A Robust Estimation of the Effects of Taxation on Charitable Contributions

Ralph Bradley
Steven Holden
Robert McClelland

February, 1999

Ralph Bradley and Robert McClelland: Bureau of Labor Statistics, Room 3105, 2 Massachusetts Ave., NE, Washington DC 20212, Steven Holden: Johns Hopkins University. We would like to thank Joe Daniel, Richard Steinberg, Gene Steurle and participants of seminars at the American Economic Association, the Association for Research on Nonprofit Organizations and Voluntary Action, the Urban Institute and George Washington University for helpful comments on this paper. All errors are our own. The views expressed here do not reflect the policy of the Bureau of Labor Statistics or the views of other BLS staff members.

INTRODUCTION

From the many studies on taxation's effect upon charitable contributions arises the stylized fact that the 'tax-price' (defined below) elasticity of giving exceeds unity. If true, then the federal policy of allowing households to deduct contributions from taxable income causes more to be contributed than is lost to the treasury in foregone tax revenue. Recent work using panels of tax data, however, suggests that the long-run tax price elasticity may be substantially lower than unity. In this case the deduction is inefficient in that the long run loss to the federal treasury exceeds the gains to charities, so that it is less costly for the government to remove the deduction and provide the funds or services directly.

Here we address this issue using cross-sectional household data from the Consumer Expenditure Survey (CEX) to estimate the price elasticity of all deductible contributions. A common solution to problems caused by the large percentage of noncontributors in the sample is to use, for example, methods such as Tobit or the two-stage method of Heckman. Specification tests, however, reject the assumptions necessary for the consistency of these estimates. This casts doubt upon the result that tax-price elasticities estimated with these maximum likelihood methods exceed unity, even when using variables designed to capture the same long run features as the panel data.

When using the semiparametric method of Ahn and Powell (1993) – which consistently estimates parameters without making analogous assumptions – we find an elasticity estimate similar to those found using panel data. This suggests that high tax-price elasticities in previous work with cross-sectional data may be as much a function of specification error as the lack of a time component.

However, our data allows us to estimate the elasticity of contributions to social welfare organizations. This more closely matches the definition of giving needed to measure efficiency because the other major recipient of contributions, religious organizations, cannot receive government expenditures as a substitute for private donations. Semi-parametric estimates using this definition find an elasticity greater than one. Therefore, while it is possible that many previous estimates of the elasticity are too high, the deduction for contributions may still be treasury efficient.

THE MODEL

We are interested in estimating a model of contributions C_i* :

(1) $C_i* = c_i(Y_i b_Y + P_i b_P + Z_i b_Z + e_i),$

where Y_i is permanent income, Z_i is a vector of demographic variables and P_i is the tax-price of charitable giving. The variable c_i represents a first stage in which the household chooses to contribute if the utility $U(C_i,X_i,T_i,Z_i)$ from giving exceeds the utility from not giving. That is,

(2) $c_i = I[U(C_i*, Y_i - P_i{\cdot}C_i*, T_i, Z_i) + n_i > U(0, Y_i, T_i, Z_i)],$

where the function $I[\cdot]$ is an indicator function equal to one if the condition is true and zero otherwise and n_i is independent and identically distributed (iid) with mean zero.

An variable unique to this first stage is T_i, the level of giving by the government. Its inclusion solely in this stage arises from the idea that the individual has enough information about T_i to make a first stage decision, but not enough to use in the second stage. In the second stage giving by others must be evaluated in a cardinal ordering to determine the precise number of dollars to contribute. But it is unreasonable to believe that households have this much information. On the other hand in the first stage households only need to know if either transfer payments are too high or too low. If they are too high then ceteris paribus the household doesn't give and if too low it gives.

Using the standard model in the charitable giving literature, C_i*, Y_i and P_i are measured in natural logs.[1] The coefficients on these variables can therefore be interpreted as elasticities. One reason for interest in the price elasticity is the idea of "treasury efficiency". The subsidy to contributions due to their deductibility is deemed "treasury efficient" if the revenue loss to the treasury from the subsidy is less than the contributions it induces. The removal of a treasury efficient subsidy while holding recipients harmless to the change by substituting government payments would therefore result in a net loss to the treasury.

The intuition relating this concept to price elasticity derives from observing that for contributions C and a flat tax rate m, the treasury loses m´C from the deduction. The deduction then raises more in contributions than is lost if

$$\frac{dC}{dm} > m\frac{dC}{dm} + C .$$

Noting that $dC/dm = - dC/d(1-m)$ and rearranging,

$$\frac{dC}{d(1-m)}\frac{(1-m)}{C} = \frac{dC}{dP}\frac{P}{C} < -1,$$

[1] A dollar is added to contributions.

where P = (1-m) for itemizers. Therefore, a price elasticity greater than 1 in absolute value indicates treasure efficiency.[2]

ESTIMATION METHODS

Because a large proportion of the households make no contributions, using simple generalized least squares on either the whole sample or on households that contribute will produce inconsistent parameter estimates since $c_i e_i$ can be decomposed as follows:

(3) $$c_i e_i = E(e_i | I[U_i(C_i^*) + n_i > U_i(0)]) + h_i$$

$$= l(Y_i, P_i, T_i, Z_i) + h_i$$

where h_i has both a conditional and unconditional expectation equal to 0 and $U_i(C_i^*)$ equals $U(C_i^*, Y_i - P_i \cdot C_i^*, T_i, Z_i)$. Rewriting (1) as

(4) $$C_i^* = c_i Y_i b_Y + c_i P_i b_P + c_i Z_i b_Z + l_i + h_i,$$

where l_i is $l(Y_i,P_i,T_i,Z_i)$, we see that the residual ($l_i + h_i$) can be correlated with the regressors, requiring us to use an estimator accounting for this selection effect. The tobit technique is commonly used in the charitable giving literature to address this selection effect. Another useful method is the two-stage Heckman model in which l_i is estimated using a probit model. Both of these models make parametric assumptions about the distribution of c_i.

In addition to the sample selection problem we must include a correction for the endogeneity of the price variable. Because contributions are deductible from federal taxation and the tax rate varies with income net of deductions, the tax rate facing the

---

[2] Roberts (1987) notes that a small adjustment should be made to this figure to account for the change in contributions

household tax-price may be written as

$$P_i(Y_i, C_i) = 1 - d_i m(Y_i - C_i)$$

where d_i is one if the household itemizes and zero otherwise, and $m(\cdot)$ is the marginal tax rate function. Therefore, because P_i is a function of C_i, it is correlated with e_i. The standard solution to this problem is to use the "first dollar" tax price – the tax price if contributions were zero – as an instrument for the "last dollar" tax price – the actual tax price of the household. We incorporate this solution into the tobit model by using the full information maximum likelihood method described in Greene (1995). This is essentially a two equation system: one equation for the instrumental variable regression and one for the charitable giving regression. As an alternative we incorporate the instrumental variable into the two-stage Heckman model by using the model of Lee, Maddala and Trost (1980). In this model there is a first stage selection equation and a second stage system of equations similar to the Greene model.

One drawback to both the tobit and Heckman models is that a failure of the distributional assumptions on e_i and n_i makes the model's coefficient estimates biased and inconsistent. An alternative is use a semiparametric procedure whose estimates are consistent without requiring specific densities for e_i and n_i.

To this end we use the two stage estimation procedure established by Ahn and Powell (1993). Unlike other semiparametric models such as those of Ichimura and Lee (1991) or Chen and Lee (1998), the estimates from this method do not depend upon starting values selected by the researcher. The primary advantage of Ahn and Powell

---

caused by a change in transfers, but we ignore this effect here because we assume that contributors do not have enough

(1993), however, is that there is no requirement to specify a parametric model of the index component c_i in (2).

Instead, the Ahn and Powell method estimates the index c_i using a kernel regression. Unlike the Heckman method, the results from the first stage are then used to create weights in the estimation of (4) using first differences of the data:

(5) $C_{ij}{*}=[cY]_{ij}b_Y + [cP]_{ij}b_P + [cZ]_{ij}b_Z + l_{ij} + h_{ij}$,

where, for example, $[cY]_{ij} = c_iY_i - c_jY_j$ and $l_{ij} = l_i - l_j$.[3] The essential idea of Ahn and Powell is to use weights that are largest when l_{ij} is closest to zero and smallest when l_{ij} is furthest from zero. The selection effect is then reduced and, asymptotically, eliminated. Formally, if the weights w_{ijn} satisfy the regularity conditions in Appendix 1 and the following properties:

$$
(6) \qquad
\begin{array}{ll}
w_{hkn} > w_{ijn} & \text{if } |l_{ij}| > |l_{hk}| \\
w_{ijn} \xrightarrow{p} 0 & \text{if } l_{ij} \neq 0 \\
w_{ijn} \xrightarrow{p} c > 0 & \text{if } l_{ij} = 0
\end{array}
$$

then a generalized weighted least square regression of (5) produces consistent parameter estimates. Since the lack of parametric assumptions prevents us from identifying l_{ij}, w_{ijn} must be a function of something that is one-to-one with l_{ij}. If this is the case, then l_{ij} being close to zero implies that w_{ijn} will be non-zero.

To solve this problem, Ahn and Powell suggest that one nonparametrically estimate:

$$Pr(c_i = 1|Y_i,P_i,T_i,Z_i) = F(Y_i,P_i,T_i,Z_i) \equiv F_i$$

---

information about government expenditures to set the level of giving.

[3] A sample of size 100 would then become a sample of size 100·99/2=495.

where c_i is defined in (2) as one for contributors and zero otherwise. Ahn and Powell then argue that $|F_i - F_j| > |F_h - F_k|$ implies $l_{ij} > l_{hk}$, in which case F_{ij} is one-to-one with l_{ij}. Although not discussed by them, this follows from the fact that

$$l_i = \frac{\int_{U_i(0)-U_i(C^*)} \int e_i f(e_i, u_i) du_i de_i}{F_i} .$$

Then if F_i is close to F_j, by definition $Pr[n_i > U_i(0)-U_i(C_i^*)]$ is close to $Pr[n_j > U_j(0) - U_j(C_j^*)]$. But because n_i is assumed to come from an i.i.d. distribution, if the probabilities are close it must be the case that $U_i(0)-U_i(C_i^*)$ is close to $U_j(0)-U_j(C_j^*)$. This implies that the lower limit of the integrals in the numerators are close and therefore the numerators are close. If both the numerators and denominators of l_i and l_j are close, then l_i and l_j themselves are close.

Thus, the Ahn and Powell method uses a kernel regression estimate of F_i in the first stage, $\hat{F}_{in}$, where the sample size subscript notes that this estimate is a function of sample size. This is estimated as:

$$\hat{F}_n(Y_i, T_i, P_i, Z_i) = \frac{\sum_{j=1}^{n} K\left(\frac{Y_j - Y_i, T_i - T_j, P_i - P_j, Z_i - Z_j}{h_n}\right) c_j}{\sum_{j=1}^{n} K\left(\frac{Y_j - Y_i, T_i - T_j, P_i - P_j, Z_i - Z_j}{h_n}\right)}$$

where h_n is the window width (or "bandwidth") that approaches zero as the n approaches infinity and in our application $K(\cdot)$ is the standard normal kernel. Larger window widths create smoother regressions weighting observations more evenly. Smaller window widths create more flexible regressions by weighting observations near to i much more than distant observations. Letting

(7) $\hat{F}_{ijn} = \hat{F}_{in} - \hat{F}_{jn}$,

we use the left-hand side of (7) as the argument to a weighting function

(8) $\hat{w}_{ijn} = K(\hat{F}_{ijn} / b_n)$

where $\hat{w}_{ijn}$ is the nonparametric estimator of w_{ijn} and satisfies the properties in (6). Because the window width b_n converges to zero as n converges to infinity, for any nonzero constant *c*, $\lim_{n \to \infty} K(c/b_n)=0$. These conditions alone are sufficient for $\hat{w}_{ijn}$ to satisfy the conditions outlined in (6).

Denoting X_i as the regressors to (5), and V_i as X_i except for the use of the first-dollar price as an instrument for the last dollar price, the parameter estimates to (4) are:

(9) $\hat{b} = (\hat{S}_{VX})^{-1} \hat{S}_{VC}$

(10) $$\hat{S}_{VX} = \binom{n}{2}^{-1} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \hat{w}_{ijn} (X_i - X_j)(V_i - V_j)'$$

(11) $$\hat{S}_{VC} = \binom{n}{2}^{-1} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \hat{w}_{ijn} (C_i^* - C_j^*)(V_i - V_j)'$$

The resemblance of (9) to the standard instrumental variable estimator is reassuring and one can now readily see the intuition behind the consistency of (9). As the sample size increases, the differences in (10) and (11) that have nonzero selection effects, l_{ij}, have weights that converge to zero. Asymptotically, the only observations receiving weights are those for which the selection effect is zero and therefore the parameter estimates are consistent. However, there are three additional issues that need to be addressed when implementing this estimator.

First, this estimator requires the use of two kernel estimators which require order conditions to guarantee the consistency of the parameter estimates. Our initial implementation carefully ensured that the kernels satisfied the Ahn and Powell's order conditions listed in Appendix 1. This was later abandoned for a simple normal kernel because the parameter estimates were, as predicted by Powell, Stock, and Stoker (1989), discontinuous functions of the bandwidth. Aside from the difficulty in meeting the order conditions one cannot satisfy the convergence conditions on the bandwidth in a single sample. Thus, we are faced with the problem of choosing the bandwidths.

The most common solution to this second problem is to “cross-validate” by choosing the bandwidth that minimizes the out-of-sample mean square error. This implied balancing between bias and variance makes sense in a model producing predicted values without parameter estimates. Here, however, cross validation is inappropriate for two reasons. First, our goal is to get a consistent estimate of b_p and so choosing bandwidths to trade-off bias and variance in the predicted value of C_i* is of no value. Even asymptotically cross validation is questionable because the cross validated window width would converge at a rate such that the order of bias in the nonparametric estimators equals the order of the variance. This could easily violate the bandwidth regularity conditions described in Appendix 1.[4] Second, the computational burden of cross-validation is very high: at ten minutes per regression it takes about 16 days to generate the out-of-sample mean square error for a single pair of bandwidths.

A third problem involves estimation of the covariance matrix. In this study our goal is to conduct hypothesis tests on the coefficients in (4). Unfortunately, Ahn and

Powell did not formally prove the consistency of their covariance matrix estimator. In addition, it is a fourth order U-statistic of kernel derivatives and therefore computationally burdensome.

We address the problems arising from using a simple kernel, bandwidth selection and the difficulty in performing hypothesis tests with a nonparametric bootstrap, using the percentile bootstrap method of Hall (1992). Following MacKinnon and Smith (1998), we use the bootstrap to estimate the finite sample bias from using a simple kernel and selecting a bandwidth without cross-validation, which is then subtracted from the parameter estimates. The bootstrap thus aids the use of a simple kernel and our bandwidth selection method, described below, by reducing the finite sample bias inherent in nonlinear estimators. It also relieves the problem of hypothesis testing by allowing us to construct confidence intervals.

Denoting a = .05 as the significance level of these confidence intervals, we choose B=999 samples from the original sample using sampling with replacement. These choices make $(a/2)(B+1)$ an integer so that the 50th largest and smallest coefficient estimates correspond to the .05 and .95 percentiles. Letting j index the bootstrap sample, we denote the j[th] estimated parameter vector by b_j^* and the estimated finite sample bias as b. Then

$$b^* = \frac{1}{B}\sum_{j=1}^{B} b_j^*, \quad b = b^* - \hat{b}$$

Our bias corrected estimate is then

$$\hat{b}_0 = \hat{b} - b = 2\hat{b} - b^*$$

---

[4] See Singh and Ullah (1985) about the properties of cross validated window widths.

and the confidence interval around this estimate is $[2\hat{b} - b^*_{1-a/2}, 2\hat{b} + b^*_{a/2}]$.[5]

DATA

Our data set is taken from the 1982-1984 CEX Interview database (see U. S. Bureau of Labor Statistics (1986) for more information). The CEX is a rotating panel survey that may begin in any month of the year for a given participating household; for example, a household's survey year may run from June 1983 through May 1984. Unlike data derived from tax files, the survey contains information on specific household consumption expenditures such as charitable contributions by type of organization (e.g. religious, civic, etc.), demographic attributes of the household members, location, income by source, and other attributes which affect income tax burdens at the federal, state and local levels.

From the CEX, we draw a sample that includes single person households, husband/wife couples without children, and husband/wife couples with one or more of their own children. Residents of Alaska and Hawaii are deleted, as well as retirees, occupants of student housing, and households that moved during their survey period. In addition, households believed by the interviewer to have answered incompletely all questions pertaining to current income are eliminated, as are those few reporting contributions drastically in excess of their incomes. The remaining households that had

[5] See Hall (1992) for the properties of the percentile method bootstrap. This correction assumes that the bias, b, is constant across all possible values of β. To test this we used the regressors to generate $C_i^{**} = X_i \hat{b} + e_i$ and a normally distributed sample selection method and averaged the estimated bias over 1,000 trials. Comparing this bias estimate to those created from the same process but different parameter values b^* yielded essentially identical estimates.

final interviews between June 1983, and May 1984 were selected and their responses annualized. This entire process results in a sample size of 2,347 households.

Contributions are defined as either the sum of giving to religious, educational and social welfare organizations (CHARITY) or simply as giving to social welfare organizations (SOCIAL). Although the CEX does contain household tax payment information, it omits marginal tax rates. The values of federal, state and local income taxes and marginal tax rates are therefore simulated for each year of the study period, using an iterative algorithm developed for computation of the tax-and-price index (TPI). For more information on this algorithm, see Appendix 2.

Contributions are defined as contributions to religious, educational and social welfare organizations. As described previously, the tax-price of giving is one for non-itemizers and one minus the applicable marginal tax rate for itemizers. It is defined as

$$P = 1 - d\,\frac{m_f + m_s b_d - m_f m_s b_f - m_f m_s b_d}{1 - b_f m_f m_s}$$

where m_f is marginal federal tax rate, m_s is the marginal state tax rate, d is equal to one if the household itemizes and zero otherwise, b_d is the equal to one if deductions are allowed on a state return and zero otherwise, and b_f is equal to one if the household can deduct federal taxation from their state return and zero otherwise.

Although a measure of current income exists, we instead use a measure of permanent income: the sum of all household expenditures except charitable giving. For examples of this approach in general, see Prais and Houthakker (1971). For an example of this with the use of the CEX, see Nelson (1989).

Unlike tax data, the CEX includes variables for age, education, and race of the head of household, all of which are included here. To allow for the effect of government expenditures upon the probability of making a contribution, we include a variable made up of the per capita government expenditures on AFDC, Supplemental Security Income, and Food Stamps in the household's metropolitan statistical area.

RESULTS

In this section we show that price elasticity estimates from several parametric methods exceed unity while the semi-parametric estimate does not. However, standard tests of the distributional assumptions necessary for the consistency of the commonly used Heckman and tobit estimators strongly rejects the null hypothesis that we are using the correct distributions. This suggests that price elasticity estimates in much previous work are too high, although the semiparametric estimate is similar to low estimates found in recent work by Broman (1989), Randolph (1995) and Barrett, McGuirk and Steinberg (1997) using tax file data.

However, our initial definition of charity and the tax file definition includes religious giving. Here the idea of the federal government payments replacing private contributions – the thought experiment behind treasury efficiency – isn't realistic. When we restrict our definition to the more reasonable definition of giving to social welfare organizations, our elasticity estimate exceeds one. It is possible, then, that previous estimates are too high *and* that the charitable giving deduction is treasury efficient.

The coefficient estimates derived from using parametric regression methods on CHARITY, defined as giving religious, educational and social welfare organizations are listed in Table 2. Because we use a log-linear model, the coefficient estimates on income and price can be interpreted as elasticities. The first two columns contain estimates from regressions using standard 2SLS methods on the entire dataset and on the set of contributors, respectively.

In the first column of Table 2 are the results from a naive 2SLS regression without considering the censoring problem. The estimate is significantly greater than unity. But because approximately 40 percent of the households do not give, these estimates should be biased and inconsistent. A 2SLS regression on just the observations is also biased, but produces an elasticity less than one, although not significantly so.

In the third column we list the results from the Heckman model of Lee, Maddala and Trost (1980). Again the price elasticity estimate exceeds one, although it is insignificantly different from one. This estimate is only consistent if the normality assumption in the first stage is satisfied. But using the test of Bera, Jarque and Lee (1984) we find a statistic of 26.93, which is a rejection at any usual significance level for a $\chi^2(2)$ distribution. This estimate is therefore biased and inconsistent.

As an alternative we use the full information maximum likelihood tobit model in Greene (1995). This again yields an estimated price elasticity well in excess of one. This model, however, also relies on the assumption of normality of the disturbance. If this assumption is violated the estimates are biased and inconsistent. Using the test of Pagan and Vella (1989) for censored normality we find a statistic of 758, which is a rather strong rejection for a $\chi^2(2)$ distribution.

This leaves the semi-parametric estimates, which are consistent without making assumptions about the distribution of the disturbances. But before describing the results the bandwidth selection method needs explanation. Our technique relies on the fact that plots of coefficient estimates as a function of bandwidths have a clear interpretation.

Specifically, when the bandwidths are very high the coefficient estimate is the same as 2SLS on the censored data.[6] This occurs because high bandwidths tend to weight data equally. As the bandwidths shrink, at some point the coefficient estimates change. This occurs because the estimated probability of selection, $\hat{F}_i$, begins to affect the weights $\hat{w}_{ijn}$, which in turn differentially weight the observations.

As expected the coefficients invariably move towards their true values. In some cases this results in a "plateau", and in other times a "well", but in either case the top or bottom of this region comes closest to the true coefficient estimate. Still smaller bandwidths cause the parameter estimates to vary widely with the bandwidth. This is analogous to the sensitivity of coefficients estimated with a small number of observations: the addition or subtraction of a few observations can dramatically change the estimate.

In figure one we show the results of this experiment on our data. At the very high bandwidths one can see that the coefficient estimate for price is –0.73. This corresponds to the censored 2SLS estimate in Table 2. At a sufficiently small pair of bandwidths we find a well. The bottom of this well represents our choice for bandwidths. Beyond this point one can see sharp fluctuations in the parameter values, as mentioned. Figure 2 plots the kernel density with the selected bandwidths as a function of income.

---

[6] See Holden (1999) for proof.

Table 3 lists our bias-corrected parameter estimates and confidence intervals using our chosen bandwidths. The results are remarkably similar to those found from the censored 2SLS. The interpretation of this is that there is, in fact, very little selection bias and that the different estimates of the Heckman and tobit models are the result of mis-specification of the likelihood function. Although not significantly below one, the coefficient estimate is well below the other estimates. Because our price and income variables are defined using a measure of permanent income, this estimate can be compared to the estimates of long-run elasticity found in Broman (1989), Randolph (1995) and Barrett, McGuirk and Steinberg (1997).

But because this definition of giving includes religious giving, its price elasticity is not of use when drawing inferences about treasury efficiency. The results from repeating the entire experiment for giving to social welfare organizations, SOCIAL, is listed in Tables 4 and 5. In this case even 2SLS on the censored data produces a parameter estimate in excess of unity although again not significantly so. A likely reason for this is that religious giving is much less sensitive to the tax price than other forms of giving. Again the tests for the distributional assumptions necessary for the consistency of the Heckman and tobit estimates reject the null hypothesis: 34.15 in the case of the Bera, Jarque and Lee test and 811.01 in the case of the Pagan and Vella test. The semiparametric estimate again is close to the censored 2SLS model and so it exceeds unity. This implies that the deduction may be treasury efficient.

CONCLUSION

In this paper, we employ a semiparametric technique to estimate a model of charitable giving. Our results show that price elasticities are substantially lower with this method than with maximum likelihood methods. However, tests on the maximum likelihood models show that they are misspecified. The low elasticity estimates of the semi-parametric approach are similar to those produced from panel data and suggest that many previous estimates are too high. But if one takes the idea of treasury efficiency seriously, then the elasticity of giving to all sources is less meaningful than the elasticity of giving to social welfare organizations. Semi-parametric estimates using this definition find an elasticity greater than one. This implies that it is possible for many previous estimates of the price elasticity to be too high but still have the deduction for contributions be treasury efficient.

Figure 1: Price Elasticity as a Function of Bandwidths

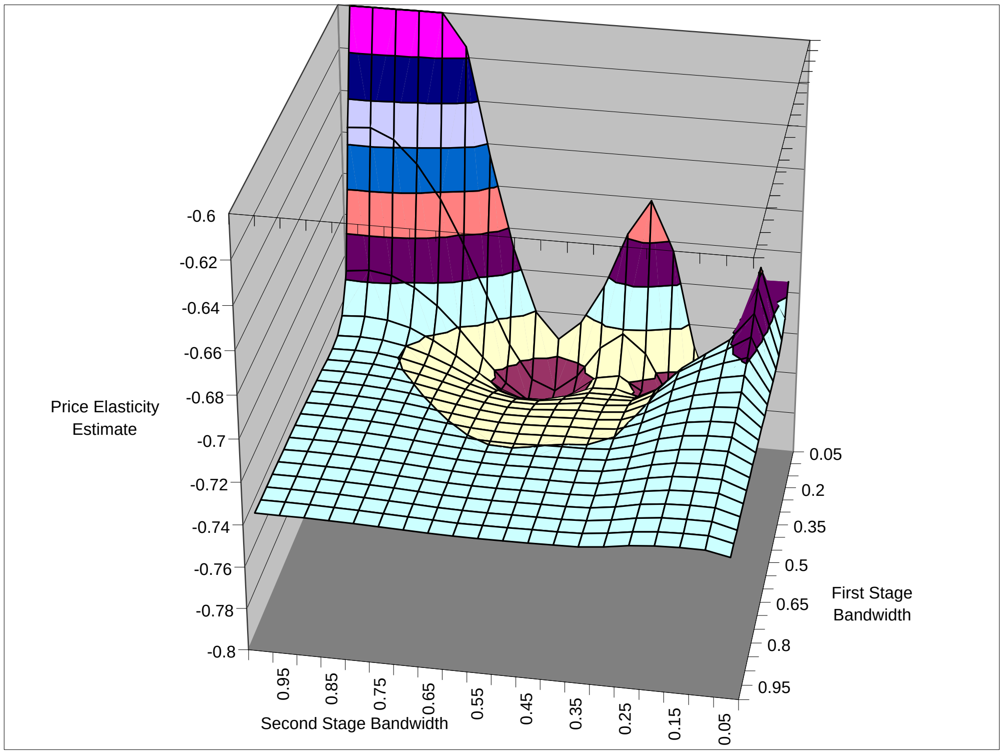

Figure 2: Kernel Density Estimate of Charitable Giving (N=1424)

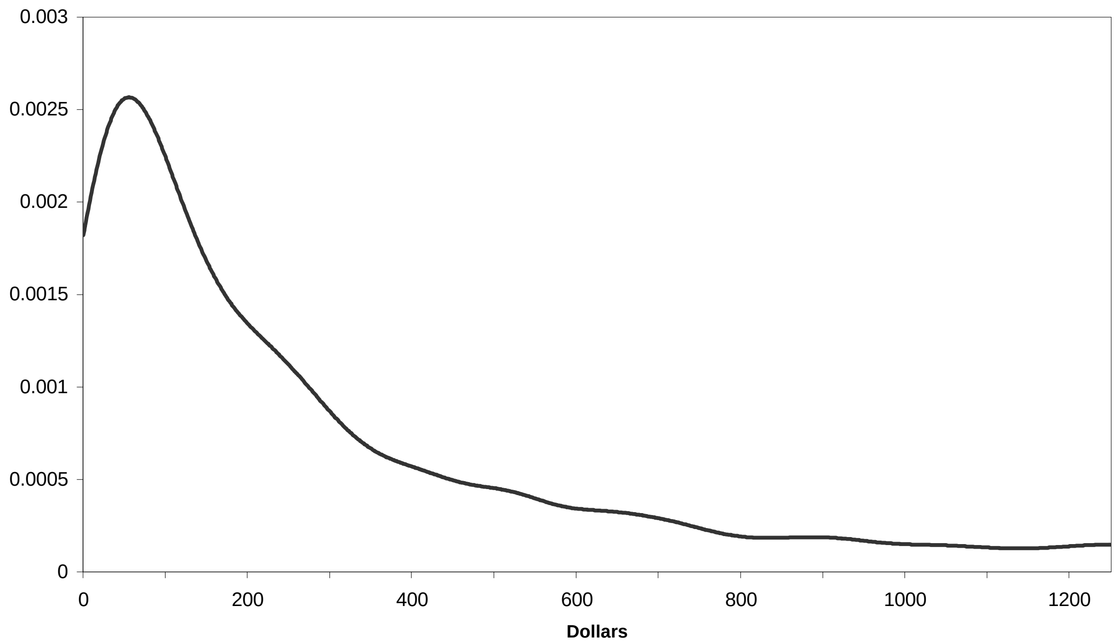

Table 1a: Descriptive Statistics: Full sample (N=2347)

Variable	Mean	Std. Dev.	Min.	Max.
Charity>0 (1424 observations)	541.93	929.28	1.00	9388.55
Social>0 (963 observations)	138.87	284.10	1.00	5000.00
Price	0.88	0.17	0.43	1.00
Income	12974.84	9320.38	174.43	102024.60
Race	0.14	0.35	0.00	1.00
Education	3.42	1.47	1.00	7.00
Age	43.63	17.91	17.00	94.00

Table 1b: Descriptive Statistics: Contributors only (N=1424)

Variable	Mean	Std. Dev.	Min.	Max.
Charity>0 (1424 observations)	541.93	929.28	1.00	9388.55
Price	0.84	0.17	0.43	1.00
Income	15087.36	9703.55	602.25	102024.60
Race	0.12	0.32	0.00	1.00
Education	3.57	1.47	1.00	7.00
Age	45.81	17.26	17.00	93.00

Table 1c: Descriptive Statistics: Social welfare contributors only (N=963)

Variable	Mean	Std. Dev.	Min.	Max.
Social>0 (963 observations)	138.87	284.10	1.00	5000.00
Price	0.81	0.18	0.43	1.00
Income	16379.07	9863.22	602.25	102024.60
Race	0.10	0.30	0.00	1.00
Education	3.76	1.42	1.00	6.00
Age	45.63	16.58	17.00	92.00

Table 2: Parametric Estimates: Dependent Variable = CHARITY

Variable	Full Sample 2SLS	Censored 2SLS	Heckman	Tobit
Price	-1.93	-0.73	-1.29	-3.06
	(-6.91)	(-3.24)	(-2.98)	(-4.95)
Income	0.95	0.55	0.83	1.48
	(12.57)	(8.69)	(5.261)	(10.91)
Race	-0.01	0.20	0.11	-0.14
	(-0.08)	(1.65)	(0.764)	(-0.59)
Education	0.26	0.15	0.23	0.32
	(6.50)	(5.25)	(4.32)	(5.01)
Age	0.03	0.02	0.02	0.05
	(10.74)	(7.39)	(4.25)	(9.17)
Constant	-8.11	-1.41	-5.79	19.01
	(0.69)	(-2.39)	(-2.42)	(-7.44)

Note: t-statistics are in parentheses

Table 3: Semiparametric Estimates: Dependent Variable = CHARITY

Variable	Estimate	Confidence Interval 5% point	95% point
Price	-0.78	-1.08	-0.48
Income	0.56	0.47	0.64
Race	0.17	-0.01	.34
Education	0.16	0.12	0.20
Age*10	0.18	0.15	0.22

Table 4: Parametric Estimates: Dependent Variable = SOCIAL

Variable	Full Sample 2SLS	Censored 2SLS	Heckman	Tobit
Price	-2.56	-1.36	-0.99	-4.28
	(-10.60)	(-6.48)	(-1.79)	(-7.12)
Income	0.53	0.36	0.31	1.56
	(9.27)	(5.40)	(1.99)	(9.06)
Race	-0.11	0.03	-.07	-0.49
	(-0.95)	(0.24)	(0.50)	(-1.64)
Education	0.22	0.15	0.12	0.47
	(7.41)	(5.02)	(2.0)	(6.18)
Age	0.02	0.007	0.05	0.04
	(6.85)	(2.93)	(0.97)	(6.14)
Constant	-5.06	-0.59	0.48	-23.50
	(-9.59)	(-0.92)	(0.19)	(-7.50)

Note: t-statistics are in parentheses

Table 5: Semiparametric Estimates: Dependent Variable = SOCIAL

Variable	Estimate	Confidence Interval 5% point	95% point
Price	-1.34	-1.61	-1.08
Income	0.38	0.29	0.46
Race	0.03	-0.12	0.18
Education	0.15	0.11	0.19
Age*10	0.07	0.03	0.10

APPENDIX 1

Here we discuss some of the application issues when using the Ahn and Powell model.

For consistency of $\hat{b}$ the kernels and the bandwidths in both stages must satisfy regularity conditions which force the bias from the two non-parametric stages to converge to zero at a rate faster than $\sqrt{n}$. Achieving this requires the use of "higher order kernels." The weighting kernel, *k*() in the second stage must satisfy:

$\int u^l k(u)du \ \ l = 1,2,3$ while the second stage band width, b_n, must be

$O(n^{-d})$, where $d \in (\frac{1}{8}, \frac{1}{6})$. Denoting *m* as the number of continuous right hand side regressors, *K()* in the first stage must satisfy the system:

$$K(u) = \sum_{j=1}^{M/2} q_j N(0, t_j C)$$

$$\sum_{j=1}^{M/2} q_j = 1, \ \sum_{j=1}^{M/2} t_j^l q_j = 0, \ l = 1,..M/2$$

$$M > m / (\frac{1}{3} - 2d)$$

where N(0,C) is the density of normal vector, and δ is the convergence rate of the second stage bandwidth. The first stage bandwidth, h_n, must be $O(n^{-g})$, $g \in \left\{ \frac{1}{2M}, (\frac{1}{6} - d)/m) \right\}$.

Notice that the regularity conditions of the window width convergence and the order of the kernels are interdependent.

APPENDIX 2

The TPI algorithm links specific tax code parameters to income, expenditure, and demographic information for a calendar year. A detailed description of this algorithm is contained in Gillingham and Greenlees (1987) and Kokoski (1990). Given the available household-specific information, the algorithm simulates the household's tax liability, including eligibility and amounts of tax credits, exemptions, and deductions at the federal, state and local levels, as well as the FICA tax. Simulated tax payments and marginal tax rates are based upon permanent income, computed from the aggregate of the consumption expenditures of the household. The algorithm endogenizes the decision to itemize based upon the criterion of which option results in the smaller tax payment, thus producing observations of household charitable contributions by both itemizers and non-itemizers. To reconcile the calendar year tax expenditures with the survey period of the households, the imputed tax rates and payments are averaged across the years represented by each household's survey period. Thus, for a household surveyed from June 1983 to May 1984, its marginal tax rate is 7/12 of the 1983 rate plus 5/12 of the 1984 rate.

Information on tax rules is derived from federal publications: Individual Income Tax Returns, and Your Federal Income Tax (Internal Revenue Service Publication 17) for the relevant years. Exemptions include those for taxpayer, spouse, dependents, and elderly status, although blind individuals could not be distinguished in the CEX data. Deductions include: medical care costs, state and local taxes, interest payments, charitable contributions, and certain nonconsumption expenditures, such as union dues and rental of safe deposit boxes. Surcharges, earned income credits (EICs), the Credit for the Elderly, the exemption for contributions to an Individual Retirement Account (IRA), and special credits for couples when both work were incorporated into the computations in the years for which they applied. Information on the FICA tax, is from the Social Security Administration. State and local income taxes are from annual issues of the State Tax Handbook and the State Tax Guide, both publications of the Commerce Clearing House, Inc. It is assumed that all taxpayers itemized deductions in computing their state tax and that there are no variations among states in the deductions for interest, energy conservation investments, medical care expenses, and other miscellaneous deductions. Because we do not have complete data on sales and excise tax rates by state and local areas we assume that there is a single rate of 4.075 per cent for items that are generally not exempted from this tax.

REFERENCES

Ahn, H. and J. L. Powell, 1993, Semiparametric Estimation of Censored Selection Models with a Nonparametric Selection Mechanism, Journal of Econometrics 58, pp 3-29

Barrett, K. et. al., 1997, Further Evidence on the Dynamic Impact of Taxes on Charitable Giving, National Tax Journal 50 (2), pp 321-334

Bera, A. et. al., 1984, Testing the Normality Assumption in Limited Dependent Variable Models, International Economic Review 25(3), pp 563-578

Broman, A., 1989, Statutory Tax Rate Reform and Charitable Contributions: Evidence from a Recent Period of Reform, The Journal of the American Taxation Association Fall 1989, pp 7-21

Chen S. and Lee L-F., 1998, Efficient Semiparametric Scoring Estimation of Sample Selection Models, Econometric Theory 14(4), pp 387-422.

Commerce Clearing House. *State tax guide.* various issues

Commerce Clearing House. *State tax handbook.* various issues

Greene, W.,1995, *Limdep User's Manual*, New York: Econometric Software, Inc.

Gillingham, Robert and John Greenlees, 1987, The impact of direct taxes on the cost of living, Journal of Political Economy 95 (Nov), pp 775-796

Hall P, 1992, *The Bootstrap and Edgeworth Expansion*, Springer-Verlag.

Heckman, J. J., 1976, The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and A Simple Estimator for Such Models, Annals of Economic and Social Measurement 5, pp 475-492

Holden, S, 1999, A Note on Bandwidth Choice for the Ahn-Powell Estimator, mimeo

Ichimura H. and Lee L-F., 1991, Semiparametric Least Squares Estimation of Multiple Index Models: Single Equation Estimation, in Nonparametric and Semiparametric Methods in Econometrics and Statistics ed. Barnett W., J. Powell and G. Tauchen, Cambridge University Press, pp 3-50.

Kingma, B., 1989, An Accurate Measurement of the Crowd-Out Effect, Income Effect, and Price Effect for Charitable Contributions, Journal of Political Economy 97(5), pp 1197-1207

_____ and McClelland, R., 1995, Public Radio Stations Are Really, Really Not Public Goods, Annals of Public and Cooperative Economics 66(1), pp 65-76

Kokoski, M., 1990, Changes in taxes, prices, and the cost-of-living: An analysis of tax reform, BLS Working Paper No. 199

Lankford, L.H. and J.H. Wycoff, 1991, Modeling Charitable giving Using a Box-Cox Standard Tobit Model, Review of Economics and Statistics 73, pp 460-470

Lee, et. al., 1980, Asymptotic Covariance Matrices of Two-Stage Probit and Two-Stage Tobit Methods for Simultaneous Equations Models with Selectivity, Econometrica 48(2), pp 491-503.

MacKinnon, J. G. and A. A. Smith, Jr, 1998, Approximate Bias Correction in Econometrics, Journal of Econometrics 85, pp 205-230

McClelland, R. and M. Kokoski, 1994, Econometric Issues in the Analysis of Charitable Giving, Public Finance Quarterly 22, pp 498-517

Nelson, J., 1989, Individual Consumption within the Household: A Study of Expenditures on Clothing, Journal of Consumer Affairs 23(1), pp 21-44

Pagan, A. and F. Vella, 1989, Diagnostic Tests for Models Based on Individual Data: A Survey, Journal of Applied Econometrics 4(0) Supplement, pp S29-59

Powell, J.L., Stock, J.H. and T.M. Stoker (1989) "Semiparametric Estimation of Index Coefficients" Econometrica vol. 57 (6) 1403-1430.

Prais, S. J. and H.S. Houthakker, 1971, *The Analysis of Family Budgets*, Cambridge: Cambridge University Press

Press, W. H., B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, 1992, *Numerical Recipes in C: The Art of Scientific Computing,* Cambridge: Cambridge University Press

Randoph, W., 1995, Dynamic Income, Progressive Taxes, and the Timing of Charitable Contributions, Journal of Political Economy 103(4), pp 709-38

Roberts, R.D., 1987, Financing Public Goods, Journal of Political Economy 95(2), pp 420-37

Singh, R.S. and A. Ullah, 1985, Nonparametric Time Series Estimation in Nonparametric of Joint DGP, Conditional DGP, and Vector Autoregression, Econometric Theory 1, pp 27-52.

U. S. Bureau of Labor Statistics, 1986, *Consumer Expenditure Survey: Interview Survey 1982-1983*, Bulletin No. 2246 (February)

U.S. Internal Revenue Service *Your Federal Income Tax*, Publication 17, various issues