Nonprofit and Voluntary Sector Quarterly
XX(X) 1–11

Reprints and permission: http://www.sagepub.com/journalsPermissions.nav
DOI: 10.1177/0899764010371010
http://nvsq.sagepub.com

Research Note: Some Econometric Issues in Studying Nonprofit Revenue Interactions Using NCCS Data

Daniel Tinkelman[1] and Daniel G. Neely[2]

Abstract

Several econometric issues arise in using nonprofit data. After controlling for unduly influential observations and heteroskedasticity, regression analysis performed on National Center for Charitable Statistics "digitized data" from 2001 to 2003 found mixed evidence of economically significant associations between donations and other revenue streams; regressions without these controls support different conclusions. Testing does indicate that government support sends greater quality signals than program support or investment income.

Keywords

nonprofit, National Center for Charitable Statistics, crowding out, donations, empirical methods

Introduction

This empirical study of the interactions of nonprofit revenues, using National Center for Charitable Statistics (NCCS) data on approximately 123,000 organizations from 2000 to 2002, seeks to add to the "crowding-out" literature. In particular, we sought to determine whether evidence was consistent with the theory that the receipt of government grants

[1]Hofstra University, Hempstead, NY
[2]University of Wisconsin–Milwaukee

Corresponding Author:
Daniel Tinkelman, Professor of Accounting, Department of Accounting, Taxation and Legal Studies, Zarb School of Business, Hofstra University, 134 Hofstra University, Hempstead, NY 11549-1340
Email: Daniel.tinkelman@hofstra.edu

provides a quality signal that moderates or counteracts "crowding-out" effects. We find some support for this theory, although results vary by nonprofit sector. The econometric issues that arose may also be instructive to other researchers.

An enormous theoretical, experimental, and empirical literature exists on the interaction of contributions and government grants. See Steinberg (1993), Vesterlund (2006), and Tinkelman (2010) for literature reviews. Early theoretical papers suggest that government grants displace private donations dollar for dollar. Other theoretical papers indicate that under varying assumptions, crowd out could be smaller or negative. For example, if donors have imperfect information, government funding may signal quality. The many empirical studies do not, in general, support substantial crowd out.

This article's primary research question is whether, empirically, there are differences in crowding out of donations for government grants, investment income, and other income that are consistent with the quality signal argument. If government grants provide a positive quality signal, then the crowding-out effect would be moderated or counteracted. We predict the measured crowding out of government grants would be less than the crowding out of investment and program service income. Our findings (below) support this argument for our overall sample and for seven subsamples.[1]

Three major econometric issues arose: the appropriate choice of model, the choice and cleaning of the sample, and heteroskedasticity. In particular, a failure to correct for heteroskedasticity and to purge the sample of unduly influential (and often erroneous) observations would have led to wildly distorted results.

Choice of Model

Because of the absence of a theory that specifies the precise mathematical relation of the variables of interest, any empirical regression model must be somewhat ad hoc. We wanted parsimonious models, grounded in prior literature's findings about relevant variables, that paid some consideration to the problem of endogeneity. While simultaneous equation models and instrumental variables are two possible alternative ways of dealing with endogeneity, we chose to use ordinary least squares (OLS) models with lagged right hand variables. The advantage to employing OLS models is that we do not have to define a set of instruments (in the case of independent variables) or develop models of exogenous independent variables for each of our independent variables suspected to be endogenous (Kennedy, 2003). Because research on nonprofits is limited to a rather small set of observable data (i.e., data derived from the Form 990), and strong theoretical models for suspected endogenous variables do not exist, we believe the prudent model selection choice is OLS.

Prior literature has employed two major forms of OLS regression models: a "changes" model, which measures variables as first differences, and a "levels" model. Theoretically, changes models are less subject to omitted variable bias, allowing for more parsimonious models. However, in nonprofit data sets, first-difference models seem to be very sensitive to data errors. See Tinkelman (1999). This article reports results using both an OLS changes model and an OLS levels model.

The following changes regression model is an extension of Horne (2005):

$$\Delta C_{03} = \alpha + \beta_1\,\Delta GG_{02} + \beta_2\,\Delta PSR_{02} + \beta_3\,\Delta OR_{02} + \beta_4\,\Delta II_{02} + \beta_5\,\Delta EA_{02} + \beta_6\,\Delta FR_{02} + \beta_7\,\Delta PR_{02} + \beta_8\,\Delta PGGSQ_{02} + \beta_9\,\Delta NGGSQ_{02} + \beta_{10}\,LAGE + \varepsilon \quad (1)$$

where ΔDC_{03} = the change in direct contributions (Line 1a of Form 990) in 2003, and ΔGG_{02}, ΔPSR_{02}, ΔOR_{02}, and ΔII_{02} equal the changes in government grants (Form 990, Line 1C), program service revenue, "other revenue," and investment income (interest, dividends, rents, security sales, and other investment) in 2002. To control for other known determinants of next period direct public support (Tinkelman 1999), we include ΔEA_{02}, ΔFR_{02}, and ΔPR_{02}, which equal the change in ending assets, the change in fundraising expenses, and the change in price. The next two variables address the possibility of nonlinear effects, as described by Brooks (2000) and Borgonovi (2006). $\Delta PGGSQ_{02}$ ($\Delta NGGSQ_{02}$) is the squared change in government grants in 2002 if the grants increased (decreased). If the grant decreased (increased), this variable equals zero. LAGE is the natural log of the age of the organization, measured using the date it applied for tax exemption from the Internal Revenue Service (IRS). See Weisbrod and Dominguez (1986).

The levels regression model is,

$$DC = \beta_0 + \beta_1 GG + \beta_2 PR + \beta_3 OR + \beta_4 II + \beta_5 SQGG + \beta_6 GGNEG + \beta_7 SQGG{*}GGNEG + \beta_8 AGE + \beta_9 EA + \beta_{10} FR + \beta_{11} PRICE + \varepsilon \quad (2)$$

DC equals 2003 direct contributions. The first six right-hand variables mirror the changes model. GG, PR, OR, and II are 2002 government grants, program revenues, other revenues, and investment income. SQGG is the square of 2002 government grant income. GGNEG is an indicator variable equal to one if the change in government grants is negative. AGE is the organizational age. Additional control variables suggested by prior literature (Posnett & Sandler, 1989; Tinkelman, 1999) are the 2002 "price" (PRICE), the 2002 ending total assets (EA), and the 2002 fundraising expense (FR). The price variable measures the cost to the donor of obtaining a dollar of charitable output from one dollar of contributions. It is numerically equal to the ratio of total expenses to program expenses. As discussed below, the changes model was more sensitive to outlying observations and heteroskedasticity.[2]

Choosing and Cleaning an Appropriate Sample

To have data for both government grants (from Part I of the 990) and government contract income (from Part V) for a wide range of nonprofit organizations, we used an extract from the NCCS "digitized data." The "core data" do not have Part V data, and the Statistics of Income (SOI) time series is skewed toward large organizations. (See the NCCS website, http://nccsdataweb.urban.org, for a description of these databases and some cautions about their proper use.) We needed 3 years of data for our changes model.

We obtained data from NCCS for all operating public charities from 2001 through 2003, other than mutual benefit organizations and the separate returns of organizations that were included in group returns. NCCS provided us with 297,909 observations. We deleted 139,760 organizations because of missing data.

As we wanted to characterize various types of nonprofit organizations, the 766 organizations with National Taxonomy of Exempt Entities (NTEE) Code Z (unknown) were not helpful and were deleted. To "clean" the data, we deleted 5,546 organizations with data that were implausible or outside of normal bounds. This includes organizations with IRS registration dates before 1903 or after 2003, coded "out of scope" by NCCS, with total expenses above US$20 billion, or with negative total revenues, total expenses, or program expenses in any year. Finally, we excluded 28,523 observations reporting zero program expenses in 2001.

We then examined descriptive statistics (not tabulated here) for the resulting sample of 123,314 organizations. The wide diversity of the organizations in the sector is evident. Some high values skew the data. Median values tended to be much smaller than mean values, and standard deviations are large relative to means. Many organizations receive relatively little revenue from government grants (median = 0), other revenues (median = 0), and investment income (median = US$842). The relative reliance on different types of revenue varied widely by NCCS category, consistent with Horne (2005).

To detect outlying observations that had the potential to skew the regressions, we relied on the "leverage" statistic in STATA, which measures the degree of influence a particular observation has on the regression. We performed every regression, for the whole sample and for each of 24 NTEE categories, three times. The "first stage" used the sample of 119,113.[3] Our "second stage" model controls for heteroskedasticity. Finally, we identified observations with a leverage statistic above 0.05, deleted them, and performed a third regression ("third stage"). These third stage regressions used, for the overall sample, more than 99.9% of the original observations.

To justify the propriety of simply deleting these high-leverage observations, we obtained copies of Form 990s from "GuideStar" for 22 of the highest leverage observations in the changes regression for the overall sample, and the high-leverage observations in the diseases and medical research categories. A majority of these observations contained data errors, questionable classifications of organizations, or inconsistencies. For example, a decimal point error in the NCCS data for an organization that reported both dollars and cents one year caused large errors in the year to year changes in financial variables. In other cases, dramatic changes in financial variables were due to events other than crowding out. These observations, while not erroneous, add noise. For example, the Salt Lake Olympic Committee put on the Winter Olympics in 2002, causing 2002 to look very different from 2001 and 2003 for reasons unrelated to the hypotheses tested herein.

Heteroskedasticity

Steps to guard against this problem included dividing the sample by NTEE category and using White's (1980) correction of *t* tests for significance. In STATA, the heteroskedasticity correction is performed by including the "robust" option at the end of the regression. While some prior research has used the log form of variables in levels models to help reduce heteroskedasticity, we deemed this inappropriate for our study because negative values cannot be tested using logs and some observations of interest could legitimately have negative investment or other income.

Summary Results

OLS changes regressions were performed for the overall sample of 123,314 observations and for subsamples for 24 NTEE categories, using Equation 1. Table 1 presents summary results. For space limitations, individual variables' coefficients and *t* statistics are not presented. Results before and after addressing the influential observations and heteroskedasticity are in the first stage and third stage columns. The second stage column reports just the results of correcting for heteroskedasticity.

The first stage column suggests the various revenue sources have significant associations with donations. In the overall sample, 8 out of 10 independent variables are significant, and the average number of significant variables in the 24 different NTEE categories was 5 out of 10. The average adjusted R^2 for the 24 NTEE categories was approximately .31, with a range of .003 to .981.

The second-stage regression does not affect R^2 but shows a dramatic change in the significance of the coefficients after controlling for heteroskedasticity. From the second-stage column, the number of variables in the overall regression with statistically significant coefficients changed from eight to zero. Across the 24 NTEE categories, the models averaged two statistically significant coefficients. The total number of significant variables in all sectors fell from 131 to 47.

The third-stage regressions eliminated only a tiny fraction of the original sample's observations but had much lower explanatory power and few significant coefficients. For the overall sample, the adjusted R^2 statistic fell from .149 to .033 after deleting just 22 observations from the full sample of 119,113. Applying the asymptotic *t* tests, three of the variables in the overall regression were significant in Stage 3, even though none were in Stage 2. The average adjusted R^2 statistic for the 24 NTEE categories tested fell from .31 to .10. In 24 separate regressions by category, 10 had no significant variables and 6 had only one. The total number of significant variables by sector fell from 47 in Stage 2 to 26 in Stage 3. The recreation category is an extreme example. The first stage had 6,747 observations, adjusted R^2 of .981, and nine significant variables; the third stage, with 30 fewer observations, had adjusted R^2 of .057 and only one significant variable.

Thus, results from the changes model are sensitive to minor changes in sample. We therefore chose to concentrate our analysis on the levels model shown in Equation 2. See Table 2. The levels model was generally less sensitive to heteroskedasticity and extreme observations than the changes model and tended to have higher R^2 statistics. Deleting high-leverage observations and using heteroskedasticity-robust *t* values had

Table 1. Key Regression Result—Change Model—Government Grant Regressions

Sample category	*n* First stage	*n* Third stage	Adjusted R^2 First stage	Adjusted R^2 Third stage	Significant variable First stage	Significant variable Second stage	Significant variable Third stage
Whole	119,113	119,091	.149	.033	8	0	3
Arts	13,047	13,018	.104	.308	5	2	1
Education	15,974	15,957	.072	.082	8	1	1
Environmental	2,791	2,767	.828	.053	6	2	2
Animal-related	2,009	1,977	.319	.075	6	1	2
Health	11,291	11,263	.074	.030	7	1	0
Mental health	4,654	4,624	.303	.008	6	1	0
Diseases	2,710	2,686	.417	.016	8	7	0
Medical research	681	652	.577	.103	5	5	0
Crime and legal	2,673	2,643	.127	.019	9	1	3
Employment	2,486	2,458	.028	.138	6	0	1
Food/agriculture	1,337	1,312	.049	.024	3	0	0
Housing	9,747	9,721	.080	.017	4	0	0
Public safety	2,025	1,999	.060	.112	7	1	2
Recreation	6,747	6,717	.981	.057	9	5	1
Youth development	3,014	2,986	.200	.013	7	5	3
Human services	21,262	21,241	.857	.030	7	6	4
International	2,091	2,059	.477	.084	5	2	1
Civil rights	983	949	.720	.365	5	1	2
Community improvement	5,470	5,439	.020	.184	5	1	0
Philanthropy	304	282	.792	–.006	5	1	0
Science	656	632	.035	.363	2	0	2
Social science	305	274	.016	.190	0	0	1
Public benefit	1,164	1,131	.003	.037	1	0	0
Religion	5,692	5,669	.230	.041	5	4	0
Total for sectors					131	47	26

Note: NTEE = National Taxonomy of Exempt Entities. This table summarizes key results of regressions of the 2003 changes in direct support on prior year changes in investment income, government grants, program service income, other revenue, ending assets, fundraising expenses, and price, the log of organizational age, and quadratic terms for the changes in government grants. "First stage" results use standard *t* tests for significance at the 95% level. The "second stage" controls for heteroskedasticity by running robust standard errors. The "third stage" controls for heteroskedasticity and removes observations with leverage >.05. Categories are listed by NTEE classification letter.

a lesser, but still important, impact on results. The greater robustness and higher R^2 statistics are consistent with Tinkelman (1999).

Five out of the 11 independent variables were significant in the third stage of the overall sample levels regression, with plausible values. The fund-raising and total asset variables were strongly significant, with positive values. Price has a strongly

Table 2. Key Regression Result—Level Model—Government Grant Regressions

	n		Adjusted R^2		Significant variable		
Sample category	First stage	Third stage	First stage	Third stage	First stage	Second stage	Third stage
Whole	123,314	123,293	.667	.507	9	6	5
Arts	13,693	13,671	.530	.531	7	2	2
Education	16,625	16,598	.849	.407	9	3	4
Environmental	2,935	2,906	.947	.355	7	4	4
Animal-related	2,127	2,094	.314	.054	4	2	4
Health	11,467	11,438	.548	.310	9	4	3
Mental health	4,742	4,720	.251	.197	8	2	4
Diseases	2,802	2,781	.963	.763	10	7	3
Medical research	705	669	.736	.340	6	6	4
Crime and legal	2,742	2,710	.587	.303	9	7	6
Employment	2,540	2,502	.208	.244	5	3	3
Food and agriculture	1,390	1,351	.490	.576	3	2	3
Housing	9,928	9,901	.181	.170	5	2	1
Public safety	2,132	2,111	.453	.060	8	7	6
Recreation	7,151	7,125	.628	.302	6	3	5
Youth development	3,120	3,095	.786	.270	9	5	4
Human services	21,784	21,764	.781	.166	8	7	7
International	2,183	2,144	.792	.437	5	3	4
Civil rights	1,020	989.777	.665	5	4	8	
Community improvement	5,666	5,648	.143	.274	4	2	2
Philanthropy	320	293	.855	.191	7	7	3
Science	678	641	.369	.275	5	1	2
Social science	317	287	.473	.442	4	3	4
Public benefit	1,202	1,164	.776	.409	8	8	2
Religion	6,045	6,020	.797	.549	6	3	4
Total for sectors					157	97	92

Note: NTEE = National Taxonomy of Exempt Entities. This table summarizes key results of regressions of the direct support on prior year investment income, government grants, program service income, other revenue, ending assets, fundraising expenses, price, organizational age, a quadratic term for government grants, an indicator variable for whether the change in government grants is negative, and the interaction of the quadratic term and the indicator variable. "First stage" results use standard *t* tests for significance at the 95% level. The "second stage" controls for heteroskedasticity by running robust standard errors. Categories are listed by NTEE classification letter.

significant negative value. Program revenues had a negative coefficient (–.01), suggesting some crowding out. Government grants had a statistically positive coefficient of .07, suggesting some crowding in. This is somewhat surprising, as Borgonovi (2006) suggests that high levels of government grants would cause greater crowd out. The coefficients on other revenue and investment income were not statistically

Table 3. Significant Coefficients From Levels Regressions

Category (*n*)	Government grants	Program revenue	Investment income
All categories combined (123,293)	0.07	–0.01	
Environmental (2,906)		–0.14	
Health (11,438)	0.17		
Mental health (4,720)	–0.02	–0.03	
Diseases (2,781)	–0.31		
Medical research (669)	0.24		
Crime and legal (2,710)		–0.09	
Employment (2,502)		–0.04	
Public safety (2,111)	–0.22	–0.06	
Human services (21,764)		–0.02	
Civil rights (989)	–0.30	–0.40	
Philanthropy (293)		–0.37	
Science (641)		–0.01	
Social science (287)		–0.11	–10.57
Range of category values	–0.31 to 0.24	–0.40 to –0.01	

Note: This table reports (where significant) the coefficients from the model results reported in Table 2 for government grants, program services, and investment income. Values are only reported when the regression coefficients were significantly different from zero at 95% significance levels.

significant. The tests on restrictions indicate that the linear combination of the coefficient on governmental grant income and the square of governmental grant income are statistically different from the coefficient on program revenue (at the 1% level) but not statistically different from the coefficient on investment income.

The regressions were repeated on the 24 NTEE categories, with generally similar results. Government grants had significant positive coefficients twice and significant negative ones 4 times. Program revenues were significantly negative 10 times. Investment income was significantly negative once. The tests on restrictions indicate that the linear combination of the coefficient on governmental grant income and the square of governmental grant income is statistically different from the coefficient on program revenue in 6 out of the 24 cases (at the 5% level) but only statistically different from the coefficient on investment income in 2 out of the 24 cases (at the 5% level). In 4 of the 6 significant cases comparing the coefficients on government grant income and program revenue, and in both significant cases comparing the coefficients on government grant income and investment income, the coefficient on government grants is positive indicating a quality signal.

We report the significant coefficients from the levels regression in Table 3. The derivative of donations with respect to government grants also includes a term related to the squared government grants. However, the regression coefficients are so small that this effect was immaterial, and the coefficient for government grant provides a measure of the rate of change of donations for each dollar change in government grants.

The results suggest that on an overall basis, government grants have a positive impact on donations but that program revenues have a negative one, with effects that vary by sector. This finding is interesting, suggesting some evidence of crowding in (government grants) and crowding out (program revenue). For the overall sample, the coefficients on government grants and program revenues are .07 and –.01. For the NTEE categories with significant coefficients, the values fell within the range of –.31 to +.24. The crowding-out effect of government grants for the public safety, civil rights, and diseases categories was at economically significant levels, ranging from –.22 to –.31, while the coefficient for mental health was only –.02. The crowding in for health organizations was .17 and for medical research it was .24. The crowding-out effect of program revenue ranged from –.01 to –.40 across the 10 sectors with statistically significant coefficients.

A somewhat surprising result is the insignificant coefficient on investment income. In fact, the coefficient on investment income was only significant for the social science sector. The significant result for the social science disappears if one influential outlying observation is deleted.

Discussion

This study finds mixed evidence about the crowding-out effects of alternative revenue sources, using a levels model, and the tests performed do not find an effect on crowding out due to investment income.

The article highlights certain econometric issues to future researchers. Researchers should employ standard measures to control for heteroskedasticity. Changes models using NCCS data may be less robust than levels models. Results, even for large samples, were strongly affected by a small number of outlying observations, which were often data errors.

Authors' Note

We gratefully acknowledge comments from participants at ARNOVA, assistance with data from the National Center for Charitable Statistics, a summer research grant from Hofstra University, and research assistance from Jessie Xiao Chen. In addition, we would like to thank the anonymous referee for useful comments during the review process.

Declaration of Conflicting Interests

The author(s) declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding

The author(s) received no financial support for the research and/or authorship of this article.

Notes

1. In one subsample, diseases, the coefficient on government grants is actually a larger negative number than the coefficient on program revenue or investment income.
2. We also investigated the effects of two other models but do not present the results here. One involved the changes of the logged values between years. The results were qualitatively similar to third stage results using the changes model. However, the model adjusted R^2 were noticeably lower for the log changes model. We also employed a levels model that used logs of the variables. The impact of heteroskedasticity and influential observations on that model is qualitatively similar to the results presented herein for the levels model. While logged levels models have been used in the past in the literature (e.g. Weisbrod & Dominguez 1986), using a log model here would force us to exclude observations where investment and other income are negative. As our object is to test revenue interactions, such observations might be important.
3. The actual number of observations used in the change analysis was 119,113. The difference in total observations from the change analysis and levels analysis reported in Table 2 are due to missing 2001 observations.

References

Borgonovi, F. (2006). Do public grants to American theatres crowd-out private donations? *Public Choice, 126*, 429-451.

Brooks, A. C. (2000). Public subsidies and charitable giving: Crowding out, crowding in, or both? *Journal of Policy Analysis and Management, 19*, 451-464

Horne, C. S. (2005). Toward an understanding of the revenue of nonprofit organizations. Unpublished doctoral dissertation, Georgia State University and Georgia Institute of Technology, Atlanta.

Kennedy, P. (2003). *A guide to econometrics* (5th ed.). Cambridge, MA: The MIT Press

Posnett, J., & Sandler, T. (1989). Demand for charity donations in private non-profit markets; The case of the U.K. *Journal of Public Economics, 40*, 187-200.

Steinberg, R. (1993). Does government spending crowd out donations? Interpreting the evidence. In A. Ben-Ner & B. Gui (Eds.), *The nonprofit sector in the mixed economy* (pp. 99-125). Ann Arbor: University of Michigan Press.

Tinkelman, D. (1999). Factors affecting the relation between donations to not-for-profit organizations and an efficiency ratio. *Research in Government and Nonprofit Accounting, 10*, 135-161.

Tinkelman, D. (2010). Crowding out: How do revenue sources interact? In B. A. Seaman & D. R. Young (Eds.), *Nonprofit economics and management: The state of research*. Cheltenham, UK: Edward Elgar.

Vesterlund, L. (2006). Why do people give? In W. W. Powell & R. Steinberg (Eds.), *The non-profit sector: A research handbook* (2nd ed., pp. 591-611). New Haven, CT: Yale University Press.

Weisbrod, B. A., & Dominguez, N. D. (1986). Demand for collective goods in private non-profit markets: Can fund-raising expenditures help overcome free-rider behavior? *Journal of Public Economics, 30*, 83-95.

White, H. (1980).A heteroskedasticity-consistent covariance matrix estimator and a Direct test for heteroskedasticity. *Econometrica, 48*, 817-838.

Bios

Daniel Tinkelman is a certified public accountant with a doctorate in accounting who has written numerous articles on accounting issues related to nonprofit organizations.

Daniel G. Neely is a certified public accountant who recently completed his doctorate in accounting. His current research interests concern accounting issues related to nonprofit organizations.