Global Journal of Management and Business Research: C
Finance
Volume 15 Issue 5 Version 1.0 Year 2015
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 2249-4588 & Print ISSN: 0975-5853

Tax Perception and Sample Selection Bias: Microeconometrics

By Amaresh Das & Adnan Omar
Southern University, United States

Abstract- This paper econometrically compares the perceived marginal tax rates and the actually computed marginal tax rates and tries to find out if consumers could accurately perceive the marginal tax rates. Econometrically, the paper highlights that sample selectivity operates through unobservable elements and their correlation with unobservables influencing the variable of primary interest. Sample selection bias will not arise purely because of difference in observable characteristics. Although our paper is illustrative, it highlights the generality of the issue and its relevance to many economic examples.

Keywords: fiscal illusion, probit model, sample selection bias, censored model, mills ratio.

GJMBR - C Classification : JELCode : H29

TAXPERCEPTIONANDSAMPLESELECTIONBIASMICROECONOMETRICS

Strictly as per the compliance and regulations of:

Tax Perception and Sample Selection Bias: Microeconometrics

Amaresh Das [α] & Adnan Omar [σ]

Abstract- This paper econometrically compares the perceived marginal tax rates and the actually computed marginal tax rates and tries to find out if consumers could accurately perceive the marginal tax rates. Econometrically, the paper highlights that sample selectivity operates through unobservable elements and their correlation with unobservables influencing the variable of primary interest. Sample selection bias will not arise purely because of difference in observable characteristics. Although our paper is illustrative, it highlights the generality of the issue and its relevance to many economic examples.

Keywords: fiscal illusion, probit model, sample selection bias, censored model, mills ratio.

I. Introduction

The question is: Do the majority of individuals make rational tax decisions based on the actual tax burden, but rather use simple decision heuristics? This leads to the importance of the tax rate being significantly overestimated and the importance of the tax base being significantly underestimated. There is a standing literature on the perception (bias) of individuals with respect to their own tax burden and its effect on economic decisions. The strands of literature being currently discussed are: perception of marginal tax rates, influence of tax complexity on tax perception, taxation and incentives to work, tax salience, tax morale and fairness and money illusion, perceived inflation and fiscal drag.

There is more evidence for than against a perception bias in the literature. .We will compare in our work the perceived marginal tax rates and the actually computed marginal tax rates and try to find if, consistent with prior studies, consumers accurately perceive the marginal tax rates they face. Statistical analyses based on non-randomly selected samples can lead to erroneous conclusions and, also, poor policy prescriptions. Heckman who in 2000 received the Economics Nobel Prize for this achievement while working at the University of Chicago, proposed a two-stage estimation procedure using the inverse Mills ratio to take into account of the selection bias. In the first step, a regression for observing a positive outcome of the dependent variable is modeled with a Probit model. If the inverse Mills ratio is generated from the estimation of a Probit model, a logit can not be used. The Probit model assumes that the error term follows a standard normal distribution. The estimated parameters will be used to calculate the inverse Mills ratio, which will then be included as an additional explanatory variable in the OLS estimation.

Our direct evidence on perceived and computed marginal federal income tax rates from our sample of Louisiana households may provide support for what in macroeconomics literature is called a' fiscal illusion' as a determinant of market behavior. This is a concept that governments find it easy to raise revenues because of consumer ignorance about the way the tax system operates[1]. The point is: if government revenues or taxes are not fully perceived by taxpayers, then the cost of government is seen to be less expensive than it actually is and in that case, the public appetite for government expenditures will increase providing politicians' incentive to expand the size of the government. Fiscal illusionists encourages tax increases (especially during times of budget deficit) because they force the public to meet excessive spending without making them feel the cost. This study will not incorporate the notion of fiscal illusion to include imperfect information where voters are unsure about how much they must pay for additional services or where they are unsure about the services received in return for higher taxes. This paper does not also incorporate other forms in which fiscal illusion may appear, for example, complexity of tax structure, recent illusion with respect to property taxation[2], income elasticity of the tax structure, debt illusion, and what is known as the 'fly paper effect'. For evaluation of the work on each of them, see Payne [12], Romer [13] and Turnbull [15] specifically for flypaper effects[3].

Author α σ: Professors, College of Business, Southern University at New Orleans, New Orleans. e-mail: adas2@cox.net

[1] Anthony Downs [1] as far back in 1970 argued convincingly that the representative voter is likely to have highly imperfect information on which to base his decisions on public sector activities. Imperfect information is not however, synonymous with fiscal illusion for its existence. Fiscal illusion refers to a systematic misperceptions of fiscal parameters – a recurring propensity, for example, to underestimate one's tax liability associated with certain public programs. Imperfect information alone might give rise to a random pattern of over-and underestimation of such tax liabilities. As such, it will give rise to recurring and presumably predictable, biases to budgetary decisions.

[2] Buchanan and Wagner [2], suggest that the complicated nature of the U.S. tax system causes fiscal illusion and results in greater public expenditure than would be the case in an idealized system in which everyone is aware in detail of what their share of the costs of government is. See, also, Breden and William [1],

[3] Chetty et al [3] demonstrate that tax salience has economically significant behavioral implications, which indicates that tax visibility matters both for consumer choice and for public policy.

II. Methodology and Data

In investigating a bias that arises from using an incomplete sample to estimate β_1, we must know why the data are missing. All of the models in the literature developed for limited dependent variables and sample selection bias may be interpreted within a missing data framework. Suppose that we seek to estimate a regression equation but for some observations from a large random sample data are missing on Y_1 in $Y_{1i} = X_{1i}\,\beta_i + U_{1i}$ In the case of a censored sample, we have access to the larger random sample but we do not know Y_1 for censored observation. In a truncated sample, we do not have access to any observations from the larger random sample except those for which data on Y_1 is available. In both cases, there is a sample of I_1 complete observations.

The population regression function for equation may be written as:

$$E(Y_{1i} \mid X_{1i}) = X_{1i}\beta_1 \quad i = 1, 2, \quad N_{1i}$$

which would be estimable from a random sample. The regression function for an incomplete sample may be written as

$$E(Y \mid X_{1i}, \text{sample selection rule})\; X_{1i}\,\beta_1 + E(U_{1i} \mid \text{sample selection rule}) \qquad (1)$$

$$i = 1, \qquad N_{1i}$$

where without loss of generality the first *N* observations are assumed to continue data on Y_1. If the conditional expectation of U_{1i} is zero, regressions fit on the subsample yield unbiased estimator of β_1.

In general it is not the case that selection into the subsample is random. For example, in Tobin's celebrated paper (Tobin [14]. Heckman [8]) on limited dependent variables, we observe Y_1 if and only if

$$Y_{1i} \geq C$$

Where *C* is constant.

Y_{1i} may be interpreted as an index of a tax payer's intensity of desire. If the intensity is sufficiently great ($Y_{1i} > C$) the tax payer expresses his desire and Y_{1i} is observed. Otherwise we cannot observe intensity and observed payment of taxes are zero. In Tobin's model the sample selection rule is given by

$$E(Y_1 \mid X_{1i}, Y_{1i} \geq 0) = X_{1i}\,\beta_1 + E(U_1 \mid Y_{1i} \geq 0)$$

We consider *a la* Tobin the following decision rule: we obtain data on Y_{1i}, if another random variable creates a threshold, i. e., if

$$Y_{2i} \geq 0$$

while if the opposite inequality holds we do not obtain data on Y_{1i}. The choice of zero as a threshold is an inessential normalization. Also, note that we could define a dummy variable $d_i = 1$ with the properties

$$d_i = 1 \;\; if\, Y_{2i} \geq 0 \quad d_i = 0 \;\; \text{otherwise.}$$

proceed to analyze the joint distribution of Y_{1i} and d_i dispensing with Y_{2i} altogether. The advantage of using selection rule representation is that it permits a unified summary of the existing literature[4].

Using this representation we may write equation (1) as[5]

$$E(Y_{1i} \mid X_{1i}\, Y_2 i \geq 0) = X_{1i}\,\beta_i + E(U_{2i} \mid U_{2i} = X_2\,\beta_2) \quad (2)$$

If U_{1i} is independent of U_{2i} the conditional mean of U_{1i} is zero. and the sample selection process into the incomplete sample is random. In the general case, the conditional mean of the disturbance in the incomplete sample is a function of X_{2i}. Moreover, the effect of such sample selection is that X_2 variables that do not belong in the population regression function appear to be statistically significant in equations fit on selected samples. To exploit the information that we observe Y_{2i} up to a positive factor of proportionality if Y_{2i} is positive.[6]

[4] A good example of this phenomenon is found in Lewis [10]. In his analysis, Y_{1i} is the wage rate which is only observable for working women, had Y_{2i} is an index of labor force attachment (which in the absence of fixed costs of work may be interpreted as the difference between market wages and reservation wages). If the presence of children affects the work decision but does not affect market wages, regression evidence from selected sample of working women that women with children earn lower wages is not necessarily evidence that there is market discrimination against such women or that women with lower market experience – as by children – earn lower wages. Moreover, regression evidence that such extraneous variables 'explain' wage rates may be interpreted as evidence that selection bias is present.

[5] If U_{1i} is independent of U_{2i} the conditional mean of $U_1 i$ is zero and the sample selection process into the incomplete sample is random. In the general case, the conditional mean of the disturbance in the incomplete sample is a function of X_{2i}. Moreover the effect of such sample selection is that X_2 variables that do not belong in the population regression function appear to be statistically significant in equations fit on selected samples.

[6] A crucial distinction between a truncated sample and a censored sample. In a truncated sample one cannot use the available data to define the probability that an observation has complete data. In a censored sample, one can.

$$E(h_i \mid X_{2i}, Y_{2i} \geq 0) = E\left[\frac{Y_{2i}}{\gamma} \mid X_2\right], Y_{2i} \geq 0 \tag{3}$$

Suppose that $h(U_{1i}, U_{2i})$, the joint density of U_{1i} and U_{2i} is bivariate normal. Using well results of the literature (Jonson and Kotz[9]).

$$E(U_{1i} \mid Y_{2i} \geq 0) = E(U_{1i} \mid U_{2i} >- X_{2i}\beta_2) = \frac{\sigma_{12}}{(\sigma_{22})^{1/2}} \lambda_i$$

$$E(U_{2i} \mid Y_{2i} > 0) = E(U_{2i} \mid U_{2i} > -X_{2i}\beta_2) = \frac{\sigma_{22}}{(\sigma_{22})^{1/2}} \lambda_i$$

where

$$\lambda_i = \frac{f(\phi_i)}{1 - F(\phi_i)}$$

$$\frac{X_2 \beta_2}{(\sigma_{22})^{1/2}} = \phi_i$$

And *f* and *F* are respectively, are the density and distribution function of the standard normal distribution. The Tobin model is special case with $h(U_{1i}, U_{2i})$ a singular density sine $U_{1i} = U_{2i}$ λ_i is the inverse of Mill's ratio [7] and is known as the hazard rate in reliability theory. There are several interesting properties of λ_i.

- Its denominator is the probability that observation *i* has data for Y_i
- The lower the probability that an observation has data on Y, the greater the value of λ for that observation.

Moreover, using a result in Feller [6]

$$\frac{\delta \lambda_i}{\delta \phi_i} > 0$$

$$E[x > \alpha] \quad \mu + o \frac{\Phi(\frac{\alpha - \mu}{o})}{1 - \phi(\frac{\alpha - \mu}{o})}$$

$$E[x < \alpha] = \mu + o \frac{-\Phi(\frac{\alpha - \mu}{o})}{\phi(\frac{\alpha - \mu}{o})}$$

Where α is a constant, ϕ denotes the standard normal density function, and Φ is the standard normal cumulative distribution function. The two fractions are the inverse Mills ratios.

$$\lim_{\phi_i \to \infty} \lambda_i = \infty \quad \lim_{\phi_i \to \infty} \lambda_i = 0$$

Thus in samples the selectivity problem is unimportant, λ_i becomes negligibly small so that least squares estimates of the coefficients have optimal properties.

a) Data

The data are survey data and are drawn from200 households from the State of Louisiana. Tax considerations are often important in making investments. The survey began asking the question - In your family if you were to earn an extra dollar of income, about what percent of that would have to be paid in federal income taxes?

Own Home Yes/ No

Age of Respondent

The Respondent and/or Spouse are 65 or not?

Sex of Respondent Male/Female

Marital Status

Education of Respondent

Number of Household

We compare *PERCEPT* with *TAXCOMP*, the actual marginal federal income tax rates that household face. Since we have access to federal income tax returns for our sample, actual rates were estimated by assuming that each household took the standard deduction. Exemptions were estimated using sample information on the number of children, marital status and whether the respondent and/or spouse were age sixty five or older. Table 1 demonstrates, 60 % of the people surveyed (150) were able to provide an estimate of their household's marginal federal income tax rate. However, among those who did, the mean perceived marginal rate (*PERCEPT*) was 17.34%, while the mean computed marginal rate (*TAXCOMP*) was higher at 20.54%. The difference of the two rates (*DIFF*) has a mean of - 3.201% indicating that computed rates were higher than perceived rates, with a standard deviation of 14.882%. A simple paired sample test, however, shows

[7] Very simplistically Mills ratio (see Maddala [11]) can be represented as follows. Use of the inverse Mills ratio is often motivated by the following property of the truncated normal distribution If *X* is a random variable having a normal distribution with mean μ and variance σ^2, then

DIFF to be significantly different from zero. The appropriate test statistic is

$$t = -3.201 / (14.881 / (sqr(150))) = -2.63$$

Table 1 : Variables Means and Standard Deviation

	KNOWTAX		*DIFFERENCE*	
	Mean	SD	Mean	SD
KNOWTAX = 1 if yes = 0 otherwise	0.877	0.54	-	-
(*PERCEPT*) Perceived Marginal Tax Rate	-	-	31.154	9.212
Computed Marginal Tax Rat (*TAXCOMP*)	25,111	9.231	38.112	18.434
DIFF = *PERCEPT* – *TAXCOMF*			-6.963	11.312
OWN HOME [=1 if yes, 0 otherwise]	.917	0.563	0.783	0.512
Age of Respondent	49.322	15.341	44.212	18.542
Sex of Respondent [=1 if male, = 0 otherwise]	0.543	0.511	0.589	0.508
Education of the Respondent	14.566	1.342	14.321	3.454
Number in Household	3.231	1.434	3.254	1.412
Number of Observations	200		171	

One reason that TAXCOMP may have succeeded PERCEPT is the possibility that the household may have itemized deductions. To account for this possibility, we have included a dummy variable for homeownership (OWNHOME) in the specification. Other explanatory variables include age, Sex, and year of education of the respondent and the number of households. This last variable was entered for additional information costs relating to the magnitude of deductions encountered as household size increases and tax returns become more complex.

Table 2

Probit Criterion Function and the Determinants of the Difference between Perceived and Computed Marginal Federal Income Tax Rate with or Without Correction for Selectivity Bias.

Dependent Variable	*KNOWTAX*	*DIFF*	*DIFF*
CONSTANT	-0.542	17.333	31.421
	(-1.31)	(2.43)	(2.32)
OENHOME	-0.045	-2.345	-3.671
	(-0.56)	(3.87)	(-1.77))
Age of Respondent	-0.057	-0.333	-0. 243
	(-3.68)	(-7.65)	(-8.45)
Sex of Respondent	0.712	-7.543	-9.133
	(4.31)	(-7.43)	(-2.35)
Education of Respondent	0.054	-0.577	-3.231
	(7.07)	(-3-23)	(-1.76)
Number of Household	-0.052	-0.670	-7.751
	(-0.45)	(-2.12)	(-2.23)
LAMBDA	–		-26.76
			(-0.67)
Likelihood Ratio Test	577.357		94.23
R^2		0.05431	0.0131

To find the factors responsible for the *DIFF,* we first checked for selectivity bias following Heckman's methodology illustrated above. We present the results establishing a probit criterion equation (*KNOWTAX*), reflecting the probability of sample inclusion. The results were then used to construct a regressor, the inverse of the Mills ratio (*LAMPDA*) that is decreasing monotonic function of the probability that an observation is selected as the sample. The selectivity bias as stated above in the methodology means that if it is not corrected then regressors that do not belong to the structural equations appear to be statistically significant.

b) Estimates

The estimates are presented in Table 2.

Table 3

Probit Criterion Function and the Determinants of the Difference between Perceived and Computed Marginal Federal Income Tax Rate with or Without Correction for Selectivity Bias.

Dependent Variable	*KNOWTAX*	*DIFF*	*DIFF*
CONSTANT	-0.542	17.333	31.421
	(-1.31)	(2.43)	(2.32)
OENHOME	-0.045	-2.345	-3.671
	(-0.56)	(3.87)	(-1.77))
Age of Respondent	-0.057	-0.333	-0. 243
	(-3.68)	(-7.65)	(-8.45)
Sex of Respondent	0.712	-7.543	-9.133
	(4.31)	(-7.43)	(-2.35)
Education of Respondent	0.054	-0.577	-3.231
	(7.07)	(-3-23)	(-1.76)
Number of Household	-0.052	-0.670	-7.751
	(-0.45)	(-2.12)	(-2.23)
LAMBDA	–		-26.76
			(-0.67)
Likelihood Ratio Test	577.357		94.23
R^2		0.05431	0.0131

We find the determinants of *KNOWTAX* and then compute the coefficients of regressors in *DIFF* with and without *LAMBDA.* The computation was performed in LIMDEP (Green[7]). The coefficients generally differ only modestly as *LAMBDA* is not statistically significant from zero. The level of education of the respondent while significant in specifications excluding *LAMBDA* are not statistically significant when adjusted for selectivity bias. The lack of significance of *LAMBDA* can be explained by the fact that overall those more likely to be included in the sample do not systematically overestimate or underestimate their marginal tax rates. In Table 3a and 3b younger and more educated respondents (more than 12 years) are more likely to report an estimate of their marginal tax rate. However young, male and more educated respondents make larger errors. The probability of reporting may not seem to be independent of home ownership. Owners make larger errors which are statistically significant.

Table 3a

Mean Value of Responses by Respondents

	n	Full Sample *KNOWTAX*
OWN HOME	98	.732
		(.521)
Age of Respondents		
Less than 25	42	.324
		(.114)
Above 50	33	234
		(.201)
Sex of Respondents		
Male	108	.789
		(.643)
Female	72	.542
		(.478)

Education (more than 12 years)	76	.753 (.345)
Number of Households 3 or more	106	.678 (.367)
Overall	130	.579 (.723)

Table 3 b

Mean Value of Responses by Respondents

		Censored Sample		
	n	*PERCEPT*	*TAXCOMP*	*DIFF*
OWN HOME	70	32.12 (11.45)	36.41 (14.71)	-14.29 (17.23)
Age of Respondents				
Less than 25	34	21.47 (17.11)	17.74 (7.84)	13.73 (17.33)
Above 50	43	13.56 (11.45)	17.46 (9.45)	-3.10 (7.43)
Sex of Respondents				
Male	91	19.43 (8.56)	23.56 (14.57)	-13.13 (13.44)
Female	59	23.77 (17.59)	28.53 (23.79)	-3.76 (14.78)
Education (more than 12 years)	60	17.41 (13.77)	19.56 (11,81)	-23.15 (19.76)
Number of Households 3 or more	91	31.79 (17.59)	37.55 (21.54)	-5.76 (12.87)
Overall	99	29.46 (19.58)	34.67 (21.61)	-5.21 (19.32)

Overall, respondents underestimate federal income tax liability by about 5 percentage points. Sex does not seem to have a significant impact on *DIFF.*

III. Conclusion

Although tax payers in general underestimate their marginal tax rates, the difference is not big and this can be explained by taxpayers' use of the standard deduction in computing marginal tax rates. Our results, consistent with the previous established result, provide evidence to the fact that tax payers accurately perceive the marginal tax rates. Consequently, there is little support for existence of fiscal illusion as a determinant of market behavior. Actually, the fiscal illusion is a concept that the government finds it easy to raise tax revenues because of the consumer's ignorance about the way the tax system works. More needs to be done to limit the government ability to collect higher tax revenues otherwise government spending has a tendency to rise 'crowding out' the more efficient private sector.

References Références Referencias

1. Breeden, C and Hunter, W (1985)'ax Structure and Tax Revenue' *Public Finance Quarterly*, 13, 4, 216- 24.
2. Buchanan J. and Wagner, R. F. (1977) Democracy in Deficit: The Political Legacy of Lord Keynes, New York, Academic Press.
3. Chetty, R., Looney, A and Kroft, K (2009) Salience and Taxation: Theory and Practice, *American Economic Review*, 99, 4, 1145-77.
4. Das, Amaresh and Adnan Omar (2014) Fiscal Illusion – Does it Exist? An Econometric Evaluation, International Journal of Economics, Finance and Management, Vol 3, No 3, 136-40.
5. Downs, Anthony (1957) An Economic Theory of Democracy, New York, Harper and Row.
6. Feller, W (1971) An Introduction to Probability Theory and its Applications, Vol 2, Wiley, New York.

7. Greene William (1983) LIMDEP, New York Graduate School of Business Administration, New York University, 1985.
8. Heckman, James (1979) 'Sample Selection Bias as a Specification Error'. *Econometrica*, 47, 1 153-61.
9. Johnson, N and Kotz S (1972) Distributions in Statistics: Continuous Multivariate Distributions, Wiley New York.
10. Lewis, Gregg (1974) 'Comments on Selectivity Bias in Wage Comparison' *Journal of Political Economy*. November, 37-61.
11. Maddala, G. S. (1985) Limited Dependent and Qualitative Variables in Econometrics, Cambridge University, Press, Cambridge.
12. Payne, j E (2003) 'A Survey of the International Empirical Evidence on the Tax-Spend Decline'. *Public Finance*, 31 (3), 302-24.
13. Romer, C. D and Romer, D. H (2007) 'Do Tax-cuts Starve the Beast?' The Effects of Tax Changes Government Spending' University of California, Berkeley, Working Paper (http://elsa.berkeley.edu-cromer/draft 1007 pdf
14. Tobin, J (1958] Estimation of Relationships for Limited Dependent Variables' *Econometrica*, 26, 24-36.
15. Turnbull, Geoffrey (1998) 'The Overspending and the Flypaper Effects of Fiscal Illusion: Theory and Empirical Evidence' *Journal of Urban Economics,* 44, (1), 1-26.