Tax Perception and Sample Selection Bias: Microeconometrics

Adnan Omar

Received: 6 February 2015 Accepted: 4 March 2015 Published: 15 March 2015

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

---

Index terms— fiscal illusion, probit model, sample selection bias, censored model, mills ratio.

1 Introduction

he 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 twostage 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 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],

, income elasticity of the tax structure, debt illusion, and what is known as the 'fly paper effect'. For evaluation of the II.

2 Methodology and Data

In investigating a bias that arises from using an incomplete sample to estimate , 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 1Y in U X Y 1i i i 1 i . + = ? 1

In the case of a censored sample, we have access to the larger random sample but we do not know 1 Y for censored observation. In a truncated Y is available. In both cases, there is a sample of 1 I complete observations.

The population regression function for equation may be written as:Y (E 1i â?"? i X 1) = Ii i N 2, 1, i X = 1 1 ?

which would be estimable from a random sample. The regression function for an incomplete sample may be written asY (E â?"? i X 1 , sample selection rule) U (E X 1i 1 i + ? 1 â?"?sample selection rule) (1) i 1 N 1, i =

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 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 ifC Y i ? 1 Where C is constant.

i Y 1 may be interpreted as an index of a tax payer's intensity of desire. If the intensity is sufficiently great (i Y 1 >C) the tax payer expresses his desire and i Y 1 is observed. Otherwise we cannot observe intensity and observed payment of taxes are zero. In Tobin's model the sample selection rule is given by1 Y (E â?"? Y , X i 1 i ? 1 0) = 1 i U (E X + 1 1 ? â?"? Y i ? 1 0)

We consider a la Tobin the following decision rule: we obtain data on i Y 1 , if another random variable creates a threshold, i. e., ifY i ? 2 0

while if the opposite inequality holds we do not obtain data on i Y 1 . 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 propertiesd i = 1 if Y i ? 2 0 d i = 0 otherwise.

proceed to analyze the joint distribution of i Y 1 and d i dispensing with i Y 2 altogether. The advantage of using selection rule representation is that it permits a unified summary of the existing literature 4 . 4 A good example of this phenomenon is found in Lewis [10]. In his analysis, i Y 1 is the wage rate which is only observable for working women, had i Y 2 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.

3 Using this representation we may write equation (1) as

5 i 1 Y (E â?"? 0 1 ? i Y X 2 i) = i 2 i i U (E X + ? 1 â?"? 2 2 i X U ? = 2) (2) 5 If i U 1 is independent of i U 2 the conditional mean of i U 1 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 i X 2 . 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.If i U 1 is independent of i U 2 the conditional mean of i U 1 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 i X 2 . Moreover, the effect of such sample selection is that 2 X 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 i Y 2 up to a positive factor of proportionality if i Y 2 is positive. 6 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.

4 Global Journal of Management and Business Research

Volume XV Issue V Version I Year 2015( ) C h (E i â?"? Y , X 2i i ? 2 0) = ? ? ? ? ? ? ? ? Y , X Y E 2i 2 2i ? 0 (3) E (Y U 2i i ? 1 0) = E (U U 2i i 1 >- 2 i X ? 2) = i / 22) (? ? ? 2 1 12 E (2i i Y U 2 > 0) = E (2 2i 2i i X - U U ? > 2) = i /) (? ? ? 2 1 22 22 Suppose that) U , U (h 2i 1i

, the joint density of U i 1 and i U 2 is bivariate normal. Using well results of the literature (Jonson and Kotz [9]).

where) (F) (f i i i ? ? ? ? = 1 i / 22 2) (X ? ? ? = 2 1 2

And f and F are respectively, are the density and distribution function of the standard normal distribution.

5 The Tobin model is special case with h (

i i U , U**2 1**

) a singular density sine2i i U U = 1 i

? is the inverse of Mill's ratio 7 i ? and is known as the hazard rate in reliability theory. There are several interesting properties of . ? Its denominator is the probability that observation i has data for i Y

? The lower the probability that an observation has data on Y, the greater the value of ? for that observation. Moreover, using a result in Feller [6] i i > ? ? ? ? 0 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 ? and variance ? 2 , thenE [x ? >]) - ( ) - (? µ ? ? ? µ ? ? ? µ ? + 1 E [? < x] =)) ? µ ? ? ? µ ? ? ? µ - (- (? +

Where ? is a constant, ? denotes the standard normal density function, and ? is the standard normal cumulative distribution function. The two fractions are the inverse Mills ratios. lim? = i ? lim i = ? 0 ? ? i ? ? ? i ?

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

6 a) Data

The data are survey data and are drawn from200 households from the State of Louisiana. 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 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. 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. The estimates are presented in Table 2.

Table 3 Probit 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. Overall, respondents underestimate federal income tax liability by about 5 percentage points. Sex does not seem to have a significant impact on DIFF.

7 III.

8 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. [1] [2] [3]

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

Figure 1:

1

KNOWTAX		DIFFERENCE	
Mean	SD	Mean	SD

Figure 2: Table 1 :

[1] 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.

2

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

Figure 3: Table 2

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

Figure 4:

3

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

Figure 5: Table 3 b

Education	76	.753
(more than 12 years)		(.345)
Number of Households		
3 or more	106	.678
		(.367)
Overall	130	.579
		(.723)
	4. Das, Amaresh and Adnan Omar (2014) Fiscal Illusion -	

Figure 6:

[Payne ()] 'A Survey of the International Empirical Evidence on the Tax-Spend Decline'. Payne . *Public Finance* 2003. 31 (3) p. .

[Breeden and Hunter ()] 'ax Structure and Tax Revenue'. C Breeden , W Hunter . *Public Finance Quarterly* 1985. 13 p. .

[Lewis (1974)] 'Comments on Selectivity Bias in Wage Comparison'. Gregg Lewis . *Journal of Political Economy* 1974. November. p. .

[Buchanan and Wagner ()] *Democracy in Deficit: The Political Legacy of Lord Keynes*, J Buchanan , R F Wagner . 1977. New York: Academic Press.

[Johnson and Kotz ()] *Distributions in Statistics: Continuous Multivariate Distributions*, N Johnson , S Kotz . 1972. New York: Wiley.

[Romer and Romer ()] *Do Tax-cuts Starve the Beast?' The Effects of Tax Changes Government Spending*, C Romer , D Romer . http://elsa.berkeley.edu-cromer/draft1007pdf 2007. University of California, Berkeley (Working Paper)

[Tobin ()] *Estimation of Relationships for Limited Dependent Variables' Econometrica*, J Tobin . 1958. 26 p. .

[William ()] *LIMDEP, New York Graduate School of Business Administration*, Greene William . 1983. 1985. New York University

[Maddala ()] *Limited Dependent and Qualitative Variables in Econometrics*, G S Maddala . 1985. Cambridge: Cambridge University, Press.

[Chetty et al. ()] 'Salience and Taxation: Theory and Practice'. R Chetty , Looney , K Kroft . *American Economic Review* 2009. 99 p. .

[Heckman ()] 'Sample Selection Bias as a Specification Error'. James Heckman . *Econometrica* 1979. 47 p. .

[Turnbull ()] 'The Overspending and the Flypaper Effects of Fiscal Illusion: Theory and Empirical Evidence'. Geoffrey Turnbull . *Journal of Urban Economics* 1998. 44 (1) p. .