Determinants of effective tax rates for firms listed on China’s stock markets: panel models with two-sided censors

Yong-Ching Chiou, Yao-Chih Hsieh and Wenyi Lin
Feng Chia University, Taiwan

Key Words
Effective tax rate, two-sided censoring, political power (cost) hypothesis, JEL:C21, C52, H25

Abstract
This paper is to investigate the determinants of ETRs (Effective Tax Rate) for the firms listed on China’s (Shanghai and Shenzen) stock markets. The panel data consists of 360 firms from 2004 to 2011 as our empirical data. Since the dependent variable, ETRs, is left censored at 0 and right censored at 1, the estimation for panel data model with two-sided censoring suggested by Alan, Honoré, and Leth-Petersen (2008) is implemented in this paper. There are two important findings are obtained: first, this model can add more observations especially the observations with tax preferences. Second, theories suggest that ETR reflects outcomes of tax preference and this paper is the first time to consider the effective tax rates set between 0 and 1 and this range is more meaningful for the ETRs.

Introduction
Due to its convenience for measuring the tax burden of a corporate, average effective tax rates (ETRs) have long been used by policy makers and interest groups in tax reform debates. Governments provide tax incentives for firms subject to high levels of risk due to large amounts of capital, a lengthy production process or uncertainty in activities such as exploration (Stickney and McGee, 1982). Reduction of the corporate tax as a tool for stimulating a certain industry and elimination of certain tax preference as a tool for pursuing social justification are also the focus of the tax policy debates. Tax incentives reduce the tax burden of firms and imply non-neutrality of the tax system. Proponents of neutrality argue that the market would more effectively price the risk factors faced by some firms. Further, tax burdens varying across firms is sometimes used to suggest that the tax system is inequitable and subsequently.

Since ETR is defined as the ratio (in percentage) of taxes paid based on a firm’s current or total income to its pre-tax accounting income, ETR should be similar for all firms in a country. Therefore, evidence that corporate ETRs vary across firms and over time has been used to suggest that the tax system is inequitable, and as a justification for a corporate tax reform. In other words, study on the variability of ETRs among firms is a way to check tax neutrality and social justification in a country.

Variability of ETRs among firms has been recognized in literature. Originated from a serial reports published by the Citizens for Tax Justice (CTJ 1984, 1985, 1986) that the largest United States corporations were not paying their fair share of taxes. Whether the ETRs are systematically related to firm size has been the focus of the tax policy debates and literature studies. Theoretically, there are two different hypotheses suggested in literature to explain the relation between firm size and level of ETRs. One, the political cost hypothesis, is that the higher visibility of larger and more successful firms might be victims of greater regulator actions and wealth transfers and as a result, incur a political cost in the form of a higher ETR (Zimmerman 1983). Another is that larger firms pay less tax because they devote more efforts to tax planning and political lobbying so that the large firms pay their share of tax less. This is referred to as the political power (or clout) hypothesis.

Empirically, studies on the relation between ETRs and firm size have produced conflicting results. Zimmerman (1983) observes a positive association between ETRs and firm size while Porcano (1986) observes a negative association. No association between ETRs and firm size is found in Stickney and McGee (1982) and Shevlin and Porter (1992). Subsequent studies have tried to reconcile the conflicting results by using modified proxies, time period, data basis, and methodologies (e.g., Gupta and
Newberry, 1997; Kern and Morris, 1992; Wilkie and Limberg, 1990; Holland, 1998; Kim and Limpaphayom, 1998; Derasid and Zhang, 2003; Liu and Cao, 2007). Panel data models have been used to overcome the problem of model misspecification in studies on determinants of ETRs (e.g. Gupta and Newberry, 1997; Harris and Feeny, 2003; Liu and Cao, 2007). Besides, Gupta and Newberry (1997) suggest that because ETRs concern firms with tax refunds, negative income and measurement issue, they constrained the ETR of their sample firms to lie between zero and one. However, the censoring characteristic of ETRs is not considered in existing literature. In our sample, there are more than 28% of firms with zero ETRs. It is well known that the regression estimators are biased and inconsistent if the existence of censoring in dependent variable is neglected. Therefore, this paper attempts to use the panel data models with two-sided censoring suggested by Alan, Honoré, and Leth-Petersen (2008) to study the determinants of ETRs for the listed on China stock markets.

The remaining of this paper is organized as follows. Section 2 introduces the estimation of panel data models with two-sided censoring suggested by Honoré and Leth-Petersen (2008). Empirical studies are investigated in section 3. Conclusions are presented in the last section.

Panel Data Models with Two-sided Censoring

Since the effective tax rates are between 0 and 1 with a significant number of observations on either of the limits. In panel data setting, the specific model is

\[ y_{it} = x_{it}'\beta + e_{it}, \]

where \( e_{it} \) is stationary conditional on \((x_{it}, \ldots, x_{Tt})\). The derivation of estimator for \( \beta \) suggested by Alan, Honoré, and Leth-Petersen (2008) is briefly summarized as follows. Define, for \( a \leq b \),

\[ m_{a;b}(y) = \begin{cases} a & \text{if } y < a \\ y & \text{if } a \leq y \leq b \\ b & \text{if } y > b \end{cases} \]

so (1) can be written as

\[ y_{it} = m_{a;b}(x_{it}'\beta + e_{it}) \]

Consider an individual, \( i \), in two time periods, \( t \) and \( s \). The distribution of \((y_{it} - x_{it}'\beta)\) will be the same as that of \( e_{it} \), except that the former is censored from below at \( a - x_{it}'\beta \) and from above at \( b - x_{it}'\beta \). The dotted line depicts the distribution of \( e_{it} \), while the solid line gives the distribution of \( y_{it} - x_{it}'\beta \) (illustrated by the fatter vertical lines). Since \( x_{it}'\beta \) will typically differ from \( x_{st}'\beta \), the distributions of \( y_{it} - x_{it}'\beta \) and \( y_{st} - x_{st}'\beta \) (given \( x_{it} \)) will differ even if \( \{e_{it}\} \) is stationary. However, it is clear that one could obtain identically distributed “residuals” by artificially censoring \( y_{it} - x_{it}'\beta \) and \( y_{st} - x_{st}'\beta \) from below at \( \max\{a - x_{it}'\beta, \ldots, a - x_{s1}'\beta\} \) and from above at \( \min\{b - x_{it}'\beta, \ldots, b - x_{s1}'\beta\} \). One can then form moment conditions from the fact that the difference in these “re-censored” residuals will be orthogonal to functions of \((X_{it} - X_{st})\).

Also define functions \( u_1(y_{it}) \) and \( u_2(y_{st}) \) over the interval \(-(b-a)\) to \((b-a)\) as follows

\[
\begin{align*}
   u_1(y_{it}) &= \begin{cases} 
   \max[y_{it} - d, 0] & \text{for } b - a \geq d \geq 0 \\
   \min[y_{it}, b + d] & \text{for } 0 \geq d \geq -(b-a) 
   \end{cases} \\
   u_2(y_{st}) &= \begin{cases} 
   \max[y_{st} - d, 0] & \text{for } b - a \geq d \geq 0 \\
   \min[y_{st}, b + d] & \text{for } 0 \geq d \geq -(b-a) 
   \end{cases}
\end{align*}
\]
With these definitions, \( u_2(y_{it}, x_{it}^\theta - x_{it}^\beta) - u_1(y_{it}, x_{it}^\theta - x_{it}^\beta) \) will give the difference in the re-censored residuals by artificially censoring \( y_{it} - x_{it}^\theta \) and \( y_{it} - x_{it}^\beta \) from below as \( \max_{y_{it}} \{ u_1 - x_{it}^\beta(y_{it} - x_{it}^\beta) \} \) and from above at \( \min \{ y_{it} - x_{it}^\theta, b - x_{it}^\theta \} \).

Let the functions \( r_1(y_{it}, d) \) and \( r_2(y_{it}, d) \) are defined over the interval \(-\{b-a\}\) to \(\{b-a\}\) as

\[
r_1(y_{it}, d) = \begin{cases} 
  db + \frac{1}{2}d^2 + \frac{1}{2}(y_{it} - b)^2 & \text{for } d \leq y_{it} - b \\
  dy_{it} & \text{for } y_{it} - b \leq d \leq a \\
  dy_{it} - \frac{1}{2}d^2 & \text{for } a \leq d \leq y_{it} - a \\
  da + \frac{1}{2}(y_{it} - a)^2 & \text{for } y_{it} - a \leq d 
\end{cases}
\]

and

\[
r_2(y_{it}, d) = \begin{cases} 
  da + \frac{1}{2}d^2 + \frac{1}{2}(y_{it} - a)^2 & \text{for } d \leq -(y_{it} - a) \\
  \frac{1}{2}d^2 + dy_{it} & \text{for } -(y_{it} - a) \leq d \leq a \\
  dy_{it} & \text{for } 0 \leq d \leq b - y_{it} \\
  db - \frac{1}{2}d^2 - \frac{1}{2}(y_{it} - b)^2 & \text{for } b - y_{it} \leq d 
\end{cases}
\]

The functions \( r_1(y_{it}, d) \) and \( r_2(y_{it}, d) \) are constructed so their derivatives are \( u_1(y_{it}, d) \) and \( u_2(y_{it}, d) \), respectively. Finally, define

\[
R(y_{it}, y_{it}, d) = r_1(y_{it}, d) - r_2(y_{it}, d).
\]

Alan, Honor´e, and Leth-Petersen (2008) show that

\[
\frac{\partial}{\partial d} R(y_{it}, y_{it}, d) = r_1(y_{it}, d) - r_2(y_{it}, d).
\]

Suppose that

\[ y_{it} = \max[a, d + \varepsilon_{it}, b] \]

and

\[ y_{it} = \max[a, e_{it}, b] \]

Where \( \varepsilon_{it} \) and \( e_{it} \) are identically distributed random variables with support on the whole real line. Then Alan, Honor´e and Leth-Petersen (2008) prove that

\[
\arg\min_{-\{b-a\}} E[R(y_{it}, y_{it}, d)] = \begin{cases} 
  -(b-a) & \text{if } \delta \leq -(b-a) \\
  \delta & \text{if } -(b-a) < \delta < (b-a) \\
  (b-a) & \text{if } \delta \geq (b-a)
\end{cases}
\]

If \( \varepsilon_{it} \) is stationary conditional on \( x_{it}^\theta \) with support on the whole real line, then the set of solutions to

\[
\max_b E[R(y_{it}, y_{it}, d) \max\{-(b-a), (y_{it} - y_{it})^\beta, (b-a)\}]
\]

is

\[ \{b : \max\{-(b-a), (y_{it} - y_{it})^\beta, (b-a)\} = 1\} \]

Therefore, when the censoring points are a and b, the sample analog estimator is

\[
\hat{\beta}_{it} = \arg\max_b \mathbb{E}[\varepsilon_{it} \varepsilon_{it-1} x_{it}^\theta y_{it} \varepsilon_{it}^{\max\{-(b-a), (y_{it} - y_{it})^\beta, (b-a)\}}]
\]

where the \( w_{it} \) are exogenous weights and \( T_i \) is the number of observations for the ith individual. \( w_{it} = 1/\hat{r}_{it} \) is a trivial choice.

It is proved by Alan, Honor´e and Leth-Petersen (2008) that \( \hat{\beta}_{it} \) is consistent and asymptotically normal under appropriately conditions. Under random sampling

\[
\sqrt{n}(\hat{\beta}_{it} - \beta) \to_d N(0, \Gamma^{-1}V + \Gamma^{-1})
\]

Where

\[
\Gamma = \mathbb{E}[\varepsilon_{it} \varepsilon_{it-1} \varepsilon_{it}^{\max\{-(b-a), (y_{it} - y_{it})^\beta, (b-a)\}}]
\]

\[
(1 - (b-a) < (y_{it} - y_{it})^\beta < (b-a))
\]

\[
(1 - (b-a) < (y_{it} - y_{it})^\beta < y_{it} - b - 1) \quad \text{and} \quad 1 - (b-a) < (y_{it} - y_{it})^\beta < y_{it} - a)
\]
Empirical Studies

The sample data used in this study is collected from the Taiwan Economic Journal (TEJ) database. It consists of 360 firms each year listed on China’s stock market from 2004 to 2011. The ETR is measured as four different ETRs measures are used. We follow the approach used by Porcano (1986), ETR1 is defined as (tax expenses - deferred tax expenses) divided by (profit before interest and tax paid) and ETR2 is another version of the measure used by Porcano (1986): (tax expenses)/(profit before interest and tax). ETR 3 is the measure used by Stickney and McGee (1982) and is given as (tax expenses)/(pre-tax profit - (deferred tax expenses/statutory tax rate)). ETR4 is the measure used by Shevlin (1987) and is calculated as (tax expenses - deferred tax expenses)/(pre-tax profit - (changes in deferred tax/statutory tax rate)).

To explore the marginal effect of firm size on ETRs, the following firm-specific characteristics are taken as control variables: leverage (total liabilities divided by total asset value, denoted as “LEV”), capital intensity (net fixed assets divided by total assets, denoted as “CI”), inventory intensity (inventory divided by total assets, denoted as “II”), and return on assets (pre-tax profits divided by total assets, denoted as “ROA”). Firm size (denoted as “SIZE”) is measured as the natural logarithm of total asset value. The state ownership variable is defined as the ratio of state-owned shares over total outstanding shares and denoted as S1. The reasons behind to choose these variables are based on previous studies (e.g., Porcano (1986), Gupta and Newberry (1997), Derashid and Zhang (2003), Liu and Cao (2007)). To account for the factors discussed above, The empirical results are presented in table 1.
Table 1 Estimated Results for Panel Data Model with Two-side Censoring

<table>
<thead>
<tr>
<th></th>
<th>ETR1</th>
<th>ETR2</th>
<th>ETR3</th>
<th>ETR4</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTERCEPT</td>
<td>-0.016</td>
<td>0.043</td>
<td>0.050</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.039)</td>
<td>(0.064)</td>
<td>(0.056)</td>
</tr>
<tr>
<td>S1</td>
<td>0.008</td>
<td>0.021*</td>
<td>0.008</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.012)</td>
<td>(0.022)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>CI</td>
<td>-0.031</td>
<td>0.002</td>
<td>0.017</td>
<td>-0.025</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.019)</td>
<td>(0.033)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>II</td>
<td>0.054**</td>
<td>0.070***</td>
<td>0.121***</td>
<td>0.096***</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.025)</td>
<td>(0.042)</td>
<td>(0.035)</td>
</tr>
<tr>
<td>LEV</td>
<td>-0.080**</td>
<td>-0.026</td>
<td>0.100*</td>
<td>0.060</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.032)</td>
<td>(0.055)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>ROA</td>
<td>0.213***</td>
<td>0.276***</td>
<td>0.207***</td>
<td>0.162***</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.034)</td>
<td>(0.055)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>SIZE</td>
<td>0.011***</td>
<td>0.004</td>
<td>0.004</td>
<td>0.008**</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
</tbody>
</table>

The coefficient for S1, government equity, does not appear to be statistically different from zero, except ETR2. This result suggests that the effects of government ownership may lie in areas other than ETR. The coefficient of ln assets, firm size, is positive and statistically significant when ETR is measured as ETR1 and ETR4. Thus, we can accept the political cost hypothesis when TER is defined as ETR1 and ETR4. The coefficient for leverage is negative and statistically significant when ETR is measured as ETR1 and ETR3. This coefficient is not statistically different from zero under ETR2 and ETR4. Thus, there is some evidence to support the intuitive notion that debt financing can be used as a tax shield for China’s firms. The coefficient for capital intensity is all statistically insignificantly for all ETR measures. This evidence does not support the notion that higher capital investment and the resultant higher depreciable costs lead to a lower ETR. This evidence is inconsistent with two previous studies in the U.S. context (Gupta and Newberry, 1997; Stickney and McGee, 1982). The coefficient for inventory intensity, on the other hand, is all statistically different from zero under all ETR measures. The coefficients of ROA, the measure for efficiency in performance is all statistically significant when ETR is measured as ETR1, ETR2, ETR3, ETR4. These results suggest that more efficient firms pay more effective tax in China. And these results are consistent to previous studies.

Conclusions

This paper is to investigate the determinants of ETRs for the firms listed on China’s stock markets. This paper adopts two-sided censoring model to China’s stock market from 1997 to 2006 as our empirical data. Since the dependent variable, ETRs, is left censored at zero and right censored at 1, the estimation for panel data model with two-sided censoring suggested by Honor’e and Leth-Petersen (2008) is implemented in this paper. Several contributions are obtained: first, this model can add more observations especially the observations with tax preferences. Second, at our best knowledge, this paper is the first time to consider the effective tax rates set between 0 and 1 and this range is more meaningful for the ETRs.

References

Alan, S., and S. Leth-Petersen (2006), Tax incentives and Household Portfolios: A Panel Data Analysis, working paper, Center for Applied Microeconometrics, University of Copenhagen.

Alan, S., Bo E. Honor’e, and S. Leth-Petersen (2008), Estimation of Panel Data Models with Two-Sided Censoring, working paper, Department of Economics, Princeton University.


---

The Business & Management Review, Vol.3 Number 1, November 2012 312