Cigarette demand and tax policy for race groups in South Africa

Gregory D. Berg & William H. Kaempfer

To cite this article: Gregory D. Berg & William H. Kaempfer (2001) Cigarette demand and tax policy for race groups in South Africa, Applied Economics, 33:9, 1167-1173

To link to this article: https://doi.org/10.1080/00036840122752

Published online: 04 Oct 2010.

Submit your article to this journal

Article views: 66

Citing articles: 4 View citing articles
Cigarette demand and tax policy for race groups in South Africa

GREGORY D. BERG* and WILLIAM H. KAEMPFER†

2050 W. 10th Ave., #204, Broomfield, CO 80020, USA and †University of Colorado at Boulder, Campus Box 256, Boulder, CO 80309, USA
E-mail: *gregtania.berg@mindspring.com and †william.kaempfer@colorado.edu

This paper calculates cigarette demand for race groups in South Africa. Elasticities are the most important information a tax policy analyst can have. Elasticities determine how the tax base will change with a change in the tax rate and thus how government revenues will respond to the tax. Elasticities also determine the excess burden that consumers will bear as a result of the tax. As such, own price, cross-price, and expenditure elasticities are calculated along with government revenue maximizing tax rates, and total and excess burdens. Parametric and semiparametric estimation techniques are used and compared. Results show that a tax on cigarettes will discourage nonsmokers from starting to smoke and mainly raise revenue from current smokers. Furthermore, it is found that consumption behaviours between groups are different implying different government revenue maximizing tax rates for each group affecting the distribution of income.

I. INTRODUCTION

It is apparent to the casual observer and expert alike that income and expenditure levels differ greatly between White and Black South Africans. Apartheid contributed to this disparity. The Gini index measures the degree of inequality between income or expenditure among households. A Gini coefficient of zero represents perfect equality whereas a Gini coefficient of 1 represents perfect inequality. The expenditure Gini index calculated from this data set is 0.584 which compares to other Upper-middle-income countries as follows: 0.503 in Mexico, 0.634 in Brazil 0.266 in Czech Republic, 0.484 in Malaysia, 0.270 in Hungary, 0.565 in Chile and 0.282 in Slovenia. Table 1 shows the distribution of consumption using the 1993 Living Standards Measurement Survey conducted by the World Bank just prior to the first all race democratic elections in South Africa.

Table 2 shows the Gini index and share of people in poverty for both race groups. There is a higher percentage of Blacks in poverty than Whites. The Gini coefficient was 0.449 for Blacks and 0.336 for Whites.

PCE stands for per capita expenditure defined as total household expenditure (rand per month) divided by the number of households for both Blacks and Whites. The poverty measure uses a poverty line of 105 rand per month per person, which is approximately US$1 per day per person in the USA. The poverty measure uses the head-count ratio defined as the fraction of the population below the poverty line of 105 rand per month. Using 105 rand per month as the poverty line there are no Whites that are in poverty as compared to 31.7% of Blacks. There are approximately 6300 Black households and 1350 White households in the survey.

Tax policy can affect the distribution of income. Elasticities are the most important information a tax policy analyst can have. Elasticities determine how the tax base will change with a change in the tax rate and thus how government revenues will respond to the tax. Elasticities also determine the excess burden that consumers will bear as a result of the tax. The more price elastic is demand the more excess burden will be associated with a tax increase. If the government can determine which commodities exhibited inelasticity or a relative inelasticity between
population groups, then a tax policy that aimed at levelling the income distribution can be established.

In most poor countries, a large fraction of government revenue is raised by indirect taxes on goods and services, and many countries subsidise the prices of commodities such as basic foodstuffs. Household expenditure surveys, by revealing who buys each good and how much they spend, tell us who pays taxes and who benefits from subsidies. They thus yield a reckoning of the gainers and losers from a proposed change in taxes and subsidies (Deaton, 1997, p. 4).

II. METHODS AND PROCEDURES

Censored maximum likelihood (ML) and censored least absolute deviation (LAD) estimation techniques are used to estimate demand equations from household level data that include the zero responses. As a result, household characteristics can be specified in the utility function and thus in the demand or share equations. However, as noted by Budney (1989), heteroscedasticity is a problem for the censored maximum likelihood if the conditional variance is not constant or the functional form of the heteroscedasticity is not correctly specified. This limitation is averted by semiparametric methods such as the censored LAD, which does not require an assumption of the error process (Budney, 1989, p. 148).

The assumptions required for consistency of the censored LAD are much weaker than the assumptions required of the censored ML. The median of the censored dependent variable does not depend on a functional form for the errors. This makes the censored LAD a semiparametric estimator (distribution-free). This property of being distribution-free is a property not shared by the mean of the dependent variable. Censored ML requires estimation of the mean and thus is not a distribution-free estimator (Powell, 1984, p. 307). This is the basis of Budney’s (1989) comments above. The implications of using a semiparametric estimator is that no specification of the error process need be made with absolutely no way of misspecifying the error process as is possible in the censored ML. Of course in both censored ML and LAD the correct functional form of the regression is crucial for consistency. Only in nonparametric estimation is this not a problem where no functional form is assumed for the error process nor the regression.

The functional form is derived from the linear expenditure system of utility specification with demographic translation as described in Pollak and Wales (1992). The top seven expenditure commodity items were used in calculation of total expenditure $\mu$.

\[
 h'(P, \mu) = b_1 - \left( \frac{a_i}{P_i} \right) \sum P_K b_K + \left( \frac{a_i}{P_i} \right) \mu
\]

where $h'(P, \mu)$ is quantity demand; $P$ is vector of seven prices; $\mu$ is expenditure on seven commodities; $P_i$ is price of commodity $i (i = 1, \ldots, 7)$; $b_K = b_{0K} + b_{1K}(NA) + b_{2K}(NC)$, $(k = 1, \ldots, 7)$; $NA$ is number of adults in household; $NC$ is number of children in household; $a_i$ and $b_K$’s are parameters to be estimated.

Deaton develops the censored LAD estimation for and the estimated asymptotic variance covariance matrix of the parameters which is used in this paper.

Deaton demonstrates that with 1000 observations the censored LAD is closer to the truth in a Monte Carlo exercise 96% of the time as compared to OLS or the censored ML in the presence of heteroscedasticity. Deaton further suggests at least comparing the censored ML and LAD to see how departures from normality or heteroscedasticity affect estimates (Deaton, 1997, pp. 89-90). The sample size of Blacks in the estimation of this thesis is 1131 and the sample size of Whites is 998.

Elasticities are calculated using the standard formulas. There are two types of quantity responses a consumer can make when prices/incomes change depending on whether the household is currently consuming the good. If the consumer is currently consuming the good, the responsiveness to a price/income change is the typical Marshallian elasticity, which represents an inframarginal response. If the household is currently not consuming the good, there might be a switch from not consuming to consuming, which is the intermarginal response. The calculated elasticities are the inframarginal responses.

The standard errors of the elasticities were calculated using the delta method. Let $C$ represent the vector of par-
ameters that are estimated, \( a_i's, b_{K1}'s, b_{K2}'s \) etc. Let \( V \) represent the variance–covariance matrix of the parameters. Let \( G \) represent the gradient of the elasticity with respect to the parameters. For the own price elasticity of commodity 1, \( \varepsilon_{11} \), \( G = \partial \varepsilon_{11} / \partial C \). Note that this \( G \) is a vector with dimensions 1 by the number of parameters so the variance of \( \varepsilon_{11} \) is \( GV^2 \). The standard error of \( \varepsilon_{11} \), s.e.\( (\varepsilon_{11}) \), is the square root of the variance. The standard errors for the other elasticities are calculated in the same manner using the delta method.

The reported \( t \)-statistics were obtained by testing the hypothesis that the elasticity was equal to zero. The household level data show that there are 1131 urban Black households where all seven prices are present. There are 998 urban White households where all seven prices are present. Cigarette consumption had a 53.8% censoring for Black households and 51.5% censoring for White households.

### III. MODEL SELECTION AND ELASTICITY RESULTS

The censored maximum likelihood estimates of the household level data is done using TSP assuming homoscedasticity where the standard deviation of the errors is equal to \( \sigma \), and again assuming the functional form of the heteroscedasticity where the standard deviations of the errors is equal to

\[
\sigma^2 [\exp(g1^* \text{ income} + g2^* \text{ number of adults} + g3^* \text{ number of children})]^{1/2}
\]

where, \( \sigma, g1, g2, g3 \) are parameters to be estimated. If all three parameters \( g1, g2, g3 \) are equal to zero then the heteroscedastic model reduces to the homoscedastic model.

The censored least absolute deviation estimation was done using the Buchinsky (1994) Iterative Linear Programming Algorithm (ILPA) to obtain Powell’s (1984) estimator. The variance–covariance matrix is computed using the asymptotic covariance matrix Powell provides.

The competing models are (1) censored ML assuming homoscedasticity, (2) censored ML assuming heteroscedasticity, and (3) censored LAD.

Censored ML (Tobit) assuming homoscedasticity (the restricted model) and heteroscedasticity (the unrestricted model) uses the household level data with the zero responses treated as a corner solution. The competing censored maximum likelihood models were compared using a likelihood ratio test. The restricted model (homoscedasti-
city) had the restrictions placed on the parameters $g_1$, $g_2$, and $g_3$ where all three are assumed to be zero. The unrestricted model (heteroscedasticity) allowed the parameters $g_1$, $g_2$, and $g_3$ to take on any value. The log likelihood for both the restricted and unrestricted models are presented in Table 5 for both Blacks and Whites along with the likelihood ratio test of the restrictions that the heteroscedastic model reduces to the homoscedastic model.

With three degrees of freedom from the three restrictions, the restricted model was rejected in favour of the heteroscedastic models. The $\chi^2$ critical value with three degrees of freedom at the 99% confidence level is 11.345. Both of the likelihood test statistics were larger than this critical value.

The two remaining models are the censored ML with heteroscedasticity and the censored LAD. A Hausman (1978) test is performed on these two models. The Hausman test requires a test of two different models estimated from the same data and the same coefficients. In the present context, the null hypothesis is that the censored ML is the correct specification and the form of heteroscedasticity is correct. Censored ML under the null hypothesis is consistent and efficient. Under the alternative hypothesis, that the functional form of heteroscedasticity is not correctly specified, censored ML is neither consistent nor efficient. The censored LAD is consistent under both hypotheses. The test statistic is

$$(B_{\text{LAD}} - B_{\text{ML}})'(V_{\text{LAD}} - V_{\text{ML}})^{-1}(B_{\text{LAD}} - B_{\text{ML}})$$

where $B$ is a $(22 \times 1)$ vector of parameter estimates from the LAD and ML models and $V$ is the variance/covariance matrix from the LAD and ML models. This test statistic is distributed $\chi^2$ with 22 degrees of freedom from the 22 parameters being tested against each other in the two competing estimation techniques. The intuition behind the statistic is that when the null hypothesis is true the quadratic form in the differences in the parameters is small. However when the null hypothesis is not true, the quadratic form is large. The Hausman test statistic and $P$-values are in Table 6 for Blacks and Whites for each demand equation.

For Blacks and Whites, at the 99% ($\alpha = 0.01$) confidence level the null hypothesis of censored ML being consistent and efficient is not rejected. Generally in cross-sectional data, as used in this paper, the heteroscedasticity is associated somehow with the level of income. The exact nature of that association is not generally known. The results of the Hausman test show that both the censored ML and the censored LAD give consistent estimators whereas the censored ML estimates are efficient. Table 7 shows the estimated elasticities for both methods of estimation.

Since both the censored ML and censored LAD are both consistent estimates and the Hausman could not distinguish between the competing models the censored LAD elasticity results are used in the remaining analysis. The reason is that the censored LAD, being a fortiori robust to heteroscedasticity, is preferred to the censored ML that must make assumptions about the exact parametric nature of the heteroscedasticity. Furthermore the censored ML gives an own price elasticity that is positive for Blacks and apparently against the law of demand. For Blacks neither the censored ML nor the censored LAD gives elasticity estimates significantly different from zero. The censored ML elasticity for Whites only gives one elasticity estimate that is statistically different from zero, the elasticity of cigarette with respect to income is positive. The censored LAD estimates gives six of ten elasticities statistically significant.

### IV. Own Price and Income Elasticity Analysis

The demand for cigarettes for both groups is calculated above with a tax analysis to follow. Also calculated is the demand for cigarettes for both groups conditional on purchase of cigarettes. Previously all households are included in the estimation process including the households who have a zero purchase of cigarettes. To compare elasticities, cigarette demand is calculated for only those households where cigarette purchases is positive. See Table 8.

Looking at the own price elasticities, the estimates using all observations are more elastic than the observations using only the positive purchase observations. However the elasticity estimates using the positive observations are positive. Normally own price elasticities are negative indicating a downward Marshallian demand function. The

<table>
<thead>
<tr>
<th>Cigarettes</th>
<th>Value</th>
<th>Black's Hausman</th>
<th>Black's P-value</th>
<th>White's Hausman</th>
<th>White's P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.50</td>
<td></td>
<td>24.34</td>
<td>0.330</td>
<td>24.34</td>
<td>0.330</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cigarettes</th>
<th>Value</th>
<th>Black's Hausman</th>
<th>Black's P-value</th>
<th>White's Hausman</th>
<th>White's P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.50</td>
<td></td>
<td>24.34</td>
<td>0.330</td>
<td>24.34</td>
<td>0.330</td>
</tr>
</tbody>
</table>

Table 6. Hausman test results for censored ML and censored LAD models.
only elasticity that is statistically different from zero is the own price elasticity for Whites using all observations.

The difference between the two rows labelled cigarettes (all observations) and cigarettes (positive observations) is that the row labelled cigarettes (all observations) used all household observations in the estimation of the censored demand function for cigarettes. The row labelled cigarettes (positive observations) uses only the households with a positive purchase of cigarettes in the estimation of a cigarette demand function. This latter demand function is not censored as all households had a positive purchase of cigarettes. The least absolute deviation estimator is used due to the robustness against heteroscedasticity.

Looking at the income elasticities, the elasticities calculated from only positive observations is less than the elasticities calculated from all observations. All of the income elasticities have a positive expected sign meaning that cigarettes are a normal good. This shows that among smokers, cigarette consumption is less income elastic than among the race group as a whole. This can be said of the White group with more certainty than the Black group. Income elasticity estimates for Whites are statistically different from zero for both demand functions.

The null hypothesis that the income elasticity of Whites using all observations is equal to the income elasticity of Whites using only the positive observations is tested. The alternative hypothesis that the income elasticity of Whites using all observations is larger than the income elasticity of Whites using only the positive observations. The t-statistic was 5.08. This statistic rejects the null hypothesis in favour of the alternative at the 99% confidence level. So with 99% confidence the elasticity as calculated from all sample observations is larger than the elasticity calculated from only positive observations.
V. GOVERNMENT REVENUE AND TAX ANALYSIS

The government revenue maximizing tax rates are calculated for both groups separately and together. Both demand models are used: (1) parameters from the censored demand estimation which included all observations (zeros included) and (2) parameters from the demand estimation which included only the observations where cigarette consumption is positive (no zeros). As can be seen the revenue maximizing tax rate calculated using only the households that did purchase cigarettes is much larger than the tax rate calculated using all observations (Table 9). This could be an indication that, among smokers, their behaviour measured in terms of purchasing cigarettes, did not change much with a large tax rate that maximized government revenues. In all three calculations of the government revenue tax rate, the tax rate is larger when just calculated from the demand parameters from only positive cigarette consumption.

Further indication of the difference between the behavior based on the entire sample and the sample conditional on purchase of cigarettes is government revenues. Since behaviour does not change by much among smokers, the average government revenue per smoking household is much higher than the average household of the entire sample. Furthermore, excess burdens are larger for the higher tax rate when calculated using positive observations due to the larger government revenue maximizing tax rate. Table 10 uses the tax rate that maximized government revenue for both Blacks and Whites together.

The average income for Blacks conditional on purchasing cigarettes is 599 compared to the average income for all Blacks, which is 327. It is consistent with the large income elasticity of Blacks when all of the households were used in the calculation. Better off Blacks smoked.

VI. CONCLUSIONS

Elasticities are the most important information a tax policy analyst can have. Elasticities determine how the tax base will change with a change in the tax rate and thus how government revenues will respond to the tax. Elasticities also determine the excess burden that consumers will bear as a result of the tax. The more price elastic is demand the more excess burden will be associated with a tax increase.

If the government could determine which commodities exhibited inelasticity or a relative inelasticity between population groups, then a tax policy aimed at levelling the income distribution can be established. Taxes on domestic goods and services comprise 34.9% of total current revenue for South Africa in 1991–1995 (World Bank, 1997, p. 197).

The implications of the cigarette analysis are consistent with arguments that raising the tax on cigarettes has two effects. It reduces smoking by smokers somewhat, but also discourages nonsmokers from becoming smokers. This is shown with the own price elasticity estimates from the sample of all households and the sample of households conditional on the purchase of cigarettes. The price responsiveness of cigarette consumption conditional on smoking is less than the price responsiveness of cigarette consumption using the entire sample of households. This shows that the general population is more likely to reduce cigarette consumption with a tax increase than the segment of the population that currently smokes. Thus a tax increase on cigarettes discourages nonsmokers from becoming smokers.

| Table 9. Government revenue maximizing tax rates for both samples of observations. |
|----------------------------------------|--------|--------|--------|
|                                       | Both groups | Whites | Blacks |
|                                       | Tax | Tax rate | Tax | Tax rate | Tax | Tax rate |
| Cigarettes (all observations)         | 1.4 | 0.52 | 1.7 | 0.6 | 1.4 | 0.5 |
| Cigarettes (positive observations)    | 5.46 | 2.06 | 4.39 | 1.65 | 14.01 | 5.30 |

| Table 10. Government revenues and excess burdens for both samples of observations. |
|----------------------------------------|--------|--------|--------|
|                                       | Government revenues (per household) | Excess burden (per household) |
|                                       | Blacks | Whites | Blacks | Whites |
| Cigarettes (all observations)         | 2.0 | 31.0 | 5.0 | 17.3 |
| Cigarettes (positive observations)    | 101.8 | 244.3 | 16.8 | 119.2 |
Cigarette demand and tax policy

However, the premise that an increase in the tax on cigarettes causes a decrease in cigarette consumption by smokers is not met with the South African data set. The price responsiveness is positive meaning that an increase in price causes more cigarette consumption by smokers. These price elasticities are not statistically significantly different from zero so no definite conclusion can be inferred.

REFERENCES