Balance Sheet Insolvency and Contribution Revenue in Public Charities

Tatiana Homonoff Thomas Luke Spreen Travis St. Clair*

April 25, 2019

Abstract

Using Form 990 data reported by public charities, we document significant bunching of nonprofits at near-zero net assets, the threshold for insolvency. Bunching occurs despite the fact that creditors cannot force nonprofit organizations into involuntary bankruptcy in the same manner as for-profits. We show that the extent of bunching is greater among organizations that rely more heavily on contribution revenue, and that by inflating their net assets, bunching organizations are able to increase their contribution revenue relative to firms that report negative net assets. Charitable donors appear to use the net assets threshold as heuristic for a charity's financial health; nonprofit managers, in turn, respond to the preferences of their donors.

JEL Codes: L31, L39, H31

Keywords: bunching, charitable giving, nonprofit finance

*Homonoff: New York University, Robert F. Wagner School of Public Service. Spreen: University of Maryland, School of Public Policy. St. Clair: New York University, Robert F. Wagner School of Public Service. We gratefully acknowledge seminar and conference participants for conversations and suggestions that have greatly improved the quality of this project. All remaining errors are our own.

I. Introduction

One of the key features that distinguishes for-profit and nonprofit entities in the United States is how federal law treats them in insolvency. The bankruptcy code permits creditors of for-profit corporations to file involuntary bankruptcy petitions, and consequently the typical outcome for corporations that are balance sheet insolvent (i.e., that report liabilities in excess of assets) is either bankruptcy or restructuring. This is not the case for nonprofit firms, whose creditors cannot initiate involuntary bankruptcy proceedings. Instead, the decision to liquidate or re-structure lies with a nonprofit organization's board of directors. Nevertheless, nonprofits that are "insolvent on the books" may still face incentives to report otherwise if doing so helps to preserve their reputation with important stakeholders such as donors. For the average public charity, donor contributions are the primary source of revenue; Internal Revenue Service Form 990 data show the median nonprofit receives 53 percent of its revenue from contributions. Consequently, how donors allocate their charitable contributions has significant implications for the strategic decisions of the nonprofit organizations. If donors use balance sheet solvency as a heuristic for assessing the financial health and viability of organizations (rather than the underlying value of net assets) when deciding where to allocate their charitable dollars, insolvent organizations may have an incentive to report financials that place them just above the insolvency threshold.

We use data from the National Center for Charitable Statistics' Core Financial Files which include all 501c(3) public charities in the United States – a panel of nearly half a million public charities spanning 2005 to 2015. Using financial information reported in the Internal Revenue Service's Form 990 or 990-EZ, we examine whether charities manipulate their income to avoid insolvency and explore the consequences of this behavior. First, we document that a considerable number of charities are balance sheet insolvent. Approximately seven percent of charities report negative net assets, or more than 22,000 charities in 2015 alone. Second, we document significant bunching at zero net assets. We examine several explanations for this result and find that the bunching is not solely a product of organization age, equity transfers between organizations, or a desire to spend down resources, consistent with a model in which a substantial portion of the bunching we observe is due to income manipulation rather than any quirk of nonprofit financial reporting.

To quantify the magnitude of the behavioral response, we estimate the size of the excess mass just above the insolvency threshold using methods from the bunching literature (Chetty et al., 2011; Dee et al., Forthcoming; Diamond and Persson, 2016; Kleven and Waseem, 2013). The size of the excess mass suggests that about one third of a percent of all public charities engage in income manipulation at the threshold and further that there is a 27 percent chance that charities falling just below the threshold inflate their net assets so as to appear solvent. For context, this estimate is roughly three times the size of the behavioral response observed in response to the requirement that nonprofits file the Form 990 (Marx, 2018).

Next, we explore the characteristics of bunching organizations using methods described in Diamond and Persson (2016). We find that bunching is most common among smaller organizations that receive a large proportion of their revenues from charitable contributions rather than from program fees. We find that our bunching estimates increase monotonically with the percent of revenues from contributions and that the extent of bunching is approximately seven times greater for organizations in the highest quartile of contribution revenue than for firms in the lowest quartile. These findings are consistent with a model in which nonprofits are motivated to manipulate their financial reporting in order to appear balance sheet solvent so as to appeal to their donor base.

Finally, we exploit the panel structure of our data to examine the consequences of bunching by comparing outcomes in later years for charities that manipulated their financial reporting to the outcomes of charities that were eligible to bunch but did not. Motivated by our prior findings, we first consider the effect of bunching on contribution revenue. We find no evidence of an effect of bunching on contribution revenue in the year that an organization bunches, consistent with the timing of the release of financial statements. However, we observe that bunching leads to an increase in contribution revenue two years later – charities near the threshold in year t experience a significant 19 percent increase in contribution revenue in year $t + 2$. In contrast, we observe no effect on expenses in subsequent years. However, we do observe an effect in year t, the year the firm bunches, suggesting that part of the mechanism by which firms engage in bunching involves decreasing expenses. Taken together, these findings suggest that donors are more likely to contribute to charities just above the insolvency threshold, and that firms manage to manipulate their income to appear solvent by reporting lower expenses.

Our analysis contributes to several strands of the literature on charitable giving and the private provision of public goods. First, we document new features of donor preferences and the allocation of charitable giving. Our findings are consistent with previous work showing that donors consider the financial health of the organizations to which they contribute. Prior research indicates that donors are sensitive to the amount of cash charities have on hand (Calabrese, 2011), their degree of leverage (Calabrese and Grizzle, 2012), and the amount of program revenue they earn (Okten and Weisbrod, 2000). Moreover, there is also clear evidence that nonprofits shift their behavior in response to these donor preferences (Calabrese, 2013; Krishnan and Yetman, 2011).

We also contribute to the behavioral economics literature on decision-making heuristics in charitable giving. For example, Karlan and List (2007) suggest that donors use the presence of a donation match as a heuristic for the price of giving: donors are more likely to contribute when there is a match, but are insensitive to the size of the match. Yoruk (2016) documents a jump in contributions for charities just above the threshold for receiving an additional star on Charity Navigator, a third-party rating agency, relative to charities just below the threshold, suggesting that donors respond to simplified benchmarks of financial health. In our context, donors may regard balance sheet solvency as a heuristic for financial health, which they use to simplify decision-making on how to best allocate scarce donative resources.

Finally, we build on a number of recent papers that have used bunching methods to study income manipulation among nonprofit firms. Marx (2018) shows that the average charity reduces reported income by $750-$1000 in order to avoid filing the full version of the Form 990, the information return required by the IRS of all tax-exempt organizations. St. Clair (2016) demonstrates how charities manipulate their revenues to avoid state audit requirements. While this paper also uses bunching methods to understand the strategic responses of firms, it differs from these other papers in that the observed bunching is not driven by a regulatory requirement, but rather by preferences of key stakeholders.

This paper proceeds as follows. Section 2 provides background on insolvency and nonprofit finance. Section 3 describes the dataset and provides summary statistics. Section 4 presents graphical evidence of bunching at zero net assets and quantifies the size of the behavioral response. Section 5 examines the characteristics of bunching charities and provides motivation for the empirical analyses. Section 6 evaluates the effects of bunching on charitable contributions as well as other financial metrics. Section 7 concludes.

II. Background on Nonprofit Finance and Insolvency

An entity is generally recognized as insolvent when it is unable to meet its outstanding obligations in full and on time. There are two versions of insolvency: cash flow insolvency and balance sheet insolvency. Cash flow insolvency occurs when an organization is unable to meet its near-term obligations due to liquidity constraints. Balance sheet insolvency occurs when a firm's total liabilities exceed its total assets; the organization is "insolvent on the books" even if it is able to service its debts in the near term.[1]

In this paper, we focus on balance sheet insolvency in order to highlight the difference in treatment between for-profit and nonprofit firms under federal law. While solvent, the fiduciary duty of the directors of for-profit companies is to serve the interest of their company's shareholders, who own the residual (surplus) value of assets over liabilities.[2] As a

[1] Financial statements are typically prepared under the assumption that the reporting entity will continue to operate as a going concern. If liquidation is imminent, the generally accepted accounting principles require that financial statements be prepared under the liquidation basis of accounting (FASB, 2014).

[2] A company's residual value is captured in the shareholders' equity account on the balance sheet. Net

for-profit organization approaches insolvency, the residual interest in the company expands to include its creditors, who gain standing to file an involuntary bankruptcy petition under the United States Bankruptcy Code.[3] A company may also voluntarily discharge their debts through liquidation under Chapter 7 of the United States Bankruptcy Code or reorganize under Chapter 11. There are also separate considerations beyond net worth that contribute to the decision to enter bankruptcy, including bankruptcy costs, tax considerations, and the ranking of interest in distributing the company's liquidated value (Bulow and Shoven, 1978; White, 1989).

While insolvent nonprofit organizations may also voluntarily file for bankruptcy protection, nonprofit creditors are ineligible to force an insolvent nonprofit into bankruptcy. While solvent, the fiduciary duty of nonprofit directors is to fulfill its chartered mission. Similar to for-profits, this duty expands to include the interests of the organization's creditors once the nonprofit approaches insolvency. Regardless, federal law does not permit a nonprofit's creditors to sue nonprofit directors, and U.S. courts have not recognized those creditors as holding a residual interest in the organization (Elliot and Hollander, 2014). While the directors of an insolvent nonprofit organizations are advised to take their creditors' interests into account, their legal duty is to fulfill the organization's mission, even if that comes at their creditors' expense (Peterman and Morissette, 2004).

If the insolvency threshold does not have legal ramifications for the nonprofit sector, then what incentives do organizations have to maintain positive net assets? Despite a lack of owners to lay claim to surpluses (the "nondistribution constraint") and research that finds that nonprofits are not revenue-maximizing (Okten and Weisbrod, 2000), nonprofit managers may wish to preserve the long-term viability of the organization so as to collect salaries and continue the firm's mission. This viability may be threatened if donors or other external parties attach negative consequences to the reporting of negative net assets.

assets are the nonprofit equivalent of shareholders' equity, as nonprofits do not have owners.

[3]The zone of insolvency is a legal concept that does not have a precise definition, but generally, organizations that are either balance sheet or cash flow insolvent fall in this region.

The ability of donors and external parties to observe the financial position of nonprofit organizations has grown in recent years with the ubiquity of third-party rating agencies such as Charity Navigator and GuideStar. These sites compile and report on the financial position of nonprofit organizations based on information extracted from their annual Form 990 information returns filed with the Internal Revenue Service. Indeed, in its "Pro Reports," GuideStar reports specifically on whether an organization has reported negative net assets in the last five years. Thus, even for unsophisticated donors with little financial knowledge, information on the insolvency threshold is available and potentially salient.

III. Data and Summary Statistics

Our data source is the National Center for Charitable Statistics' (NCCS) 2005-2015 Core Financial Files for public charities, which are based on the IRS' annual Return Transaction Files. The public charities core files contain approximately 50 financial variables for all 501c(3) public charities reporting at least $50,000 in gross receipts that filed either the Form 990 or the Form 990 EZ.[4] The public charities files contain data only on 501c(3) public charities, and consequently our analysis does not include private foundations or exempt organizations that are not 501c(3)s. The data contain information on 578,282 charities for a total of just under 3.5 million annual returns.

A. Graphical Evidence of Bunching and Sample Selection

Figure 1a presents a density plot of public charities, with the x-axis showing net assets scaled by total assets and the y-axis showing the number of organization-years.[5] The figure shows substantial bunching just above the insolvency threshold. This bunching occurs despite the

[4] The Form 990 is an information return required by the IRS of all tax-exempt organizations. Organizations with gross receipts of less than $200,000 and total assets of less than $500,000 can file the Form 990 EZ, a simpler version of the form. Organizations with gross receipts of less than $50,000 can file the Form 990-N (e-Postcard).

[5] Scaling by total net assets enables us to examine organizations of disparate size. For the remainder of the paper we use scaled 'net assets' as a shorthand for net assets as a share of total assets.

fact that there is no discontinuity in policy such as the requirement for organizations above a certain threshold to be audited, as in other analyses of bunching in the nonprofit sector (St. Clair, 2016; Marx, 2018).

While we are unaware of any policies that change for nonprofits across the insolvency threshold, there may be several mechanical reasons for observing a large number of firms reporting exactly zero net assets. Therefore, to investigate whether we observe bunching that is plausibly due to income manipulation, we make several sample restrictions. Appendix Table 1 summarizes the restrictions and their effect on the size of the sample. These restrictions are most likely conservative and, if anything, bias our bunching estimates downwards.

First, new charities are unlikely to accumulate significant net assets in their first year of operation. To ensure that bunching is not merely a feature of "new" charities, we exclude organizations that have been in operation for less than ten years, where age is based on the year in which the IRS recognized the organization's tax exempt status.[6]

Next, we allow for the possibility that some tax-exempt organizations may seek to spend down their resources in every period. We exclude organizations with zero assets that report an average net income of zero, or that never report any contribution revenue. We also remove organizations that begin the year with zero net assets, ensuring that our sample does not include organizations that maintain just enough assets to service their debts. This restriction also has the indirect effect of removing firms that persistently bunch at the threshold, and consequently the bunching that we document in our final sample is not due to repeated bunching among a relatively small group of charities.

Finally, we consider whether bunching may be driven by subsidiary organizations that transfer all of their equity at year's end to a parent organization. We do find cases of charities who provide specialized services in the form of fund-raising or investment management for a closely related organization and who frequently transfer net assets. These transfers are

[6] Appendix Figure A1 shows density distributions for charities that have been in operation for different lengths of time, ranging from 2 to 24 years. While the extent of bunching declines somewhat with age, bunching persists even among organizations that have been in existence for decades.

reported on the 990 as "other changes in net assets," and consequently we remove all charities that report a nonzero value in this field. These restrictions leave us with a final sample of 1,187,838 observations and 268,184 distinct charities.

Figure 1b repeats the analysis in Figure 1a for our final sample. While the extent of bunching declines somewhat due to these restrictions, as expected, we still observe substantial bunching just above the insolvency threshold. This suggests that the bunching we observe in our final sample may likely be due to charities manipulating their net assets rather than any structural features of nonprofit financial reporting.

B. Summary Statistics

Table 1 reports summary statistics for the final sample. The financial variables are highly skewed, and consequently for the analyses in sections 5 and 6 we apply log transformations to the outcome variables of interest. There are two features of the data that are worth highlighting. First, a very large fraction of public charities have either no or very low liabilities. The median charity in our sample has a net assets to assets ratio of 0.99, implying that liabilities are only one percent of assets; the mean is 0.78. While only about 60% of the organizations in the sample earn program revenue – revenue earned through the provision of goods or services to clients – those that do engage in revenue-generating activities collect more revenue on average from their programs than from contributions. The median charity in the final sample collects 48% of their revenues from contributions. Second, approximately five percent of the charities in our analytic sample – and seven percent in the raw data – report liabilities in excess of assets, i.e. are balance sheet insolvent.[7] The relatively large proportion of insolvent firms highlights that the insolvency threshold has different consequences for nonprofits than for for-profits.

[7] Bowman (2011) draws a distinction between for-profit and nonprofit balance sheet insolvency, arguing that nonprofits are balance sheet insolvent when their unrestricted net assets, rather than total net assets, drop below zero, since organizations with negative unrestricted net assets will be unable to discharge their obligations to their creditors. However, we focus on total net assets because it is more salient to users of financial statements and also because it is among the financial metrics highlighted by GuideStar.

IV. Bunching at Zero Net Assets

A. Measuring the Excess Mass

We employ standard methods for measuring the extent of bunching and the size of the behavioral response. The bunching design was first introduced by Saez (2010) and further developed by Chetty et al. (2011) and Kleven and Waseem (2013) to identify tax-induced behavior distortions using kink points in tax schedules. Although initially developed to study the elasticity of taxable income, it has since been employed to study behavioral responses in other contexts, including among small businesses (Onji, 2009) and nonprofits (St. Clair, 2016; Marx, 2018).

The basic bunching design divides the running variable into bins and counts the number of observations within each bin. The number of excess bins on one side of the threshold is then compared to a counterfactual distribution in which no bunching occurs, with the identifying assumption being that the counterfactual distribution is smooth across the threshold. Borrowing the notation of Kleven (2016), we estimate bunching as follows:

$$c_j = \sum_{i=0}^{p} \beta_i \cdot (z_j)^i + \sum_{i=z_-}^{z^+} \gamma_j \cdot 1[z_j = i] + v_j \tag{1}$$

where c_j represents the number of organizations in bin j and z represents the level of scaled net assets (net assets / total assets) in bin j. The left-hand side of the equation represents the counterfactual, estimated as a polynomial function that expresses the association between the organization count and net assets, with p as the degree of the polynomial. We use a fifth order polynomial based on the sharp drop that occurs in the Akaike Information Criterion (AIC) between orders four and five (see Appendix Table 2 for more information on the different polynomial choices). The right-hand side measures the extent of bunching by estimating the difference in the bin counts around the threshold (between z- and z+) relative to the counterfactual, obtained using a series of dummy variables for bins z- through z+. Bins z- to 0 represent the region of missing mass below the threshold, while bins 0 to z+ represent the region of excess mass.

Figure 2 shows a density plot of public charities along with a fitted counterfactual distribution as described above. As in Figure 1, we observe significant bunching to one side of the zero net assets threshold. The asymmetric density around the threshold is consistent with empirical distributions associated with "notches," where agents face a discontinuous jump in their choice set and a region of strictly dominated choice, rather than "kinks," where the bunching can be symmetric around a threshold (Saez, 2010; Chetty and Saez, 2013). As noted previously, the bunching occurs despite the fact that there is no discontinuity in policy.

While much of the earlier bunching literature selected the manipulated or "excluded" region based on visual inspection (Saez, 2010; Chetty et al., 2011; Kleven and Waseem, 2013), Diamond and Persson (2016) develop a fully automatic estimator that does not require manually choosing the parameters. Most papers identify the width of the manipulated window based on an "integration constraint;" that is, they set the area under the counterfactual distribution equal to the area under the empirical distribution, which implies that any excess mass should equal the missing mass on the other side of the cutoff (Chetty et al., 2011; Diamond and Persson, 2016; Dee et al., Forthcoming).

In our case, there is a high degree of attrition near the notch due to the poor financial health of the charities in the region of the data we examine. Because attrition increases as organizations decline in net assets and approach insolvency, it is unlikely that the missing mass on the left side of our cutoff will be the same magnitude as the excess mass. As a result we report our bunching estimates using the size of the excess mass alone rather than as an average of the excess and missing mass and we also report our estimates using a range of choices for the manipulated range.

Table 2 reports the estimates of the size of the excess mass, and thus the extent of bunching. Under the main specification, we choose a region of -0.16 to 0.08 net assets based on visual inspection of Figure 2 with an estimation range of net assets between -

0.5 and 0.5.[8] However, we also show that our bunching estimates are fairly insensitive to the choice of window. Following Dee et al. (Forthcoming), we provide estimates for both total manipulation and in-range manipulation. Total manipulation is the excess mass as a percent of the total sample size, which corresponds to the percentage of total charities in the sample that bunch. In-range manipulation is the excess mass as a percent of the number of charities in the counterfactual range in the region of missing mass (bins -z to 0), which can be interpreted as the probability of bunching conditional on falling just below the solvency threshold. We calculate standard errors using a parametric bootstrap procedure, similar to the one used in Dee et al. (Forthcoming) and Chetty et al. (2011). We draw with replacement from the distribution of residuals estimated in Equation 1 to generate a new density distribution from which we can generate bootstrapped estimates of the excess mass. The standard error we report is the standard deviation of 200 of the bootstrapped estimates.

Column 1 presents our total manipulation estimate and shows that the number of excess organizations above the threshold represents approximately 0.33 percent of all public charities. This is equivalent to a 27 percent probability that charities falling just below the threshold will manipulate their financial reporting so as to appear solvent (column 2). By comparison, Marx (2018) finds that the number of excess organizations above the filing threshold for the Form 990 is equal to 0.1 percent, or approximately one third the size of our estimate. Columns 3 and 4 show that changing the manipulation window has very little effect on the estimated size of the excess mass–the total manipulation estimates range from 0.32 to 0.34 percent of all charities.

B. Size of the Behavioral Response

In addition to understanding the causes and consequences of bunching, our measurement of the excess mass enables us to estimate another behavioral parameter of interest: the extent of avoidance behavior demonstrated by charities at the notch. By measuring how far the

[8] The fact that our missing mass extends further into the distribution than the excess mass appears to be common in empirical distributions around notches (Kleven, 2016)

excess mass can be distributed into the counterfactual density distribution below the notch, we can estimate δ, the distance that the average buncher "traveled" to move above the notch.

$$\delta = \frac{\left(\sum_{i=0}^{z^+} \gamma_j \cdot 1[z_j = i]\right) \cdot \rho}{f(0)} \tag{2}$$

Specifically, we multiply the number of excess organizations that we obtain from our preferred specification (column 1, table 3) by ρ, the size of the bins (0.02 in our specification), and divide this by the height of the counterfactual density distribution at the notch, $f(0)$. This follows the practice in other bunching studies of assuming that because the density is not very steep at the threshold, the counterfactual density distribution is approximately flat in a narrow range around the notch. Equation 2 yields an estimate of 0.034, suggesting that the average buncher inflates their net assets by an amount equivalent to three percent of their assets.

V. Who are the Bunchers?

In this section, we turn away from estimating the extent of manipulation and focus instead on characterizing the bunchers in our sample. We seek to understand why certain types of charities might be motivated to bunch and what distinguishes the bunchers from charities that are otherwise in similar financial health. In the next section, we explore the downstream consequences of bunching.

We follow the general approach of Diamond and Persson (2016) with only slightly adjustments due to the attrition in our sample. This involves comparing the mean characteristics of the bunchers to those charities that fall just below the insolvency threshold and might have chosen to bunch but did not. We first estimate a counterfactual by fitting second order polynomials separately to both sides of the manipulation region, i.e., re-estimating Equation 1 for the outcome of interest rather than for the density distribution. This allows us to calculate the average values for the bins above the notch in the region of excess mass (Y^{up_all}) as well as the average values for the bins in the empirical distribution in the region of the missing mass (Y^{down}). Y^{up_all} represents the average values of a combined group consisting of the "compliers" (i.e., the bunchers), and the "always-takers" (i.e., those charities that would have reported net assets just above the threshold even in the absence of income manipulation).

Once we have calculated the average characteristics for both the empirical distribution and the counterfactual by multiplying the estimated outcome for each bin by the number of charities falling into that bin, we can distinguish the characteristics of the compliers from the characteristics of the always-takers by using the counterfactual to determine the average characteristics of the always-takers:

$$\bar{Y}_{compliers} = \frac{N_{up}^{tot}}{N_{up}^{tot} - N_{up}} * \bar{Y}^{up\_all} - \frac{N_{up}}{N_{up}^{tot} - N_{up}} * \bar{Y}^{up} \tag{3}$$

where N_{up}^{tot} represents the total number of charities above the notch in the empirical distribution, N_{up} represents the number of charities above the notch in the counterfactual, $\bar{Y}^{up\_all}$ represents the mean value of charities in the region of excess mass in the empirical distribution, and $\bar{Y}^{up}$ represents the mean value of charities in the region of excess mass in the counterfactual. The region of excess mass extends from net assets of 0 to 0.08 in our preferred specification. We calculate standard errors using the same parametric approach in the previous section, except that each simulation now includes a multi-step procedure: estimating the frequency counts as well as the outcomes of interest for the empirical and counterfactual distributions.

As noted above, we must make one adjustment. The size of our excess mass does not equal our missing mass due to the fairly high degree of attrition in the region of the data we examine; charities increasingly discontinue operations and leave the sample as they approach insolvency. The approach that Diamond and Persson (2016) outline implicitly leverages the integration constraint, the fact that a fixed number of bunchers leave the region of the missing mass and cross the threshold to fall in the region of excess mass. To account for the fact that the integration constraint is not satisfied in our context, we make the simplifying assumption that the additional "missing mass" that is unaccounted for comes evenly from the manipulation region below the notch. That is, we adjust the density distribution below the cut-off so that the missing mass on one side equals the excess mass on the other. This enables us to more accurately estimate the differences between charities on one side of the threshold to the other. However, the results may be slightly biased if the charities that leave the sample would instead fall in different regions of the distribution.

Table 3 reports our results. In the first five rows, we first report the types of charities that bunch. In the bottom seven rows, we examine their financial characteristics. The nature of the charity is of course pre-determined. This is not the case with the financial metrics; as we show in the following section, bunching does affect the financial attributes of charities. Nevertheless, we summarize the characteristics of the bunchers – both pre-determined and potentially endogenous – to motivate further analysis. To limit the potential endogeneity of the financial characteristics, we choose attributes that are fairly static over our time window. This does not foreclose the possibility that bunching may have influenced some of the outcomes, but it does limit some of the more obvious concerns that would accompany income or expense measures that demonstrate a high degree of intertemporal variability.

Column 1 presents the characteristics of the bunchers. Column 2 presents the mean characteristics of the charities falling just below the threshold, and column 3 presents the difference. The charities that bunch are less likely to be human-services charities (such as Boys & Girls clubs, YMCA's, and the Boy Scouts of America) and more likely to be charities in the "other" category (including environmental, international, public benefit, and religious organizations). The bunching charities also appear to be significantly smaller than the charities that do not bunch, with substantially fewer assets and a lower probability of holding a mortgage. We also find that compensation of employees represents a higher proportion of expenses for bunching charities.

One characteristic of note is that bunching charities receive a higher proportion of their revenue from donor contributions. This is consistent with the nature of the charities in the "other" category; environmental and international organizations, unlike for example non-profit hospitals, typically rely on donor contributions to support their activities rather than program revenue (e.g., client fees in exchange for services). To further probe the association between bunching and reliance on contribution revenue, Figure 3 splits the sample into four quartiles according to the percentage of revenue that charities receive from donations and replicates the density distributions from Figure 1 separately for each quartile. While we observe graphical evidence of bunching in all four quartiles, the extent of bunching appears more pronounced among charities in the fourth quartile.

Table 4 provides the corresponding in-range manipulation estimates by contribution revenue quartile. Overall, we find that the extent of bunching grows monotonically with the reliance on contribution revenue. Among charities in the bottom two quartiles, we find an 11 to 12 percent probability that charities falling just below the threshold manipulate their financial reporting in order to cross the insolvency threshold. This estimate increases to 20 percent for charities in the third quartile. However, for charities in the top quartile, we observe a substantial increase in the extent of bunching – over 80 percent of charities falling just below the threshold manipulate their income to appear solvent.

VI. Effect of Bunching

Based on our findings in the previous section, we investigate the effects of bunching. That is, we examine the outcomes for public charities that manipulated their financial reporting so as to remain balance sheet solvent and compare these outcomes to charities that were eligible to bunch but did not.

We start by once again using the general framework outlined by Diamond and Persson (2016). We construct an estimate of what outcomes would have looked like for charities in the manipulation region absent income manipulation. We then compare this counterfactual to the actual distribution of outcomes for charities in the same range. The difference between these two estimates represents the reduced form effect, an intent-to-treat (ITT) estimate of the effect of falling in the manipulation region. We then scale this effect by the "first stage," the probability of being a buncher, which we previously calculated and reported in Table 2 as the estimates of in-range manipulation. This constitutes the local average treatment effect (LATE) of bunching. The difference between these two estimates stems from separating the characteristics of the always-takers from the compliers. As in the previous section, we estimate the counterfactual by regressing the outcome of interest on second order polynomials, which we estimate separately on both sides of the cut-off, thus allowing for the possibility of a discontinuous jump in the outcome at the threshold.[9]

One challenge we face is that the financial outcomes we are interested in studying are mechanically related to net assets; by definition, bunching organizations "inflate" their net assets by somehow increasing their reported revenue, decreasing their reported expenses, or some combination. If we focus on outcomes that are themselves subject to manipulation, the estimates we calculate may not be the effects of bunching but rather the cause. However, the panel nature of the data gives us some insight into which variables are subject to manipulation as well as which variables are affected by manipulation. Specifically, we assume that it is not possible for bunching to have any causal effect in year t, the year in which manipulation occurs. Therefore, any "effect" that we observe in that year may be part of the mechanism by which charities inflate their net asset position. In contrast, if we do not observe any distortion in the outcome distribution in year t, but observe effects in subsequent years, this would suggest that the outcome in question was not subject to manipulation, but may instead have been affected by the decision to bunch. In effect, by examining outcomes across a variety of years, we can use our estimates from year t (the year in which bunching occurs) as a benchmark against which to compare the effects of bunching in future years.[10] This

[9] Our choice of second order polynomials is based on the specification that minimizes the AIC.
[10] This would seem to suggest event study methods as an alternative empirical strategy. However, it is not

approach has the further advantage of mitigating concerns about differential attrition; the outcomes we focus on are conditional on charities remaining in the sample.

We examine several financial variables as outcomes. First, motivated by our descriptive findings in the previous section, we examine the effects of bunching on contribution revenue. Consistent with the panel approach discussed above, if charities appear solvent by inflating their reported contribution revenue, we should see an increase in contribution revenue across the threshold in year t. In contrast, if contribution revenue is not the source of income manipulation but rather donors are more likely to donate to solvent charities, we would expect to see an increase in contribution revenue in the years following the decision to bunch, but not in the year in which bunching occurs. Our assumptions regarding the timing of donor response are based on the dynamics of the release of financial information to the public. Specifically, financial statements and Form 990s are not completed and released until several months after the fiscal year-end. Moreover, websites that disseminate information regarding the financial health of charities, such as Charity Navigator or GuideStar, do not obtain or publish this information until many months after financial statements are released. It follows that if bunching has an effect on contribution revenue, it would not occur for at least one or more years later.

Tables 5a and 5b provide estimates of the impact of falling in the manipulation region (ITT) and the impact of bunching (LATE) on contribution revenue in years t, $t+1$, and $t+2$ using the methodology described above. Column 1 shows that in the year that bunching occurs, the estimated impact of falling just above the threshold is small and not statistically significant, suggesting that charities are not manipulating contribution revenue as a means of appearing solvent. In fact, the point estimate is negative, whereas an income manipulation hypothesis predicts a positive coefficient. In contrast, we observe positive estimates in the two years after bunching occurs. Specifically, organizations that fall in the manipulation region in year t receive 17 log points (or 19 percent) more in contribution revenue in year

feasible to precisely identify which specific public charities are bunching at any point in time.

t+2. When the reduced form effect is scaled by the probability of bunching, the LATE indicates that bunching organizations increase their contribution revenue by 66 log points (93 percent). Although not significant, the coefficient on contribution revenues in year *t+1* indicates that bunching may have increased contributions in the year following the decision to bunch as well.

Figure 4 complements the regression analysis by plotting charitable contributions in year *t+2* as a function of net assets. The figure shows a sharp increase in log contributions at the zero net assets threshold. While it is not obvious whether there is a discontinuous change at the threshold or if the underlying function is instead highly non-linear, the kink in the distribution at the insolvency threshold is consistent with the hypothesis that donors exhibit preferences for solvent firms, which results in bunching charities receiving increased contribution revenue in the following years.

Table 6 repeats the analysis in Table 5 for expenses. Again, if firms inflate their net assets by manipulating their expense reporting, we would expect to see a negative impact of falling in the manipulation region on expenses in the year that bunching occurs. If instead, bunching has an effect on expenses, differences should emerge in the years following bunching, but not necessarily in the year that bunching occurs. Column 1 shows that charities in the manipulation region have expenses that are 11 log points (12 percent) lower in the year in which bunching occurs and that organizations that bunch decrease their expenses by 42 log points (52 percent).[11] Columns 2 and 3 show that there are no significant effects on expenses in years *t+1* or *t+2*, respectively. This suggests that manipulating expenses might be a cause of bunching, but we find no evidence that bunching has an effect on reported expenses in later years.

[11] Note that here we refer to the effect of bunching, not the effect of crossing the insolvency threshold. In appendix figure 2, we show that simply crossing the threshold has no exogenous effect, which is consistent with the discussion in section 2 on the difference between for-profit and nonprofit firms. The figure plots expenses as a function of net assets, with observations in the manipulated range omitted, and shows that the two trend lines from opposite sides of the threshold intersect, confirming that expenses do not change discontinuously. The test is weakened by the need to remove observations near the threshold, but provides some evidence that the change in expenses we report in this section is entirely endogenous.

Lastly, Table 7 presents results for program revenues. Here we find no significant effects of falling in the manipulation region in the year in which bunching occurs in subsequent years. This suggests that manipulation of program revenues does not appear to be a cause of bunching, nor does bunching cause a change in future program revenues.

VII. Discussion

In this paper, we document significant bunching of public charities at the zero net assets threshold. This bunching persists even after limiting our sample to firms that are unlikely to have a mechanical reason to report zero net assets, suggesting that firms may manipulate their income to appear balance sheet solvent. This is somewhat surprising given that non-profits cannot be forced into involuntary bankruptcy in the same manner as for-profits, and there are no direct financial penalities for reporting negative net assets.

We estimate the characteristics of the bunchers and show that the firms that inflate their financial position tend to be smaller charities that earn most of their revenues from donations and are relatively undiversified in their activities. Next, by modeling the counter-factual distributions of various financial outcomes, we explore the causes and consequences of bunching. We find evidence that one mechanism by which charities inflate their net assets is to under-report their expenses. Turning to the consequences of bunching, we find that bunching charities report higher contributions from donors in the years after bunching than they otherwise would, suggesting that one motivation for bunching is to appeal to donor preferences.

These findings are consistent with prior research indicating that donors care about the financial health and viability of the charities they donate to. Specifically, it relates to a literature in which donors rely on heuristics to simplify their decision-making about where to spend their contribution dollars (Karlan and List, 2007; Yoruk, 2016). Our results suggest that donors may not wish to give to insolvent organizations, even though there may be very little difference between a charity that is just barely solvent and one that is just barely insolvent. The emphasis of the insolvency threshold as a binary metric of financial health may be fostered in part by third-party rating agencies, such as GuideStar, which track pefrormance metrics, including specifically whether charities have reported negative net assets. In the face of these donor proferences, nonprofit managers respond by inflating their reported financial health.

What, if any, conclusions can we draw about welfare? If the use of heuristics and third-party ratings lead to the misallocation of charitable dollars, the frequent use of financial metrics in the nonprofit sector may have significant negative welfare consequences. Alternatively, if the use of heuristics improves upon the allocation of charitable donations, they may enhance the efficiency of public goods provision. Additionally, if the increase in charitable contributions that we document reflects an increase in overall donations for the sector as a whole, it is possible bunching could actually increase public goods provision. In the absence of clear performance metrics, it is difficult to definitively conclude that bunching firms operate less efficiently. Moreover, we have confined our analysis to a specific region of the data and our findings may have limited generalizability to firms in better financial condition. We leave to future work to expand on the implications of donor preferences and the use of heuristics for the allocative efficiency of charitable contributions.

References

Bowman, Woods. 2011. *Finance Fundamentals for Nonprofits.* Vol. 2, John Wiley & Sons.

Bulow, Jeremy I, and John B Shoven. 1978. "The Bankruptcy Decision." *The Bell Journal of Economics*, 437–456.

Calabrese, Thad. 2013. "Running on Empty: The Operating Reserves of US Nonprofit Organizations." *Nonprofit Management and Leadership*, 23(3): 281–302.

Calabrese, Thad, and Cleopatra Grizzle. 2012. "Debt, Donors, and the Decision to Give." *Journal of Public Budgeting, Accounting & Financial Management*, 24(2): 221–254.

Calabrese, Thad D. 2011. "Do Donors Penalize Nonprofit Organizations with Accumulated Wealth?" *Public Administration Review*, 71(6): 859–869.

Chetty, Raj, and Emmanuel Saez. 2013. "Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients." *American Economic Journal: Applied Economics*, 5(1): 1–31.

Chetty, Raj, John N Friedman, Tore Olsen, and Luigi Pistaferri. 2011. "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." *The Quarterly Journal of Economics*, 126(2): 749–804.

Dee, Thomas S, Will Dobbie, Brian A Jacob, and Jonah Rockoff. Forthcoming. "The Causes and Consequences of Test Score Manipulation: Evidence from the New York Regents Examinations." *American Economic Journal: Applied Economics*.

Diamond, Rebecca, and Petra Persson. 2016. "The Long-term Consequences of Teacher Discretion in Grading of High-stakes Tests." National Bureau of Economic Research 22207.

Elliot, Dana, and Evan Hollander. 2014. "Navigating a Nonprofit Corporation Through Bankruptcy." *Nonprofit Quarterly*.

FASB. 2014. "Presentation of Financial Statements—Going Concern: Disclosure of Uncertainties about an Entity's Ability to Continue as a Going Concern." Financial Accounting Standards Board.

Karlan, Dean, and John A List. 2007. "Does Price Matter in Charitable Giving? Evidence from a Large-scale Natural Field Experiment." *American Economic Review*, 97(5): 1774–1793.

Kleven, Henrik Jacobsen. 2016. "Bunching." *Annual Review of Economics*, 8: 435–464.

Kleven, Henrik J, and Mazhar Waseem. 2013. "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan." *The Quarterly Journal of Economics*, 128(2): 669–723.

Krishnan, Ranjani, and Michelle H Yetman. 2011. "Institutional Drivers of Reporting Decisions in Nonprofit Hospitals." *Journal of Accounting Research*, 49(4): 1001–1039.

Marx, Benjamin M. 2018. "The Cost of Requiring Charities to Report Financial Information."

Okten, Cagla, and Burton A Weisbrod. 2000. "Determinants of Donations in Private Nonprofit Markets." *Journal of Public Economics*, 75(2): 255–272.

Onji, Kazuki. 2009. "The Response of Firms to Eligibility Thresholds: Evidence from the Japanese Value-added Tax." *Journal of Public Economics*, 93(5): 766 – 775.

Peterman, Nancy A, and Sherri Morissette. 2004. "Directors' Duties in the Zone of Insolvency: The Quandary of the Nonprofit Corporation." *American Bankruptcy Institute Law Review*, 23: 12.

Saez, Emmanuel. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy*, 2(3): 180–212.

St. Clair, Travis. 2016. "How Do Nonprofits Respond to Regulatory Thresholds: Evidence From New York's Audit Requirements." *Journal of Policy Analysis and Management*, 35(4): 772–790.

White, Michelle J. 1989. "The Corporate Bankruptcy Decision." *Journal of Economic Perspectives*, 3(2): 129–151.

Yoruk, Baris K. 2016. "Charity Ratings." *Journal of Economics & Management Strategy*, 25(1): 195–219.

Figure 1a: Bunching in Raw Data

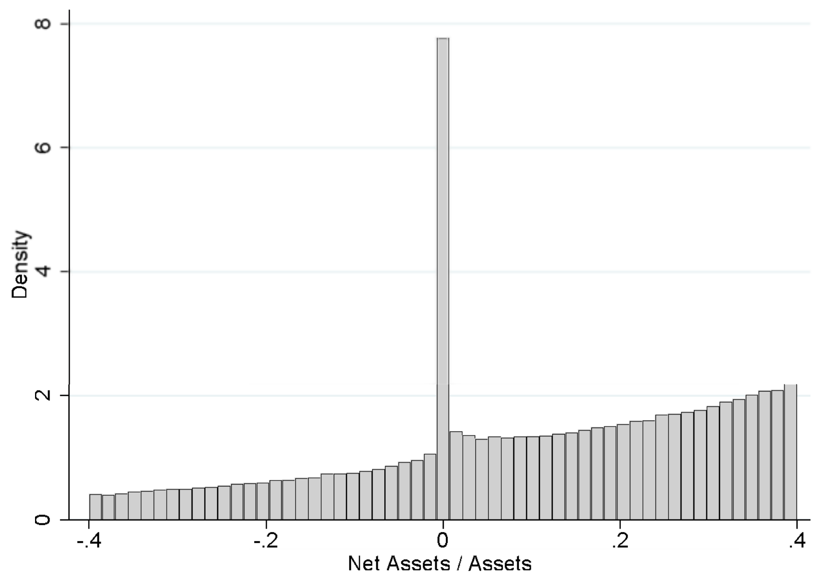
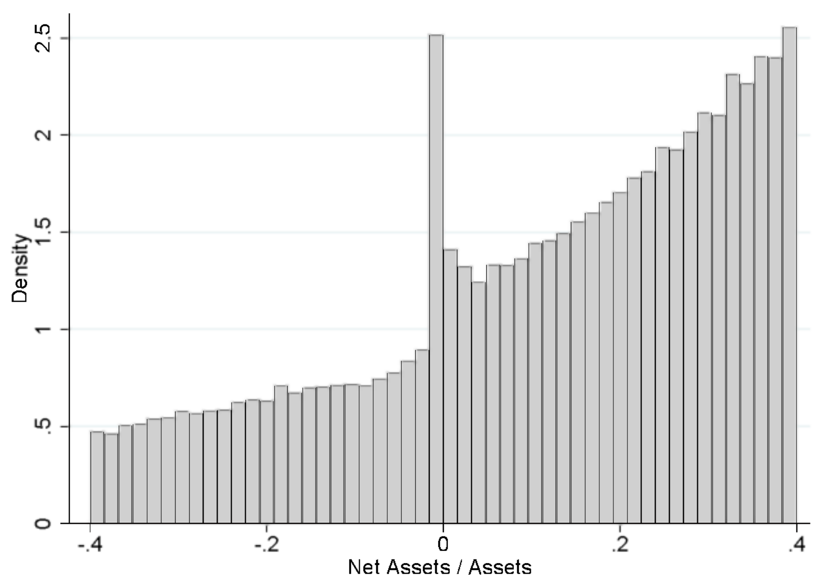
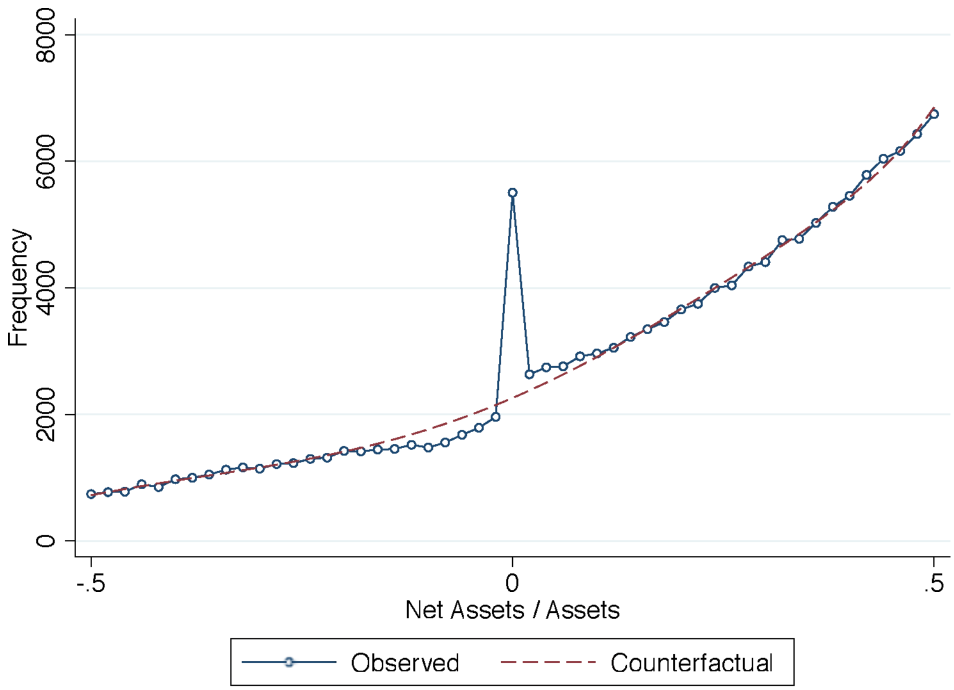


Figure 1b: Bunching in Final Sample



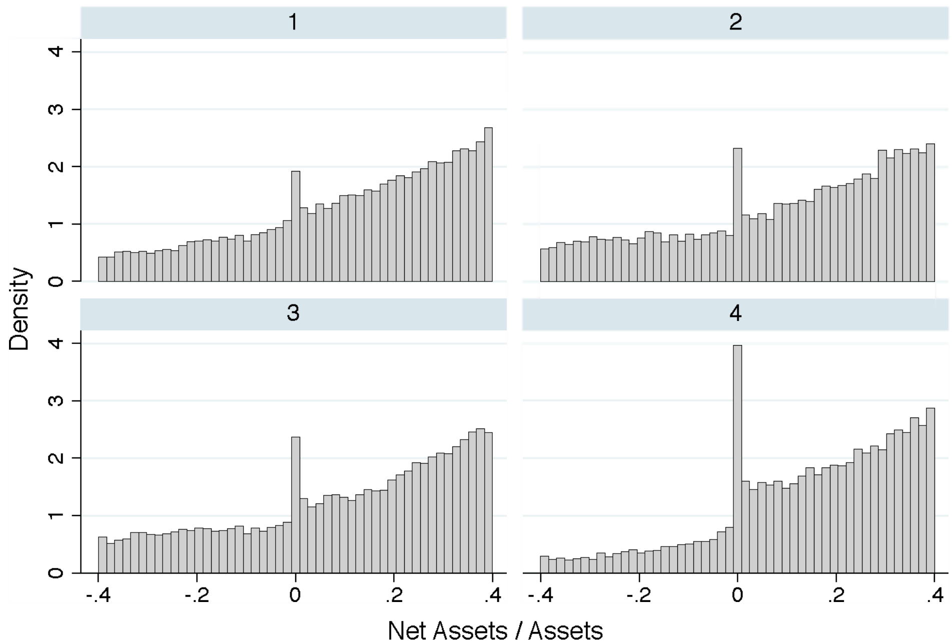
Note: Figure 1a shows a density distribution of the full sample by net assets. Figure 1b shows the same density distribution for our analytic sample after placing a series of restrictions on the sample.

Figure 2: Bunching at Near-Zero Net Assets



Note: This figure shows the density distribution of public charities in the vicinity of zero net assets. The dashed line beneath the observed distribution is a fifth degree polynomial fitted to observations outside the manipulated region (-0.16 - 0.8). Each point represents the number of charity-years in a bin of size 0.02.

Figure 3: Bunching by the Percentage of Revenues from Contributions (Quartiles)



Note: This figure shows heterogeneity in the extent of bunching by splitting the sample into quartiles according to the percentage of revenues that charities receive from contributions.

Figure 4: Log Contributions, Year t+2

Note: The figure plots log contributions in year t+2 by net assets in year t where year t is defined as the year in which bunching occurs. Each observation represents the local sample mean for bins of size 0.03. The dashed-lines are linear trends fit to the observations that fall outside of the manipulation range.

Table 1: Summary Statistics

	Analytic Sample		
	Mean	Median	SD
	(1)	(2)	(3)
Assets (millions)	37	0.14	61
Liabilities (millions)	1.5	0.001	0.056
Net Assets / Assets	0.78	0.99	0.52
Contributions (millions)	0.68	0.044	8.2
Program Revenue (millions)	1.7	0.008	55
Contributions/ Revenue	0.50	0.48	0.40
Expenses (millions)	2.5	0.11	55
Age of Organization	25	22	14
Insolvent (Net Assets < 0)	0.05	0	0.23

Note: Data come from the National Center of Charitable Statistics' (NCCCS) 2005-2015 core files for public charities. N = 1,187,838 observations, 266,184 charities. The ratio variables have been windsorized at the top and bottom 1 percent of the distribution.

Table 2: Measures of Bunching

	(1) Total Manipulation	(2) In-Range Manipulation	(3) Total Manipulation	(4) Total Manipulation
Excess Mass	0.326***	26.6***	0.344***	0.322***
	(0.016)	(1.41)	(0.017)	(0.017)
Manipulation Region	-0.16 - 0.08	-0.16 - 0.08	-0.20 - 0.10	-0.12 - 0.06

Note: *** $p < 0.001$. This table presents estimates of the manipulation in the sample, which is measured by the excess mass to the right of the solvency threshold. Total manipulation is the excess mass as a percentage of all charities in the sample. In-range manipulation is the excess mass relative to the counterfactual distribution in the range of the missing mass, or the probability of manipulation conditional on reporting net assets just below the cut-off. Standard errors are calculated using the parametric bootstrap procedure described in the text. N = 1,187,838 (total sample), N = 144,486 within the range of estimation (-0.5 < Net Assets/Assets < 0.5).

Table 3: Characteristics of Bunchers

	(1) Bunching Charities	(2) Falls in Region of Missing Mass	(3) Difference
Type of Charity			
Arts, Culture, & Humanities	0.095	0.098	-0.004
			(0.015)
Education	0.124	0.084	0.040*
			(0.017)
Health	0.101	0.111	-0.010
			(0.019)
Human Services	0.394	0.530	-0.136**
			(0.030)
Other	0.287	0.177	0.109**
			(0.016)
Financial Characteristics			
Log Assets	5.96	13.3	-7.67**
			(0.23)
Mortgage (Yes/No)	0.102	0.505	-0.403**
			(0.037)
Percent Revenue from Contributions	0.678	0.381	0.297**
			(0.025)
Reported Unrelated Business Income (Yes/No)	0.013	0.049	-0.036**
			(0.009)
Reported Program Revenue (Yes/No)	0.381	0.818	-0.437**
			(0.026)
Sold Inventory (Yes/No)	0.030	0.093	-0.063**
			(0.011)
Total Compensation as Percent of Expenses	0.419	0.264	0.155**
			(0.019)

Note: ** $p < 0.01$,* $p < 0.05$. This table presents characteristics of the bunching charities and compares these characteristics to all charities that fall just below the threshold and thus also might have chosen to bunch. The characteristics of the bunching charities are distinguished from the charities that otherwise fall above the notch (the "always-takers") by using an estimated counterfactual to determine the characteristics of the always-takers. To obtain the counterfactual estimates, we fit polynomials to the observed distribution outside of the manipulation region. Column 3 presents the difference between the estimates. Charities in the "other" category include environmental, international, public benefit, and religious organizations. Standard errors are calculated using the parametric bootstrap procedure described in the text.

Table 4: Bunching by the Percent Revenues from Contributions (Quartiles)

	(1) In-Range Manipulation	(2) In-Range Manipulation	(3) In-Range Manipulation	(4) In-Range Manipulation
Excess Mass	11.2***	12.4***	19.8***	83.4***
	(2.43)	(2.22)	(2.57)	(7.27)
N	282,422	282,427	282,425	282,428

Note: *** $p < 0.001$. This table presents measurements of the excess mass by quartiles of the percentage of revenue that charities receive from contributions. All estimates reflect in-range manipulation, i.e., the probability of manipulation conditional on reporting net assets just below the cut-off. Standard errors are calculated using the parametric bootstrap procedure described in the text. We remove observations with percentages less than zero or greater than one.

Table 5a: Reduced Form: Impact of Falling in Manipulation Region on Contribution Revenue

	(1)	(2)	(3)
	Log Contributions		
	Year t	Year t+1	Year t+2
Intent-to-Treat	-0.030	0.131	0.174**
	(0.061)	(0.081)	(0.069)

Table 5b: Impact of Bunching on Contribution Revenue (LATE)

	(1)	(2)	(3)
	Log Contributions		
	Year t	Year t+1	Year t+2
LATE	-0.113	0.494	0.656**
	(0.231)	(0.303)	(0.258)

Note: ** $p < 0.01$, * $p < 0.05$. Table 5a presents estimates of the impact of falling in the manipulation region on log contributions, while Table 5b presents estimates of the effect of bunching on log contributions. The counterfactual is estimated from regressions of log contributions on 2rd order polynomials, estimated separately on both sides of the cut-off. The counterfactual uses only data from outside the manipulation region. Standard errors are calculated using the parametric bootstrap procedure described in the text.

Table 6a: Reduced Form: Impact of Falling in Manipulation Region on Expenses

(1)	(2)	(3)
Log Expenses		
Year t	Year t+1	Year t+2
-0.111**	0.075	0.0034
(0.045)	(0.055)	(0.053)

Table 6b: Impact of Bunching on Expenses (LATE)

(1)	(2)	(3)
Log Expenses		
Year t	Year t+1	Year t+2
-0.419**	0.283	0.013
(0.170)	(0.207)	(0.198)

Note: Table 6a presents estimates of the impact of falling in the manipulation region on log expenses, while Table 6b presents estimates of the effect of bunching on log expenses. The counterfactual is estimated from regressions of log expenses on 2rd order polynomials, estimated separately on both sides of the cut-off. The counterfactual uses only data from outside the manipulation region. Standard errors are calculated using the parametric bootstrap procedure described in the text.

Table 7a: Reduced Form: Impact of Falling in Manipulation Region on Program Revenue

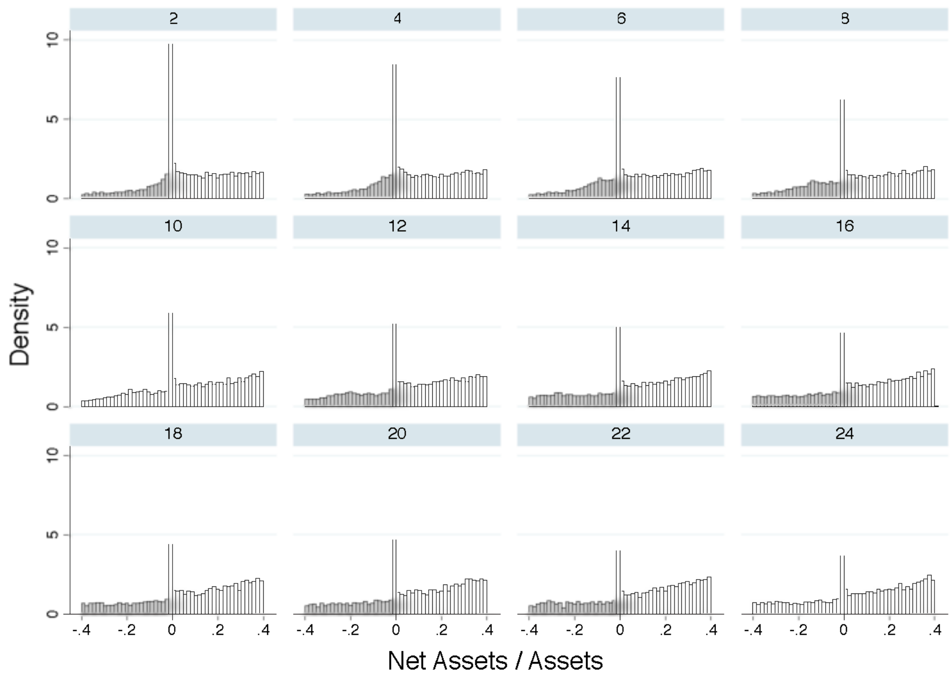
(4)	(5)	(6)
Log Program Revenue		
Year t	Year t+1	Year t+2
-0.055	0.095	0.0163
(0.047)	(0.077)	(0.065)

Table 7b: Impact of Bunching on Program Revenue (LATE)

(4)	(5)	(6)
Log Program Revenue		
Year t	Year t+1	Year t+2
-0.209	0.359	0.061
(0.178)	(0.292)	(0.0243)

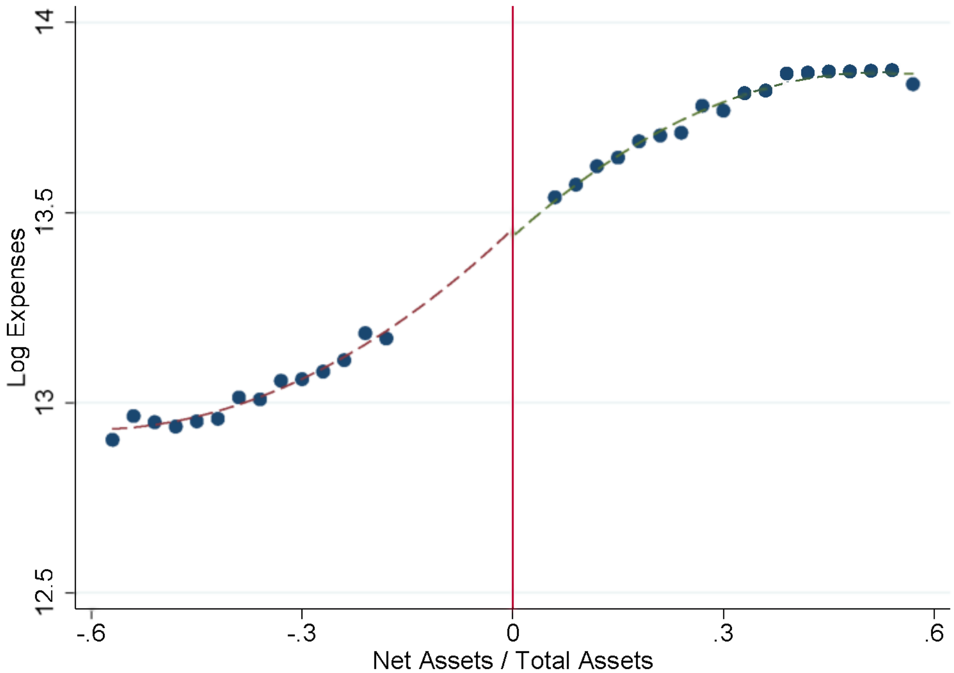
Note: Table 7a presents estimates of the impact of falling in the manipulation region on log program revenue, while Table 7b presents estimates of the effect of bunching on log program revenue. The counterfactual is estimated from regressions of log program revenue on 2rd order polynomials, estimated separately on both sides of the cut-off. The counterfactual uses only data from outside the manipulation region. Standard errors are calculated using the parametric bootstrap procedure described in the text.

Appendix Figure 1: Density Distribution by Age



Note: This figure presents the density distribution by net assets for charities of different ages. The age of the charity is based on the year in which the IRS recognized the organization's tax exempt status, as reported on the Form 990.

Appendix Figure 2: Expenses



Note: This figure plots log expenses by net assets in the region of the insolvency threshold. Each observation represents the local sample mean for bins of size 0.03. The trend lines are estimated using second order polynomials. The excluded range includes net assets greater than -0.16 and less than 0.08. For increased power, we use the full raw data file, consisting of 3,485,306 observations and 578,282 charities from the National Center of Charitable Statistics' (NCCCS) 2005-2015 core files for public charities. Within the region of the data shown in the figure, N = 520,914 observations and 148,687 unique charities.

Appendix Table 1: Sample Restrictions

Restriction	Number of Observations Lost	Sample Size
		3,485,306
Exclude charities in operation for less than ten years or lacking data on year of IRS recognition	-1,254,320	2,230,986
Exclude charities with zero assets, zero average income, or zero average contributions	-154,699	2,076,287
Exclude charity-year observations with zero net assets at the beginning of the year	-25,702	2,050,588
Exclude charity-year observations with "other changes in net assets"	-862,747	1,187,838

Note: The raw data consists of 3,485,306 observations and 578,282 charities and comes from the National Center of Charitable Statistics' (NCCCS) 2005-2015 core files for public charities. The final sample consists of 1,187,383 observations and 268,184 charities.

Appendix Table 2: AIC for Different Polynomial Choices

(1) Polynomial Order	(2) AIC
First	-188.7963
Second	-191.1385
Third	-196.2652
Fourth	-194.5196
Fifth	-210.5207

Note: The table reports Akaike Information Criteria for a series of regressions of nonprofit frequency on various polynomials of net assets.