ISSN 0709-9231

CAHIER 8735

Expenditures on Lotteries: What do People Say and What Do They Do? An Econometric Analysis

by

G.A. Brenner[1], *C.Montmarquette*[2]
and
R. Brenner[2]

[1] Ecole des Hautes Etudes Commerciales, Université de Montréal.

[2] Département des sciences économiques and C.R.D.E., Université de Montréal.

September 1987

We thank André Tremblay for his research assistance, Sophie Mahseredjian for the helpful comments and the Social Sciences and Humanities Research Council of Canada for providing financial support.

Cette étude a été publiée grâce à une subvention du fonds F.C.A.R. pour l'aide et le soutien à la recherche. Ce cahier a également été publié par le Centre de recherche et développement en économique (Publication #3087).

RESUME-ABSTRACT

On reconnait généralement que les individus sous-estiment volontairement leurs dépenses de consommation d'alcool et de tabac dans les données d'enquêtes. Le même problème existe pour les dépenses de loteries: les dépenses déclarées ne représentent que 60 à 65% des revenus des loteries d'états. En utilisant les revenus officiels des loteries gérées par les gouvernements, il est possible de tenir compte des sous-déclarations des dépenses par les individus et d'obtenir des élasticités-revenus plus précises des dépenses de loteries. L'analyse statistique permet de tester plusieurs hypothèses du modèle de loterie de Brenner et suggère des implications importantes à la fois pour l'analyse économétrique et la confiance que l'on peut accorder aux estimés d'élasticités dérivés des données aggrégées sectorielles pour l'ensemble des dépenses de consommation.

Mots-clés: Sous-déclaration de consommation, dépenses de loteries.

* * * * * * *

Underdeclarations are typical when alcohol, tobacco and gambling consumptions are questioned in surveys. Recent surveys on expenditures on lotteries have similar problems: the declared expenditures equal between 60 to 65 percent of the revenues of the various state-run lottery enterprises. By using the relatively accurate data on the revenue side of this industry one can deal with the problem of underdeclarations of consumption patterns in surveys and obtain better income elasticity estimates. The statistical analysis permits to test specific hypotheses on a lottery model developed by Brenner, and suggests broader implications both for future econometric analysis and the confidence one gives to elasticity estimates derived from aggregate, sectorial data for all consumption expenditures.

Keywords: underdeclaration of consumptions, lottery expenditures.

INTRODUCTION

For reasons of shyness or embarrassment, it is generally recognized that information on expenditures for some goods, as revealed in surveys, are biased downward. For example, it has long been recognized that surveys on expenditures on alcohols and cigarettes have this default[1]. Recent surveys on expenditures on lotteries have similar problems : the declared expenditures equal between 60 to 65 percent of the revenues of the various state-run lottery enterprises[2]. Analysts of the determinants of lottery expenditures, if aware of this difficulty, have rarely mentioned it and they never dealt specifically with the problem in their statistical analyses. We propose in this paper a way to take it into account and thus be able to make better predictions on the future revenues of this industry. The examination may be important not only because it shows a way to correct such bias by statistical methods, but also because it tests a hypothesis (presented in *Brenner (1983, 1985)*) concerning participation in games of chance where one can win a relatively large prize.

The paper is organized as follows. In section I, we summarize the model of lottery expenditures developed by *Brenner (1983, 1985)* and we present some predictions based on the model. Section II, shows a specific way to correct for the biased declarations of lottery expenditures. Section III discusses the empirical results, and conclusions follow.

I. A MODEL OF THE DETERMINANTS OF LOTTERY EXPENDITURES

Brenner (1983, 1985) argues that one participates in games of chance where there is a possibility of winning a big prize either because one is relatively poor, or because one has suddenly become significantly poorer[3]. Brenner's model assumes that one's satisfaction depends not only on his wealth, W_o, but also on his expected position in the wealth distribution:

$$(1) \qquad U = U(W_o, \alpha(W > \bar{W}_o)/W_o \in \tilde{W})$$

where α denotes the percentage of people who are richer than the mean wealth, $\bar{W}_o$, of the class to which the individual belongs ($\tilde{W}$ being the distribution of wealth within this class). Let us assume than an increase in the individual's wealth increases his satisfaction that is $\partial U/\partial W_o > 0$, but an increase in the fraction of people who are richer decreases his satisfaction, $\partial U/\partial\alpha < 0$. Let h be the price of a lottery ticket where there is a probability p of winning a large prize H.

Let the utility function of equation (1) be linear in W_o and α:

$$(2) \qquad U = aW_o + b\alpha(W > \bar{W}_o) \qquad a > 0 \quad , \quad b < 0$$

An individual will by a lottery ticket if

$$(3) \qquad (1 - p)a(W_o - h) + (1 - p)b\alpha(W > \bar{W}_o) + pa(W_o + H) + pb\alpha(W > \bar{W}_1) > aW_o + b\alpha(W > \bar{W}_o)$$

where $\bar{W}_1$ is the mean wealth of the higher class of wealth. If the lottery is fair, $(1-p)h = pH$, equation (3) reduced to:

$$(4) \qquad \alpha(W > \bar{W}_1) < \alpha(W > \bar{W}_o)$$

Equation (4) implies that the relatively poor will always participate in lottery. If the lottery is unfair, this conditions is weakened.

With a similar argument, *Brenner (1985, appendix 1.2)* shows that an individual will increase his participation to lottery when his wealth suddently diminishes.

The model yields the following predictions:

- that the relatively poor will plan to spend a greater fraction of their wealth on lotteries than the relatively rich, and that people of all classes (upper, middle, lower), who have not previously gambled, may decide to do so when they suddenly lose part of their wealth (for example when they are suddenly fired).

- that older, rather than younger, people (with the same measured income) will gamble. The reason is that a $15,000 income for a fifty-year-old is a different indicator of wealth than the same income for a twenty-year-old. The older man, knowing that he can no longer become significantly richer by pursuing opportunities in the labor market, knows that only by gambling he still may achieve that goal.

- that people who have more children will tend to gamble more. The rationale behind this statement is that for a person with one child, a $15,000 income assures a higher position in the distribution of wealth than it does for one with four children : the greater the number of children, the poorer the family (for the same income).

These statements already suggest why some studies about gambling may have been misleading[4]. Information was gathered in them on the income of the buyer, or his household, without looking at either the gamblers' age or family structure. At times, such information on income only seemed sufficient to reach the conclusion that lower-income groups tend to gamble more. At other times, the income of the gambling population seemed, at first sight, too high to support this claim. However, as pointed out above, income may provide very biased information about changes in one's position in the distribution of wealth, and additional considerations must be made to correct the bias : we shall make them below and check the facts.

The aforementioned problem, related with the attempt to get better information on people's *"wealth"*, is one that surfaces frequently, and is thus not so unusual[5]. What is more unusual with respect to expenditures on lotteries is the fact that these expenses have been frequently condemned in the past, and even today the word *"gambler"* has a negative connotation (among Protestants in particular)[6]. The negative attitude probably explains why many people do not tell the truth about their expenses on lotteries. Already in 1895 *Henry Higgs*, an early investigator of people's expenditure patterns, noted that his efforts to collect data on *"amusement"* expenditures (which included those on gambling) encountered severe resistance since people thought that the data will be used to demonstrate *"the bad management of the working classes"*. He had to abolish the category and published data on food, clothing and housing only[7].

The underdeclaration, as pointed out in the introduction, is about 35 to 40 percent, the sum of declared expenditures being equal to 60 to 65 percent of the revenues of the lottery enterprises. A possible way to deal with this problem of underdeclaration in a statistical analysis is examined next.

II. HOW TO CORRECT FOR UNDERDECLARED LOTTERY EXPENDITURES

Let equation (5) defines the determinants of the exact annual spending on lotteries of the i^{th} family as a percentage of its total after tax income:

$$(5) \qquad \ln (E^*/INC)_i = \gamma_0 + \gamma_1 INC_i + \gamma_2 INC_i^2 + \sum_{k=3}^{p} \gamma_k x_{ik} + \varepsilon_i$$

with E_i^* : the unobserved exact lottery expenditures by the family unit

INC_i: the after tax income of family i.

x_{ik}, k=3, ..., p : the other variables of the model, defined in Table 1.

ε_i : an error term with the usual properties.

Consider that

$$E_i^* = (1 + B_i)E_i \tag{6}$$

where E_i is the reported expenditures and B_i is an adjustment coefficient, which equals the ratio of the difference between the family unobserved exact lottery expenditures and the reported ones, and the reported survey expenditures.

Since, according to the model, the poorer have greater incentives to gamble, but it is their expenditure pattern that is condemned, let us assume that the adjustment coefficient varies among families according to the following relationship[a] :

$$B_i = \beta_0 + \beta_1 INC_i + \beta_2 INC_i^2 \tag{7}$$

This relationship implies that the adjustment coefficient depends non linearly on the income of the family unit i. Depending on the signs and sizes of the coefficients β_1 and β_2, this equation may examine the assumption that families with lower incomes underdeclare significantly their expenditures on lotteries.

With <u>reported</u> survey lottery expenditures, equation (5) becomes:

$$\ln (E/INC)_i = \gamma_0 + \gamma_1 INC_i + \gamma_2 INC_i^2 + \sum_{k=3}^{p} \gamma_k x_{ik} - \ln(1+B_i) + \varepsilon_i \tag{8}$$

Assume that $0 \leq B_i \leq 1$. With equation (7) and a series expansion of ln $(1 + B_i)$ up to the second term, we obtain after some manipulations the following relationship :

$$(9)\qquad \ln(E/INC)_i = (\gamma_o + 1/2\ \beta_o^2 - \beta_o) + (\gamma_1 - \beta_1 + \beta_o\beta_1)INC_i + (-\beta_2 + \beta_o\beta_2 + \beta_1^2/2)INC_i^2 + \beta_1\beta_2\ INC_i^3 + \beta_2^2/2\ INC_i^4 + \sum_{k=3}^{p}\gamma_k\, x_{ik} + \varepsilon_i$$

$$= \theta_o + \theta_1 INC + \theta_2 INC^2 + \theta_3 INC^3 + \theta_4 INC^4 + \sum_{k=3}^{p}\gamma_k\, x_{ik} + \varepsilon_i$$

The coefficients of the model are identified except for β_o. However, from the sum of declared expenditures, E, and the observed revenues of lottery entreprises, E^*, we can obtain an aggregate ajustment coefficient B from the following equation:

$$(10)\qquad E^* = (1 + B)E$$

From equation (6) and (7):

$$(11)\qquad E_i^* = (1 + B_i)E_i = E_i + \beta_o E_i + \beta_1 INC_i E_i + \beta_2 INC_i^2\ E_i.$$

Summing up all the family units, assuming $\Sigma E_i^* = E^*$, and using equation (10), (recalling that the value of B is known) we obtain :

$$(12)\qquad \beta_o = B - \beta_1 \sum_i INC_i E_i/E - \beta_2\ \Sigma\ INC_i^2 E_i/E$$

III. THE EMPIRICAL RESULTS

The data set we use is derived from the Survey of Consumer Expenditures done by Statistics Canada in 1982. The initial sample size is 10,937 but since it is weighted according to the ten provinces to represent the Canadian population, the adjusted sample size for each regression may vary accordingly[9].

Thirty-five percent (35%) of the respondents in the sample we took declared that they did not spend on lotteries. Thus one has to deal with this specification bias issue too in explaining lottery expenditures[10].

In order to obtain consistent estimates for the regressions coefficients of equation (9), we used the two stage estimators proposed by *Heckman (1979)*. The results of the probit analysis on the probability of spending on lotteries reject the assumption of randomly missing observations[11] (see Appendix), but the inverses of the Mill's ratio introduced to correct for the sample selection bias proved statistically insignificant.

Table 2 presents the regression results both without the corrections for the underdeclared lottery expenditures and with the corrections. As expected, there are important differences between the two results for the income variables. But for most of the other variables the differences are minor. One exception is the EDUC2 variable whose coefficient is negative and statistically significant in the corrected regression, a change that is in the direction of supporting the predictions of the model.

Focusing on the corrected version, we see that many results are in accord with predictions made in section I. The French-speaking Canadians (LANG2) spend more on lotteries than the English speaking families, and it is well known that French Canadians earn less than their English speaking brethren. (This may not be the only reason for the French Canadians' greater propensity to gamble : Catholics everywhere gamble more than Protestants, and French Canadian are mostly Catholic, but we did not have data on the religious background). Immigrants are persons uprooted from their traditional way of life, whose incomes are lower than those of non-immigrants with similar educational backgrounds and who possibly did not yet settle in their new country. Table 2 shows that recent immigrants (YM4) spend more than both less recent ones (YM3, YM2) and native-born Canadians (the omitted category). Also as expected, up to the age of 72.5, people gamble more with age. Strong results supporting the model allude to the education variables. When one is more educated he spends much less on lottery expenditures. However, the results concerning the number of children living with the family and the number of members in the family receiving unemployment benefits do not support the model.

Table 1

Definition of Variables and Symbols

E/INC	:	annual spending on lotteries of the family as a percentage of its total after tax income.
REG1	:	= 1, if the family lives in the Atlantic province. = 0, otherwise.
REG2	:	= 1, if the family lives in Quebec. = 0, otherwise.
REG4	:	= 1, if the family lives in the Western Provinces except British Columbia. = 0, otherwise.
REG5	:	= 1, if the family lives in British Columbia. = 0, otherwise.
	:	- The omitted category is Ontario.
MS2	:	= 1, if the head of the family was never married. = 0, otherwise.
MS3	:	= 1, if the head was neither married, nor ever married. = 0, otherwise. - The omitted category is when the head of the family is married.
LANG2	:	= 1, if the maternal language of the head of the family is French. = 0, otherwise.
LANG3	:	= 1, if the maternal language of the head of the family is neither English nor French. = 0, otherwise.
	:	- The omitted category is English as maternal language.
YM2	:	= 1, if the head of the family immigrated before 1960. = 0, otherwise.
YM3	:	= 1, if the head of the family immigrated between 1961 and 1970. = 0, otherwise.
YM4	:	= 1, if the head of the family immigrated since 1971. = 0, otherwise. - The omitted category is a native-born for the head of the family.

Table 1 Cont'd

Definition of Variables and Symbols

AGE	:	Age of the head of the family.
AGE^2	:	Age squared.
NCHD	:	Number of children in the family under the age of 16.
NUB	:	Number of members of the family receiving unemployment insurance benefits.
EDUC2	:	= 1, if the head of the family has either some secondary education or completed his secondary education. = 0, otherwise.
EDUC3	:	= 1, if the head has some post-secondary education. = 0, otherwise.
EDUC4	:	= 1, if the head has a college degree or certificate of post secondary education. = 0, otherwise.
EDUC5	:	= 1, if the head has a university degree. = 0, otherwise. - The omitted category refers to fewer than 9 years of primary education.
INC	:	After tax family income.
INC^2	:	Income squared.

Table 2

The Determinants of Lottery Expenditures

Variables	With correction for underdeclarations	Without correction for underdeclaratins
REG1	.0200 (.19)	-5.610 (-.59)
REG2	-.0087 (-.15)	.0021 (.03)
REG4	-.2750 (-1.82)	-.4077 (-3.04)
REG5	-.1960 (-1.68)	-.2897 (-2.76)
MS2	.0150 (.14)	-.0627 (-.66)
MS3	-.1179 (-1.28)	-.1831 (-2.22)
LANG2	.2332 (3.62)	.2598 (4.12)
LANG3	.1043 (1.83)	.1180 (2.09)
YM1	.0024 (.04)	-.0224 (-.36)
YM3	.1276 (1.57)	.1162 (1.44)
YM4	.2921 (3.36)	.2986 (3.44)
AGE	.0435 (2.63)	.0574 (3.87)
$(AGE)^2$	-.0003 (-2.00)	-.0005 (-3.17)

Table 2 Cont'd
The Determinants of Lottery Expenditures

Variables	With correction for underdeclarations	Without correction for underdeclarations
NCHD	-.0841 (-2.94)	-.1050 (-3.97)
NUB	-.0099 (-.18)	.0323 (.64)
EDUC2	-.0089 (-2.01)	.0023 (.05)
EDUC3	-.1599 (-2.69)	-.1714 (-2.89)
EDUC4	-.2425 (-2.65)	-.3084 (-3.61)
EDUC5	-.4597 (2.08)	-.6496 (-3.31)
INC	-.5078 (-3.59)	-.2999 (-4.63)
$(INC)^2$	.1405 (2.58)	.0050 (.40)
$(INC)^3$	-.0200 (-2.62)	
$(INC)^4$	.0010 (2.71)	
CTE	-.5.89 (-9.45)	-6.62 (-12.06)
$(\text{Mill's Ratio})^{-1}$	-.2593 (-.62)	.1244 (.34)
$\bar{R}^2$	.1409	.1403
Number of Observations	7083	7083

The numbers in parentheses are the t-statistics.

Before considering the results on the income variable, let us discuss the adjustment coefficients for the underdeclared lottery expenditures. Let us assume that the adjustment coefficient B of equation (10) is .42. This value is obtained for 1982 in comparing the weighted sums people said they spent on lotteries with the revenues of the respective Provincial lotteries adjusted to account for foreigners buying lottery tickets in Canada[12]. Table 3 presents the corresponding adjustment coefficients for different levels of family income .

Table 3

Adjustment Coefficients and Income Elasticities

Income (in 10^4\$)	Adjustment equation: $B_i=1.291-.4474INC_i+.0447INC_i^2$	Income Elasticities: $\frac{\partial(E/INC)\ INC}{\partial INC\ (E/INC)}$ with correction $-.3726INC_i-.1454INC_i^2$	without correction $-.2999INC_i+.01INC_i^2$
.5	1.07	-.22	-.15
1	.88	-.52	-.29
2.5	.45	-1.84	-.69
3.0	.35	-2.43	-.81
6.0	.21	-7.47	-1.44

These results show the importance of correcting for the underdeclared lottery expenditures: for a low income family, the underdeclaration reaches 100%. Table 3 also shows the income elasticities of money spent on lotteries relative to the family income. There are considerable downward biases in the income elasticities when not correcting for the underdeclared expenditures. As the family income rises, the uncorrected income elasticities represent between 68 to a mere 19 percent of the corrected income elasticities. The corrected negative income elasticities seem to strongly support the model of Section I.

IV. CONCLUSIONS

The statistical analyses above seem useful not only because they enable to test the hypotheses advanced in Section I, but more because they show a way to deal with problems of under and overdeclarations of consumption patterns in surveys. As mentioned, underdeclarations are typical when alcohol, tobacco and gambling consumptions are questioned (overdeclaration, bragging that is, seems to be a problem when people's sexual appetite is questioned). But by using the relatively accurate data on the revenue side of these three industries one can obtain better elasticity estimates. A way to obtain them was suggested in this study.

The study also points to a more general problem concerning demand curves. We found few studies, one of the exceptions being *Houthakker and Taylor (1970)*, where attempts were made to contrast the estimated total expenditures with the revenues for survey data with time series data of the respective sectors in order to check the reliability of the results. For many sectors the lack of this step may not pose a major problem, since there is no reason to believe that people would systematically declare under or over-spending on bread or clothing. Still, if people significantly underdeclare expenditures on gambling, alcohol, tobacco (not to mention those on drugs, prostitutes or investment in the Bahamas), this must imply that they will consciously declare either more than they spend on other, *"ordinary"*, legal items with a positive connotation, or under-declare revenues. Since expenditures on all items with a negative association, illegal or even legal, are well into the tenth of billions of dollars, there must be significant overdeclaration on ordinary items for which, a priori, one could have thought that only random, uncorrelated, errors existed. Thus, the problems we have dealt with here may have broader implications both for future econometric analysis and the confidence one gives to elasticity estimates derived from aggregate, sectorial data for *"ordinary"* items too.

Appendix

Probit Analysis of Lottery Expenditures

Variables	Coefficient (t-statistic)	Variables	Coefficient (t-statistic)
REG1	-.2561 (-5.12)	$(AGE)^2$	-.0005 (-8.39)
REG2	.0469 (.95)	NCHD	-.0716 (4.93)
REG4	-.4275 (-11.61)	NUB	.1519 (5.92)
REG5	-.3178 (-7.67)	$EDUC^2$	.0465 (1.29)
MS2	-.2838 (-6.50)	$EDUC^3$	-.0386 (-.75)
MS3	-.2519 (-6.83)	$EDUC^4$	-.2190 (-4.48)
LANG2	.0850 (1.76)	$EDUC^5$	-.5948 (11.02)
LANG3	.0385 (.85)	INC	.2771 (10.20)
YM2	-.0800 (-1.68)	$(INC)^2$	-.0238 (-7.24)
YM3	-.0363 (-.53)	CTE	-.7771 (-5.75)
YM4	.0238 (.33)		
AGE	.0472 (8.16)		

Likelihood Ratio Test: 1154.29
(degree of freedom): 21
Number of Positive Observations: 7174
Total Observations[13]: 11032

FOOTNOTES

1) See *Rubner (1966), Houthakker and Taylor (1970), Brenner and Brenner (1987a).*

2) The expenditures on Canadian lotteries by U.S. citizens cannot explain the difference, since they were estimated to be around 200 million dollars, when the revenues of the Canadian lotteries exceeded the two billion mark. See *Brenner and Brenner (1987a).*

3) The prediction of the model are tested there and confronted with additional evidence, namely, those available for winners of the big prizes. At the time that data set was the only one available. Additional examinations are done in *Brenner (1986).*

4) Reference to these studies can be found in *Brenner (1983, 1985), Brenner and Brenner (1987a).*

5) See it discussed in another context in *Lazear and Michael (1980).*

6) See *Brenner and Brenner (1987b), Devereux (1980), Tec (1969).*

7) See *McKibbin (1979).*

8) Other specifications were tried but they were not supported by the data. Note that a more general specification will add an error term to equation (7).

9) The sample size may even exceed the 10,937 figure, some observations being repeated to respect the weights corresponding to the provinces.

10) Some people might not have reported buying lottery tickets at all. But it is difficult to assess the importance of this problem because the information is from surveys only (which, as shown in the text are not entirely reliable).

11) The probabilities of a χ^2_{p-1} exceeding the calculated χ^2 are less than 0.01. Thus we reject the model subject to the constraint that all regression coefficients (except the constant term) are zero at a 99 percent confidence level.

12) We have assumed that $E/E^* = .70$ by excluding the foreigners about whom see footnote 2. Therefore from equation (10) we obtain a conservative value of $B = .42$.

13) See footnote 9. Note also that by referring to a weighted sample, the results from the probit analysis provide general parameters to correct for selectivity bias affecting the regressions of our study.

REFERENCES

Brenner, Gabrielle A., "Why Do People Gamble? Further Canadian Evidence", Journal of Gambling Behavior, Vol. 2, no 2, Fall/Winter 1986, pp. 121-129.

Brenner, Gabrielle and Reuven Brenner, "A Profile of Gamblers", cahier de recherche 1187, Centre de recherche et de développement en économique, Université de Montréal, 1987a.

Brenner, Reuven, History - the Human Gamble, Chicago : University of Chicago Press, 1983.

Brenner, Reuven, Betting on Ideas : Wars, Invention, Inflation, Chicago : University of Chicago Press, 1985.

Brenner, Reuven and Gabrielle A. Brenner, "Changing Attitudes Toward Games of Chance : from Antiquity to Our Times", working paper, Department of Economics, Université de Montréal, 1987b.

Devereux, Edward G. Jr., Gambling and the Social Structure, 1950, Ph.D. dissertation Harvard University, reprinted, 1980.

Heckman, James J., "Sample Selection Bias as a Specification Error", Econometrica, January 1979, Vol. 47 : 153-61.

Houthakker, H.S. and Lester D. Taylor, Consumer Demand in the United States: Analyses and Projections, 2nd ed., Cambridge, Massachusetts: Harvard University Press.

Lazear, Edward P. and Robert T. Michael, "Family Size and the Distribution of Real per Capita Income", American Economic Review, Vol. 70, 1980 : 90-107.

McKibbin, Ross, "Working-class Gambling in Britain 1880-1939", Past and Present, no 82, February 1979 : 147-78.

Rubner, Alex, The Economics of Gambling, London : MacMillan, 1966.

Tec, Nehama, Gambling in Sweden, Totowa, N.J. : Bedminster Press, 1964.