	1.00	1558
Direct Export Ratio	0.20	0.24	0.01	0.99	1663
Change in Tax Rate, 1997 – 1998	-0.057	0.060	-0.30	0.09	1617

Table 2(II) – Summary Statistics, Restricted to Products with Direct Export Ratio above the Median

	Mean	Std. Dev.	Min.	Max	Obs.
log(Value_Export)	6.10	2.49	-4.51	12.72	839
log(Value_Import)	5.53	2.68	-5.12	12.20	839
Gap_Value	0.57	2.01	-7.92	9.64	839
log(Qty_Export)	8.18	4.47	-2.30	20.17	577
log(Qty_Import)	7.01	4.34	-2.30	20.32	797
Gap_Qty	0.01	2.35	-8.63	11.74	526
Tax Rate (Tariff+VAT) (at the 6-digit level)	0.37	0.10	0.13	1.04	839
Avg(Tax_o) (at the 4-digit level)	0.37	0.9	0.13	0.67	746
Exemption ratio	0.82	0.29	0.00	1.00	791
Change in Tax Rate, 1997 – 1998	0.063	0.060	-0.30	0.082	828

Table 3 – Median Evasion Gaps and Tax Rates
(across and within 4-digit classes)

(1) **Different Evasion Gaps at Different Tax Rates (across 4-digit categories)**

	(1) **Median Evasion Gap** Log(Export) – Log(Import)
A. **At Low Tax Rates** (Below Median, 4-digit level)	**-0.16** (356 products)
B. **At High Tax Rates** (Above Median, 4-digit level)	**0.23** (322 products)
C. Difference in Evasion = (B) – (A)	**0.39**

(II) **Different Evasion Gaps at Different Tax Tates (within 4-digit categories)**

	(1) **Median Evasion Gap** All 4-digit categories	(2) **Median Evasion Gap** 4-digit categories with Max(Tax) – Min(Tax) > 0.05
A. **At the Minimum Tax Rate** (within a 4-digit Product Line)	**-0.02** (293 products)	**0.14** (149 products)
B. **At the Maximum Tax Rate** (within a 4-digit Product Line)	**0.06** (385 products)	**0.54** (175 products)
C. Difference in Gap = (B) – (A)	**0.08**	**0.40**

Notes: In both tables, entries reflect the median evasion gaps, or *Gap_Value* = log(Reported Exports) – log(Reported Imports), with number of observations in parentheses. In part (I), Row (A) represents observations on 4-digit product lines where the tax rates are below the median; Row (B) represents observations on 4-digit product lines where the tax rates are above the median. Note that the number of observations between rows (A) and (B) in part (II) differ because of multiple minima and maxima within each 4-digit class.

Table 4 – Median Evasion Gaps, Cross-Tabulated by Different Tax Rates Both Within and Across 4-Digit Categories

(Limited to 4-digit categories where Maximum Tax - Minimum Tax > 0.05)

	(1) **Lower Tax Rates** (below median, at the 4-digit level)	(2) **Higher Tax Rates** (above median, at the 4-digit level)	(3) **Difference in Gap** = (2) – (1)
A. At the Minimum Tax Rate (within 4-digit categories)	-0.23 (57 products)	0.36 (92 products)	**0.59**
B. At the Maximum Tax Rate (within 4-digit categories)	0.25 (48 products)	0.65 (127 products)	**0.40**
C. Difference in Gap=(B)-(A)	**0.48**	**0.29**	

Notes: Entries reflect the median value of *Gap_Value* = log(Reported Exports) – log(Reported Imports), with number of observations in parentheses. Column (1) represents observations on 4-digit product lines where the tax rates are below the median; column (2) represents observations on 4-digit product lines where the tax rates are above the median. Note that the number of observations between rows (A) and (B) differ because of possible multiple minima and maxima within each 4-digit class.

Table 5: The Effect Of Tax Rates on Evasion (Measured in Value)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tax Rate	2.93	2.46	3.21	3.57	2.98	2.61	3.4
	(0.74)	(0.67)	(0.87)	(0.89)	(0.81)	(0.79)	(0.96)
Constant	-1.31	-1.04	-1.31	-1.48	-1.29	-1.12	-1.46
	(0.29)	(0.23)	(0.30)	(0.31)	(0.29)	(0.27)	(0.34)
Excluding Outliers?	No	Yes	No	No	Yes	Yes	Yes
Excluding products lacking tax on similar products?	No	No	Yes	No	No	Yes	Yes
Excluding products lacking Obs. on Quantities?	No	No	No	Yes	Yes	No	Yes
No of Observations	1663	1639	1470	1102	1087	1450	968
R^2	0.020	0.017	0.022	0.031	0.025	0.017	0.029

Note: Dependent Variable: log(Value of Exports from HK to China) – log(Value Imports to China from HK). Robust standard errors in parentheses, accounting for clustering of standard errors by 4-digit HSC.

Table 6: Aggregating the Evasion Gap by Tax Brackets

	(1)	(2)	(3)	(4)
Tax Rate	2.90	2.41	3.29	2.56
	(1.17)	(0.76)	(1.02)	(0.89)
Constant	-1.17	-0.97	-1.37	-1.101
	(0.37)	(0.28)	(0.33)	(0.32)
Method of Aggregation	Mean	Mean	Median	Median
Weighting of Data by Number of Observations per Tax Rate?	Yes	No	Yes	No
No of Observations	42	42	42	42
R^2	0.23	0.28	0.30	0.29

Note: For columns 1 and 2, the dependent Variable is Mean[log(Value of Exports from HK to China) – log(Value Imports to China from HK], with means taken for each level of tax rate. For columns 3 and 4, the dependent variable is Median[log(Value of Exports from HK to China) – log(Value Imports to China from HK]. Robust standard errors in parentheses.

Table 7: Incorporating the Average Tax on Similar Products

	(1)	(2)	(3)	(4)	(5)
Tax Rate		6.07	5.31	8.32	7.46
		(1.37)	(1.25)	(1.56)	(1.42)
Tax on Similar Products	2.62	-3.16	-2.98	-4.65	-4.45
	(0.90)	(1.39)	(1.33)	(1.58)	(1.53)
Constant	-1.09	-1.20	-1.02	-1.56	-1.33
	(0.034)	(0.31)	(0.28)	(0.38)	(0.35)
Excluding Outliers?	No	No	Yes	No	Yes
Excluding products lacking Obs. on Quantities?	No	No	No	Yes	Yes
No of Observations	1470	1470	1450	981	968
R^2	0.014	0.025	0.020	0.041	0.035

Note: Dependent Variable: log(Value of Exports from HK to China) – log(Value Imports to China from HK). Robust standard errors in parentheses, accounting for clustering of standard errors by 4-digit HSC.

Table 8: Evasion in Physical Quantities

	(1)	(2)	(3)	(4)	(5)	(6)
Tax Rate	1.13	1.14	0.69		8.37	8.12
	(0.93)	(1.11)	(0.83)		(2.20)	(1.77)
Tax on Similar Products				0.08	-7.93	-8.31
				(1.13)	(2.23)	(1.84)
Constant	-1.05	-1.08	-0.92	-0.68	-0.85	-0.65
	(0.33)	(0.40)	(0.30)	(0.40)	(0.40)	(0.36)
Excluding products lacking Obs. on Avg(Tax_o)?	No	Yes	No	Yes	Yes	Yes
Excluding Outliers?	No	No	Yes	No	No	Yes
No of Observations	1102	981	1082	981	981	962
R^2	0.003	0.002	0.001	0.000	0.015	0.019

Note: Dependent Variable: log(Quantity of Exports from HK to China) – log(Quantity of Imports to China from HK). Robust standard errors in parentheses, accounting for clustering of standard errors by 4-digit HSC.

Table 9: Controlling for the Effect of Tariff Exemptions

	Dependent Variable: Gap_Value				Dependent Variable: Gap_Qty			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tax Rate		2.96	5.30	16.69		1.05	7.07	18.73
		(0.74)	(1.35)	(4.94)		(0.89)	(2.22)	(6.19)
Tax on Similar Products			-2.20	-8.62			-6.61	-12.16
			(1.40)	(4.83)			(2.32)	(6.03)
Exemption	-1.06	-1.02	-1.10	-1.42	-1.41	-1.37	-1.47	1.73
	(0.23)	(0.23)	(0.25)	(0.94)	(0.31)	(0.31)	(0.33)	(1.16)
Exemption*(Tax Rate)				-15.34				-16.10
				(8.53)				(7.17)
Exemption*(Tax on Similar Products)				8.91				8.54
				(5.85)				(7.02)
Constant	0.72	-0.40	-0.41	-2.38	0.48	0.05	0.30	-2.32
	(0.23)	(0.32)	(0.38)	(0.69)	(0.31)	(0.40)	(0.46)	(0.97)
No of Observations	1558	1558	1362	1362	1028	1028	905	905
R^2	0.022	0.043	0.052	0.066	0.031	0.033	0.048	0.061

Note: Dependent Variable: Columns (1) – (4): log(Value of Exports from HK to China) – log(Value Imports to China from HK). Columns (5) – (8): log(Quantity of Exports from HK to China) – log(Quantity of Imports to China from HK). Robust standard errors in parentheses, accounting for clustering of standard errors by 4-digit HSC.

Table 10: Tax and Evasion in First Differences, 1997-98

	Dependent Variable: Change in Gap_Value		Dependent Variable: Change in Gap_Qty	
	(1)	(2)	(3)	(4)
Change in Tax Rate	1.71	5.60	1.88	5.78
	(0.85)	(1.92)	(1.48)	(2.91)
Change in Tax on Similar Products		-3.97		-4.72
		(2.30)		(3.55)
Constant	0.036	0.01	0.11	0.05
	(0.060)	(0.063)	(0.08)	(0.091)
No of Observations	1617	1430	1042	938
R^2	0.004	0.008	0.002	0.005

Note: Dependent Variable: Columns (1) and (2): Change in Gap_Value between 1997 and 1998. Columns (3) and (4): Change in Gap_Qty between 1997 and 1998. Robust standard errors in parentheses, accounting for clustering of standard errors by 4-digit HSC.

Table 11: The Effect of Tax Rates on Evasion – Flexible Functional Form

	(1)	(2)
Tax rate in 1st quartile	-2.72	-1.21
(0 ≤ Tax Rate< 29)	(3.03)	(3.47)
Tax rate in 2nd quartile	0.57	3.51
(29 ≤ Tax Rate< 34)	(3.89)	(4.28)
Tax rate in 3rd quartile	6.33	9.54
(34 ≤ Tax Rate< 42)	(2.88)	(3.43)
Tax rate in 4th quartile	3.12	6.51
(42 ≤ Tax Rate)	(1.52)	(2.49)
Avg Tax on Similar		-3.06
Products		(1.42)
Constant	0.37	0.77
	(0.81)	(0.83)
No of Observations	1663	1470
R^2	0.025	0.032

Note: Dependent Variable: log(Value of Exports from HK to China) – log(Value Imports to China from HK). Robust standard errors in parentheses, accounting for clustering of standard errors by 4-digit HSC.

Table 12: Differing Cutoffs for the Direct Export Ratio

(Evasion in Values)

	(1)	(2)	(3)	(5)	(6)	(7)
Tax Rate	2.82	3.10	4.03	6.32	5.27	5.11
	(0.78)	(0.75)	(0.83)	(1.77)	(1.65)	(1.88)
Avg Tax on Similar Products				-3.57	-2.16	-1.31
				(1.70)	(1.71)	(2.06)
Constant	-1.63	-0.86	-0.96	-1.58	-0.87	-0.86
	(0.27)	(0.26)	(0.30)	(0.35)	(0.30)	(0.36)
Cutoff for (Direct Exports/(Total Exports)	0.00	0.05	0.10	0.00	0.05	0.10
No of Observations	2043	1157	863	1760	1028	764
R^2	0.015	0.026	0.041	0.018	0.029	0.037

Note: Dependent Variable: log(Value of Exports from HK to China) – log(Value Imports to China from HK).
Robust standard errors in parentheses, accounting for clustering of standard errors by 4-digit HSC.

Table 13: Differing Cutoffs for the Direct Export Ratio

(Quantity Regressions)

	(1)	(2)	(3)	(5)	(6)	(7)
Tax Rate	0.92 (0.84)	1.58 (0.99)	2.21 (0.83)	7.59 (2.08)	9.54 (2.72)	8.14 (2.79)
Avg Tax on Similar Products				-7.40 (2.07)	-9.14 (2.78)	-7.17 (2.91)
Constant	-1.39 (0.30)	-0.76 (0.37)	-0.76 (0.43)	-1.17 (0.41)	-0.38 (0.44)	-0.35 (0.50)
Cutoff for Direct Export Ratio	0.00	0.05	0.10	0.00	0.05	0.10
No of Observations	1368	770	592	1191	692	534
R^2	0.002	0.006	0.011	0.011	0.024	0.017

Note: Dependent Variable: log(Quantity of Exports from HK to China) – log(Quantity of Imports to China from HK). Robust standard errors in parentheses, accounting for clustering of standard errors by 4-digit HSC.

Figure 1 - Frequency Distribution of Direct Export Ratio
(See data appendix for definition)

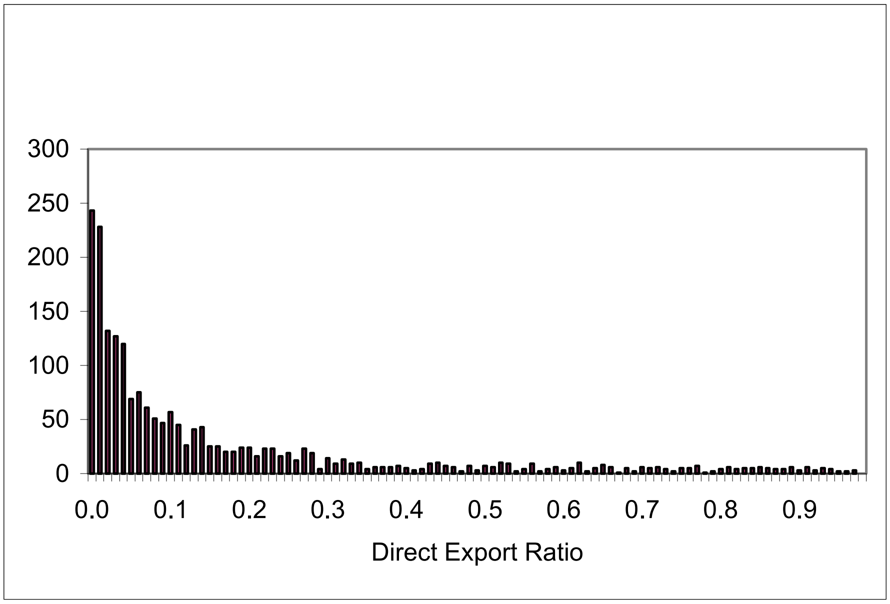

Figure 2(I) - Frequency distribution of tariff rates by 6-digit HS category

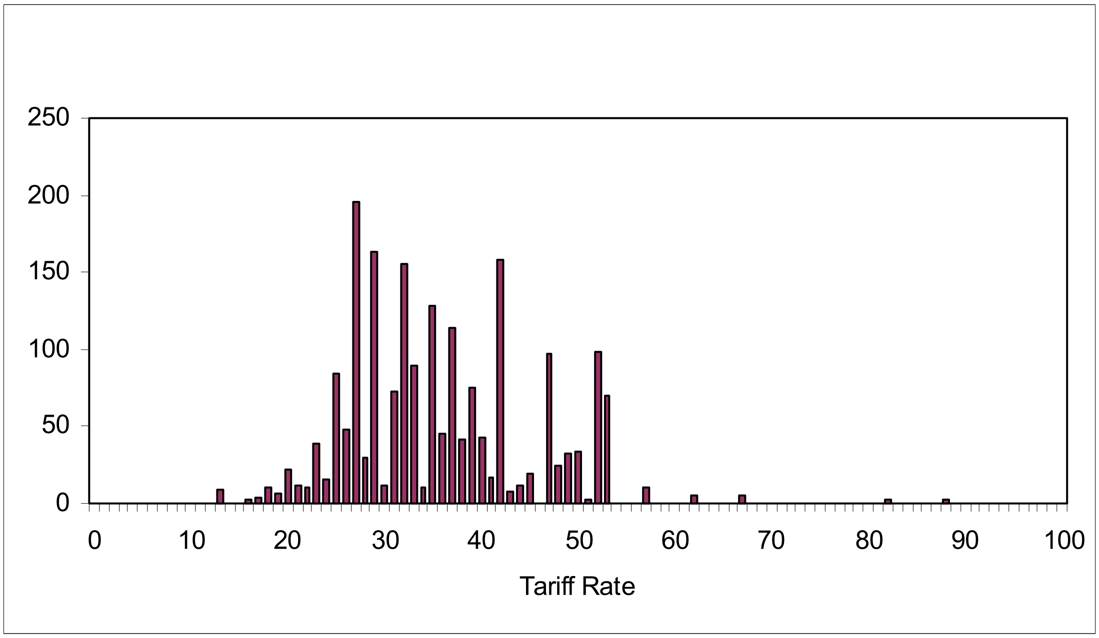

Figure 2(II) - Frequency distribution of differences between maxima and minima of tariffs within each 4-digit HS category

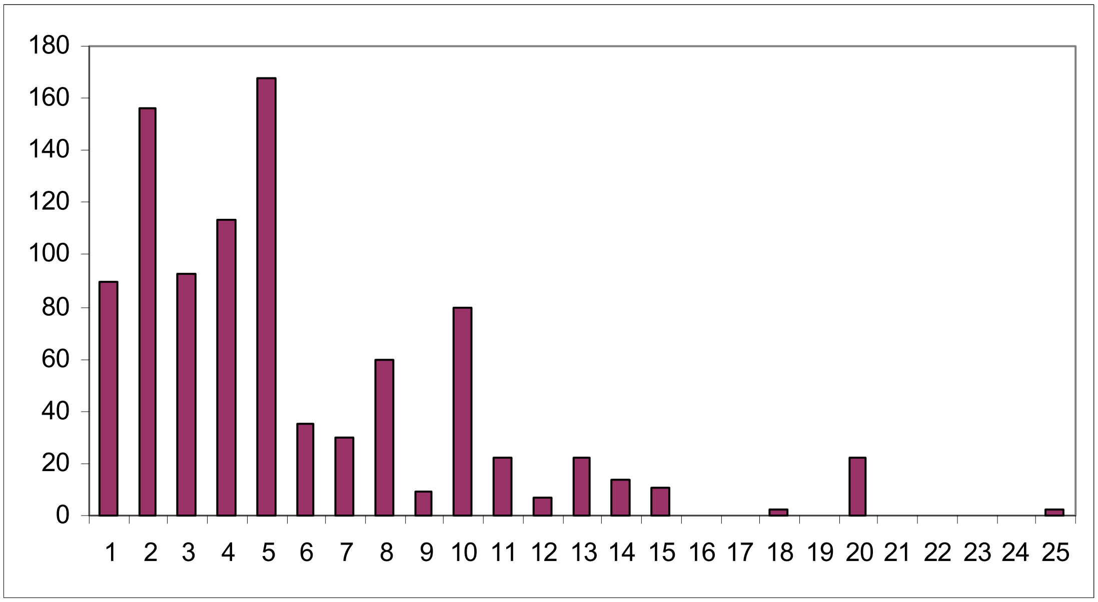

Figure 3 - Relationship Between Mean(Gap_Value) and Tax Rates, 1998