The Heterogeneous Effects of Accrual Accounting: Evidence from Municipal Borrowers

Travis St.Clair
Wagner School of Public Service
New York University

January 14, 2019

Abstract

The literature on accounting quality has long held that high quality accounting information reduces information asymmetries between borrowers and lenders, however prior work suggests that there may be substantial heterogeneity in the effects. I explore this question further in the context of municipal borrowers. In 1999, the Governmental Accounting Standards Board (GASB) introduced a new reporting model for state and local governments, requiring governments for the first time to report on a government-wide full accrual basis. I exploit the staggered phase-in of the new reporting model to examine its impact on the cost and use of municipal debt. While reporting on a full-accrual basis appears to have had a slightly beneficial effect for the average government borrower, regression discontinuity results show that for larger governments, the use of debt actually decreased while the cost of debt rose. To explain these results, I draw on findings from the credit ratings literature.

JEL codes: M41, H74

Keywords: government accounting, municipal debt, GASB, financial reporting

1 Introduction

The literature on accounting quality has long held that high quality information reduces information asymmetries between borrowers and lenders and lowers the cost of capital. This finding has proved consistent across various sectors and types of information, including the use of accrual accounting among small businesses (Cassar, Ittner, and Cavalluzzo, 2015), the auditing of privately held firms' financial statements (Minnis, 2011), and requirements mandating GAAP for municipal governments (Baber and Gore, 2008; Gore, 2004; Fairchild and Koch, 1998). While this work is informative about the effects on the average borrower, more recent work suggests that there may be substantial heterogeneity in the effects across different types of borrowers. For example, Cassar, Ittner, and Cavalluzzo (2015) show that the benefits of accrual accounting are greatest when lenders have few alternative sources of information and little prior relationship with borrowers. On the other hand, when other sources of high quality information exist, accrual information has fewer incremental benefits.

This paper further explores the heterogeneous effects of accrual accounting in the context of municipal borrowing. In 1999, the Governmental Accounting Standards Board (GASB) introduced a new reporting model for state and local governments, significantly altering the format and basis of governmental financial statements. For the first time, governments were required to prepare government-wide financial statements on a full accrual basis. The Statement of Net Assets, akin to a corporate balance sheet, would provide information about the long-term assets and liabilities of a government entity as a whole, while the Statement of Activities, akin to a corporate income statement, would provide information on accrual-based revenues and expenses as well as the change in net assets for the year. This change marked a significant shift in the scope of government financial accounting. Instead of focusing solely on near-term financial resources, the new reporting model enabled users to obtain a picture of a government's economic condition in its entirety and better assess the long-term impacts of past decisions. Upon the release of the new standard, GASB's Chairman described it as "the most significant change in the history of governmental accounting"

(Plummer, Hutchison, and Patton, 2007).

Given the extensive nature of the changes, GASB phased in the implementation of the new standard, allowing smaller governments more time to comply. Phase 1 governments, defined as those governments with more than $100 million in total revenues, were required to implement GASB 34 in the first period beginning after June 15, 2001. Phase 2 and phase 3 governments, those with total revenues of less than $100 million and less than $10 million respectively, were given one and two additional years to comply.

Using a regression discontinuity approach, I examine the impact of the accounting change on municipal borrowing. Given that there are few opportunities to leverage randomization or discontinuities in the roll-out of new accounting standards, the staggered phase-in of GASB 34 represents an opportunity to rigorously evaluate the impact of a major change in financial reporting using a quasi-experimental design. It also represents an improvement over previous analyses of government accounting standards, many of which have relied on selection-on-observables strategies and strong identification assumptions.

Others have written previously on the information relevance of GASB 34. Plummer, Hutchison, and Patton (2007) examined 530 Texas school districts to specifically investigate the information relevance of the new reporting model, finding that GASB 34's Statement of Net Assets provides information on default risk above and beyond that provided by governmental fund statements, but that the Statement of Activities does not. Reck and Wilson (2014) reached a more mixed conclusion, finding that the GASB 34 accrual information does not provide additional explanatory power in a model of interest costs relative to aggregate modified accrual information.

Notably, these studies are cross-sectional in nature and rely on selection-on-observables strategies, leaving open the possibility that their results may be biased by unobservable factors that influence borrowing costs. Moreover, they focus on the average borrower and do not investigate how the effects may vary across different types of borrowers . This paper differs in two dimensions. First, by exploiting the discontinuity in the roll-out of GASB 34, I am able to control for unobservable differences between issuers that may have been a source of bias in previous studies. Second, by utilizing both regression discontinuity and panel methods, I am able to estimate *local* average treatment effects (LATE) at distinct thresholds in addition to an average treatment effect (ATE). The local effects may differ substantially from the average effect if accrual accounting affects certain types of issuers differently, as Cassar, Ittner, and Cavalluzzo (2015) suggest.

Indeed, this paper shows that while the average effect of GASB appears to have been slightly positive, the effect on large issuers was actually negative. That is, when higher quality accounting information was introduced, for large issuers the use of debt decreased and the cost of debt rose. Although this finding is novel in the accounting literature, it is consistent with recent findings in the finance literature showing that the credit ratings of large issuers are sometimes inflated as a result of the market power that they possess. To provide evidence of this channel – that the negative effect on large governments was due to inflated credit ratings – I demonstrate that the phase 1 governments in my sample were more likely to experience rating declines in the year after GASB 34 than phase 2 governments, even though the two groups had similar ratings in the prior period. The findings suggest that, in the absence of high quality financial reporting, borrowers relied on credit rating information that was in some cases biased.

This paper proceeds as follows. Section II provides background on GASB 34 and the changes that it brought to government financial reporting. Section III discusses the data and methods, including both panel data and regression discontinuity approaches. Results are presented in Section IV. Section V examines the role of credit ratings and explores other possible explanations for the findings. Section VI concludes.

2 Background on GASB 34

Prior to the adoption of GASB 34, the focus of state and local government financial statements in the U.S. was on summarizing financial information on a modified accrual basis by the type of fund. Funds are established by governing bodies to show restrictions on the use of resources and to comply with finance-related laws, such as balanced budget restrictions. The adoption of GASB 34 resulted in three major changes: 1) government-wide financial statements prepared on a full accrual basis, including a Statement of Net Assets and a Statement of Activities, 2) a Management's Discussion and Analysis (MD&A) section providing a readable analysis of the government's financial condition, and 3) separate reporting of major funds that constitute a significant proportion of resources rather than only reporting by fund type.

The most significant change was the requirement for goverment-wide financial statements. The statements, prepared using the economic resources measurement focus and a full-accrual basis, would consist of a Statement of Net Assets and a Statement of Activities, requiring governments for the first time to report on all capital assets, including infrastructure assets, as well as to report depreciation expenses.[1] In the case of general infrastructure assets that had to be retroactively capitalized, governments were given a grace period of four years to complete the capitalization. While the new reporting model required government-wide statements, it did not abandon fund reporting, as governments were still required to prepare governmental fund statements on a modified accrual basis.

The focus of modified accrual accounting is on near-term financial resources rather than total economic resources. It is unique to the government sector because it demonstrates whether the reporting entity acquired and used its resources according to its legally adopted budget. So whereas previously users were able to assess whether, for example, the general fund of a municipality was in balance, with the new reporting model they were able to

[1] Governments using an asset management system were allowed to use a "modified" approach for reporting infrastructure assets if they could document that the assets were being preserved at a certain level. See GASB (1999) for full details.

more comprehensively assess the finances of the government in its entirety.[2] Moreover, they could now assess the medium and long-term effects of past decisions rather than simply the short-term effects and determine whether governments were shifting costs to future years.

It is important to note that the new model did not in practice capture all economic resources. Future GASB statements would specifically address pension and other post-employment benefit (OPEB) liabilities. Furthermore, other pieces of information, such as the total outstanding principle on general obligations bonds, were previously available in the notes even if they were not directly visible on the financial statements. Nevertheless, the new reporting model provided a great deal more insight into a government's overall financial condition than had been previously available in the financial statements, including information on capital assets, compensated absences, and the full extent of bonds payable.

The other changes stemming from GASB 34 were less significant, but nevertheless resulted in major changes to the format and presentation of the statements. The MD&A section gave financial managers an opportunity to share their insights by providing a readable analysis of the government's recent operations and a discussion of any significant changes that occurred in particular funds or significant budget variances. Instead of simply reporting funds in the aggregate by type on the fund statements, the new model also required separate columns for major funds, defined as funds whose revenues, expenditures/expenses, assets, or liabilities were at least 10 percent of corresponding totals for all governmental or enterprise funds and at least 5 percent of the aggregate total for all governmental and enterprise funds. The change required governments for the first time to provide more specific information on funds of particular consequence, such as water and sewer funds, that previously were lumped into columns summarizing all activity within a certain fund type.

Whether GASB 34 in particular resulted in significantly new information to investors however is a different question. As noted above, Plummer, Hutchison, and Patton (2007) found that the Statement of Net Assets provides information on default risk above and

[2] Fiduciary activities, whose resources are not available to fund government programs, are not included in the government-wide statements.

beyond the information provided by governmental fund statements, but that the Statement of Activities does not, suggesting that the accrual measure of "earnings" is not more informative than the modified-accrual measure. However, their analysis focused only on Texas school districts and is correlational in nature, lacking a research design that would enable them to make stronger causal claims. This paper, in contrast, draws on data from local governments nationwide and leverages quasi-experimental variation.

3 Data and Methods

3.1 Data

This paper uses data from three different sources: 1) the Census of Governments and the Annual Survey of State and Local Finances, 2) data collected from close to 300 individual government financial statements, and 3) SDC Platinum's municipal bond database. The primary source of data is the Census of Governments and the Annual Survey of State and Local Government Finances.[3] The Census contains comprehensive data on government revenues and debt issuance back to 1967; indeed, it is "the only comprehensive source of information on the finances of local governments in the United States" (Pierson, Hand, and Thompson, 2015).

Since not all governments follow GAAP, the sample excludes local governments for whom GASB 34 did not apply. This drops school districts in Indiana, Missouri, Washington, and West Virginia as well as local governments in New Jersey and Indiana.[4] The sample also

[3]The year of the census data does not always correspond to the fiscal year of the government; a "survey year" includes each government's fiscal year that ended between July 1 of the previous year and June 30 of the survey year, meaning that governments with fiscal years ending after June 30 are represented in the following year's survey data. There are a few exceptions for school districts; see Census of Government user note files.

[4]This list is based on findings from the data collection process described in section four. It does not coincide perfectly with Baber and Gore (2008), who list ten states that did not impose reporting requirements during 1995-2002. In part this is because the sample does not exclude states who may not otherwise follow all aspects of GAAP but implemented GASB 34 on schedule. As Khumawala, Marlowe, and Neely (2014) point out, it is not uncommon for local governments in non-GAAP states to prepare financial statements that have many of the same components of GAAP. Moreover, if the list is incomplete, it is likely that the results will

excludes school districts in Texas since, as Plummer, Hutchison, and Patton (2007) note, the state of Texas mandated that all local governments implement GASB 34 early regardless of revenue level. Given these sample restrictions, the estimand can be interpreted as the impact of GASB 34 on governments that follow GAAP.

Two challenges with the data present themselves. Since the Census collects data on the amount of debt issued by governments but does not collect data on borrowing costs, it is necessary to match the census data with the primary market data from SDC. However, there is no common id that facilitates a crosswalk between the two datasets, which requires that individual governments must be matched to their debt issues using "fuzzy name matching." In practice, this severely restricts the number of observations that can be feasibly matched, and as a result the amount of data available to analyze the cost of debt is much less than the amount of data available to analyze the use of debt. Appendix Part 2 details the process used to match government borrowers with particular debt issues. The other challenge is that the Census data consists of survey responses, which results in measurement error in certain key variables. I discuss this issue, and the effect it has on the analysis, in more detail in the next section.

For the cost of debt, I look at the true interest cost of the issue. For the use of debt, I draw on two variables in the census data pertaining to public borrowing: 1) a binary measure for whether a government issued long-term debt in a given year[5], and 2) the amount of new debt issued in a given year. The amount of new debt is calculated by adding the total amount of long-term debt issued to the amount of short-debt outstanding at the end of the year and subtracting the amount of long-term debt retired.[6] Note that this measure of new

simply understate the effects of GASB 34, as the sample will include some states for whom the assignment to treatment had no effect.

[5]The surveys that governments fill out ask about long-term debt for public purposes, long-term debt for private purposes, and short-term debt. Although they refer to debt "issued," the questions are intended to capture all long-term debt and do not distinguish between public and private debt.

[6]An alternative measure, almost exactly equivalent, would be to measure the change in total debt outstanding. The only difference is in the measurement of short-term debt (relying on the year-end balance rather than the change). The problem with this approach is that the measurement of total debt outstanding may have been affected by the treatment. That is, governments that implemented GASB 34 may have calculated this measure differently based on their new accrual-based accounting system, despite the fact that the

debt can take negative values if a government retired more debt than it issued.

3.2 Empirical Strategy

3.2.1 Regression Discontinuity

GASB issued Statement No. 34 in June of 1999 and required governments to implement the new standard on a staggered basis based on their total revenues in the period ending after June 15, 1999. Since the majority of state and local governments in the United States use a fiscal year of July 1 to June 30, for most governments the date of implementation was determined based on their revenue collection in fiscal year 1999. For phase 1 governments, those with more than \$100 million in total revenues in 1999, the new reporting model was effective for the first period beginning after June 15, 2001. For phase 2 governments, those with less than \$100 million and more than \$10 million in revenues, the standard was effective for the first period beginning after June 15, 2002. For phase 3 governments, those with less than \$10 million in revenues, the standard was effective for the first period beginning after June 15, 2003. Thus, for the "typical" phase 1 government with a fiscal year ending on June 30, GASB 34 was effective in fiscal year 2002. For the typical phase 2 government, it was effective in fiscal year 2003.

The staggered roll-out of the standard created exogenous variation at the cut-off that can be exploited using a regression discontinuity (RD) design. An RD design represents a more credible means of assessing the impact of GASB 34 than is typically available in the analysis of major accounting standards. Although the use of unaudited financial data presents some challenges (discussed in more detail below), RD designs represent a much more credible and transparent identification strategy than simple observational studies using OLS regression. And while RD designs are invalid in settings where agents can manipulate the running variable, GASB issued Statement No. 34 in June of 1999 and based its phase-in

Census posed questions about total debt outstanding even in years prior to 2002. Hence, I prefer to construct a measure of new debt that is independent of the measure of total debt outstanding.

schedule on the revenues that governments collected that same year; consequently there was little opportunity for governments to manipulate their fiscal year 1999 revenues post-hoc. Appendix Figure 1 shows a density plot of governments near the threshold, confirming that there was no evidence of bunching below the revenue cut-off.

RD designs can be analyzed in one of two ways: as a randomized experiment at the threshold or using a continuity-based approach. The continuity-based framework, rather than relying on as-if randomization, makes the less restrictive assumption that, in the absence of treatment, the outcome variable is continous at the cut-off (Sekhon and Titiunik, 2017). Although it can be advantageous for reasons of precision to adopt the local randomization assumption, that approach also requires stronger assumptions, and consequently this paper mainly employs regression methods based on the continuity-based framework. (Results based on the local randomization assumption are presented in an appendix.)

One limitation of RD – no matter how the treatment effect is estimated – is the local nature of the estimand; the design permits identification of the local average treatment effect (LATE) at a cut-off. In this case, there are actually two cut-offs – \$100 million in total revenues and \$10 million in total revenues. Although the local nature of the RD estimand is typically thought of as a shortcoming, in this case it presents some advantages, as it provides an opportunity to separately estimate treatment effects for both small and large issuers. In 1999, governments with \$100 million in revenues fell at approximately the 94th percentile in a ranking of governments by revenue, while governments with \$10 million in revenues fell at the 59th percentile. (It's worth noting that a large number of government entities consist of authorities that issue debt but do not independently collect much revenue, and thus fall into GASB's Phase III category by default.) Given Cassar, Ittner, and Cavaluzzo (2015)'s finding that accrual accounting primarily benefits smaller entities, for whom fewer alternative sources of information are available, the a priori hypothesis is that GASB 34 benefitted the smaller issuers more than the larger issues. However, one countervailing factor is that very small governments are less likely to be GAAP-compliant (Khumawala, Marlowe, and Neely,

2014) and thus are less likely to be affected by changes in financial accounting.

3.2.2 Measurement Error and Noncompliance

The census' measure of total revenues includes all tax revenues, intergovernmental revenues, sales and service revenues but excludes transfers between accounts of the government, including internal service funds. This is consistent with the total revenue measure that GASB uses to differentiate between phase 1 and phase 2 governments, which includes all revenues (not other financing sources) of the primary government's governmental and enterprise funds.[7] However, despite the census reporting a measure of total revenues, the variable that it reports comes from surveys, meaning that the running variable contains measurement error. Moreover, some of the self-reported data are based on unaudited financial statements, while the total revenue number that GASB uses is based on audited financials. Several recent papers in the economics literature discuss RD designs with measurement error in the assignment variable and the challenges that it poses for identification (Hullegie and Klein, 2010; Pei and Shen, 2016; Davezies and Le Barbanchon, 2017). With classical (smoothly distributed) measurement errors in the running variable, the discontinuity in the assignment probability vanishes, rendering the RD design invalid.

However, in this case not all of the data is measured with error. Not only is some of the self-reported data based on verified information, but the census also *edits* the data, checking the reported numbers against audited financial statements; in fact, the census' division of state and local finance spends a good portion of each year checking the responses against financial statement data.[8] Hence, the running variable in this case contains a mixture of values that are measured correctly and values that are reported with error, which Horowitz and Manski (1995) refer to as the contaminated sampling model. Specifically, the analysis assumes that

[7] If component units are considered to be separate special purpose governments, then they receive their own survey forms from the census. They do not receive a separate survey if they are part of the parent government.

[8] This was confirmed via e-mail correspondence with census staff.

$$R_{obs} = R^* Z + R(1 - Z) \tag{1}$$

where R^* and R denote the true and error-prone measurements of revenues, Z is a binary indicator for those who report correctly, and R_{obs} is the survey measurement actually observed. In an example that is very similar to the one discussed here, Battistin et. al (2009) show that the RD design is still valid in a contaminated sampling model as long as there is still a discontinuity in the take-up of treatment. However, the sharp RD estimate that is obtained from error-prone data will be downward biased and subject to a correction factor. Specifically, under a sharp RD design, the causal effect of being assigned to treatment, β, is measured by the difference in the mean outcomes of governments marginally above and below the threshold $\bar{r}$:

$$E\{\beta|\bar{r}^+\} = E\{Y|\bar{r}^+\} - E\{Y|\bar{r}^-\} \tag{2}$$

where Y is the outcome of interest and r is the running variable. Battisten et. al (2009) show that, under the assumption that the process generating measurement error is orthogonal to the outcome of interest, Y:

$$E\{Y|R_{obs} = \bar{r}^+\} - E\{Y|R_{obs} = \bar{r}^-\} = (E\{Y|R^* = \bar{r}^+\} - E\{Y|R^* = \bar{r}^-\})E\{Z|R_{obs} = \bar{r}^-\} \tag{3}$$

In other words, the amount of downward bias in the RD estimate will be equal to the factor $E\{Z|R_{obs} = \bar{r}^-\}$, which represents the proportion of observations that are without error.

In order to confirm that there is a discontinuity in the take-up of treatment, Figure 1 and Table 1 present compliance data that was collected from all governments within a \$10 million window of the \$100 million threshold. Of the 326 governments within this window, there were 277 who comply with GAAP and thus formed part of the sample. Of the 277, it was possible to obtain financial statements for 217 (78%). Figure 1 plots the results, indicating the extent to which governments above and below the cut-off implemented GASB 34 in the initial year of implementation. Table 1 shows the same results in table form, with the first row showing the distance from the threshold, and the third row showing the percentage of governments that implemented the standard.

The figure confirms that, despite the presence of measurement error in the running variable, there is still a discontinuity in treatment take-up at the threshold, confirming that the RD design remains valid. The two trend lines have intercepts that are approximately 25 percentage points apart. As an additional robustness check, the following sections also present placebo results from before GASB 34 was implemented, providing further evidence that there is indeed a discontinuity in the probability of take-up, and furthermore that any causal effects of GASB 34 are not merely the product of misspecification.

The relatively small size of the discontinuity in Figure 1 does not only stem from measurement error. As with the Battistin et. al paper, there is also noncompliance. While GASB clearly delineated effective dates for phase 1, phase 2, and phase 3 governments, the organization also encouraged governments to adopt the standard early if possible, and many elected to do so. In addition to early implementation on the part of phase II governments, there may have also been some noncompliance for phase 1 governments that misunderstood which effective date applied to them or were simply unable to implement the standard on time.

As equation 2 shows, correcting for the downward bias in the RD estimate requires an estimate of the proportion of observations that are measured without error. This is complicated by the noncompliance; it is impossible to know how much of the take-up on the left side of the threshold in Figure 1 (as well as the lack of take-up on the right side) is due to noncompliance and how much is due to measurement error in the running variable. However, the figure does provide a clear upper bound. Based on both the left-hand and right-side intercepts, the proportion of error-prone measurements appears to be no higher than 0.8

(since an intercept of 0.5 would potentially indicate all observations are measured with error). Given that there is noncompliance, it is likely that the true measure is significantly less than that. Hence, the proportion of accurate (error-free) measurements is between 0.2 and 1.0, but with a noncompliance rate of around 20%, the more likely range is between 0.6 and 1.0. This implies that the results from RD regression should be divided by a factor between 0.6 and 1 (or equivalently, multipled by a factor between 1 and 1.67).

Although it is common in RD studies to report sharp and fuzzy RD estimates, this paper will focus on the sharp estimates, i.e. the effect of being *assigned* to treatment rather than the "treatment effect on the treated". I focus on the sharp RD results for two reasons. First, other than for governments in close proximity to the $100 million threshold and presented in Figure 1, there is no data on compliance. Second, even with compliance data within a $10 million bandwidth of the $100 million threshold, the relatively small size of the discontinuity (due to both measurement error and noncompliance) leads to a weak first stage and IV estimates that are imprecise. As a result, while I present fuzzy RD results for the $100 million threshold in an appendix (with the assumption of 100 percent compliance outside of the $10 million bandwidth), the discussion will focus primarily on the intent-to-treat results.[9]

3.2.3 Panel Methods

Although the RD approach outlined above estimates a LATE rather than an ATE, the comprehensive nature of the Census data means that it is also possible to analyze the effect of GASB 34 using panel methods. The existence of panel data strengthens the RD analysis in two ways. First, it provides a means of validating the RD results using difference-in-difference

[9] Battistin et. al (2009) show that, in the case of Fuzzy RD, the bias factor cancels out, and that the traditional instrumental variable estimator identifies the treatment effect on the treated (the causal effect, G, of implementing GASB 34) without any correction term being necessary. The fuzzy RD estimator is:

$$\frac{E\{Y|R_{obs} = \bar{r}^{+}\} - E\{Y|R_{obs} = \bar{r}^{-}\}}{E\{G|R_{obs} = \bar{r}^{+}\} - E\{G|R_{obs} = \bar{r}^{-}\}} \tag{4}$$

designs; by limiting the difference-in-difference designs to a subsample of governments near the two thresholds, the average treatment effects that they estimate should be somewhat comparable to the LATEs at the two different thresholds. Second, and perhaps more importantly, it allows for the estimation of an average treatment effect that uses all of the governments in the sample. This is crucial because it provides some basis for comparing the local effects at the two thresholds to an overall average effect. Hence, the RD and the panel methods in combination illuminate the extent to which the effects of GASB 34 are heterogeneous.

The downside of the Census panel data is that, as noted above, it lacks information on borrowing costs, and the challenges of matching the census data to the primary market data limited the collection of borrowing costs to governments near the two thresholds ($10 million and $100 million in revenues). Hence, while the RD methods will examine both the use of debt and the cost of debt at the two different thresholds, the panel analysis will be limited to analyzing just the use of debt.

All of the methods will focus on the year after the initial implementation of the reporting model; since it takes several months for governments to issue financial statements following the close of the fiscal year, the effects of the reporting change would not have been felt immediately.[10] For phase 1 governments, the year after implementation was the first fiscal year beginning after June 15, 2002 (fiscal year 2003 for most governments). For phase 2 governments it was the first fiscal year beginning after June 15, 2003. It is also possible to look for effects in other years; since some governments implemented the standard early, there may have been an early treatment effect. And to the extent that there is an effect of having received the treatment for one year longer (a difference of treatment *duration*), there may also be evidence of a treatment effect in later years. Results for additional years are discussed

[10]For the cost of debt, I looked for debt issues with transaction dates as close as possible to 9 months following implementation. However, in order to match the census data with cost of debt information, I drew from a dataset that included debt issues from the two years following implementation, and hence some of the issues in the final sample fell more than one year after implementation. See Appendix Part 2 for more details.

and presented in Appendix Table 2, but they do not form a large part of the discussion that follows.

4 Results

4.1 Regression Discontinuity Results

Figure 2a plots the proportion of governments around the phase 1/phase 2 threshold that issued long-term debt in the year following the initial implementation of GASB 34. The figure shows clear evidence of a treatment effect, with governments just above the threshold approximately 12 percentage points less likely to issue debt. Figure 2b presents the results for the amount of new debt.[11] Once again, there is a difference in the intercepts at the threshold, with phase 1 governments issuing approximately 4 million less debt than phase 2 governments.

Table 2 presents summary statistics. Table 3 presents the regression results that correspond to Figures 2a and 2b. All of the estimates come from nonparametric regressions, using mean-squared-error optimal bandwidths and robust standard errors based on Calonico, Cattaneo, and Titiunik (2014); relying on an "optimal" data-driven bandwidth prevents specification searches that may yield false positives. Each table presents three specifications: a linear model without covariates, a linear model with covariates, and a quadratic model with covariates. Standard RD estimation does not require covariates, but they can increase efficiency (Calonico, Cattaneo, Farrell, and Titiunik, 2016). The covariates include the amount of new debt issued in 1998, total taxes collected in 1998, the amount of cash and securities on hand in 1998, and a year dummy indicating the survey year (2003 or 2004). The financial variables are from 1998 to ensure that the covariates could not have been affected by the treatment.

The first three columns in Table 3 focus on the likelihood of issuing long term debt

[11] In order to minimize the effect of outliers, the sample excludes observations near the cut-off with amounts in the top/bottom 1 percent of the entire distribution.

at the phase 1/phase 2 threshold. Although the dependent variable in this case is binary, it is standard in RD to use a linear probability model since the running variable is unrestricted on both sides of the cut-off (Lee and Lemieux, 2010; Lemieux and Milligan, 2008). The results are consistent with Figure 2a. There appears to be a difference in the likelihood of issuing long term debt between governments that implemented the standard early (above the threshold) and those that did not, with phase 1 governments near the threshold approximately 17-24 percentage points less likely than phase 2 governments to issue debt in the year after they implemented the standard. Two of the three estimates are significant at the 5 percent level.

Columns 4-6 in Table 3 present the results for the amount of new debt. The regressions again yield estimates that are consistent with the graphical analysis. Governments above the threshold (phase 1 governments) issued approximately \$6 million less in new debt than phase 2 governments near the threshold, with two of the three estimates significant at the five percent level.

How large are these effect sizes? As noted in section 4, these effects are subject to a correction factor and should be divided by a number approximately between 0.6 and 1 in order to properly account for the measurement error in the running variable. For the probability of issuing debt, the coefficients in Table 3 fall between 16.5% and 24.8%, implying that the true effect of GASB 34 lies somewhere in the range of 16.5-41% (inflating the larger estimate by the maximum amount). Since the correction factor is almost certainly inflated somewhat by the presence of noncompliance, a more realistic estimate might be to take the middle estimate in the table and divide it by 0.8, yielding an estimate of approximately 22%. The corresponding estimates for the amount of new debt is 10 million. For context, the average debt load for governments within a \$50 million window of the threshold (\$50 - \$150 million in total revenues) in the year following phase 1 implementation was \$85 million in total debt outstanding. Hence, a decline of 10 million in new debt represents roughly 12% of the average total debt load.[12]

[12] Appendix Tables 3 presents the fuzzy RD results that account for noncompliance at the phase 1/2 threshold. As noted above, these estimates assume that there is full compliance outside of the band for which

Figure 3 presents RD results for the effect of GASB 34 on the cost of debt.[13] There appears to be a discontinuity in the outcomes – with phase 1 governments at the threshold facing higher borrowing costs – though the discontinuity is not as large relative to the variance in the data as in Figure 2. Table 4 presents the corresponding regression results. The results are all positive, indicating that governments that implemented the standard early faced higher borrowing costs. However, the coefficients lack precision, with only one of three estimates significant at the five percent level.

The results from Tables 3 and 4 suggest that phase 1 borrowers reduced the amount of debt they issued in the wake of GASB 34 and also faced potentially higher borrowing costs. These results hold at the $100 million threshold since they are local average treatment effects. What about results from the lower part of the revenue distribution? Figures 4 and 5 present RD plots for the use of debt and the cost of debt at the lower threshold. Tables 5 and 6 present the corresponding regression results. All of the estimates for the use of debt are very close to zero and not statistically significant. The estimates for the cost of debt are also not significant, though they are consistently negative, i.e. in the opposite direction of the results at the phase 1/2 threshold. It appears that, while GASB 34 reduced the use of debt and raised borrowing costs among larger governments, it had little to no effect on smaller governments near the $10 million threshold; if anything, for those governments it may have lowered borrowing costs, which would be consistent with the results of Cassar et. al (2015).

4.2 Robustness and Placebo Tests

Appendix Table 1 presents the results of additional specifications for both the use of the debt and the cost of debt at the phase 1/2 threshold. Since the main results all rely on optimal bandwidths based on Calonico, Cattaneo, & Titiunik (2014), the table present results for alternative bandwidth choices, as well as from the local randomization approach. All of

data has been collected ($90 - $110 million) and thus must be interpreted with caution. The fuzzy estimates are extremely noisy and not statistically significant, but 5 of 6 carry the same sign as the sharp RD results.

[13]The range of the x-axis is slightly smaller than Figure 2 because of the more limited amount of data.

the additional specifications are in line with the main estimates, although once again the coefficients for the true interest cost are imprecise, with only one significant at the five percent level.

While these additional specifications demonstrate that the results are not sensitive to the choice of bandwidth or functional form, they do not speak to the validity of the RD design in a setting such as this one with measurement error in the running variable. To further validate the use of the RD design and ensure that the results are not the product of misspecification, Figure 6 plots the use of debt – both a binary measure (Figure 6a) and a continuous measure (Figure 6b) – for fiscal year 1998. Given that there was no treatment in 1998, there should be no evidence of a discontinuity at the threshold. While there is a small difference in the intercepts in Figure 6a, it is much smaller than the discontinuity observed in Figure 2a, and the two lines in Figure 6b nearly intersect, together providing support for the validity of the design. Table 7a presents the corresponding regression results. All of the estimates are near zero and not statistically significant. Even the largest estimate (column 6) is considerably smaller than any of the estimates in Table 3. Hence, the results support the validity of the RD design as an empirical strategy.

4.3 Panel Data Results

The RD results above provide estimates of the local average treatment effect at the two different policy thresholds – \$100 million and \$10 million in revenues respectively – but they don't reveal what the average treatment effect of the policy is across all governments in the sample. This section exploits the longitudinal nature of the Census data to 1) provide additional validation of the RD results, and 2) estimate an average treatment effect.

Figure 7a depicts the proportion of phase 1 and phase 2 governments issuing long-term debt three years before and three years after GASB 34 implementation (corresponding to fiscal years 1999-2005 for most governments). For the sake of comparability with the RD estimates, the sample is limited to governments within a \$50 million window of the \$100 million threshold (i.e. phase 1 governments with between $100 and $150 million in total revenues and phase 2 governments with between $50 million and $100 million in total revenues).[14] Up until the year of implementation (year 0), phase 1 and phase 2 governments follow a similar trajectory, with the two lines roughly parallel. In the year following implementation (year 1), there is a drop in the proportion of phase 1 governments issuing debt relative to phase 2 governments and relative to their prior trajectory.

The visual evidence thus corroborates the RD findings from the previous section. It shows that phase 1 governments were less likely to issue debt in the wake of GASB 34 implementation than they had been previously. To further investigate, Table 9 shows the results from regressions that estimate the treatment effect using a difference-in-difference framework. That is, the regressions estimate the difference in phase 1 governments relative to phase 2 governments in year 1 after accounting for pre-existing differences between the groups. The difference-in-difference results are smaller than the RD results presented in Table 3, but they have the same sign and two of the four estimates are statistically significant. Thus, the pattern of results – both in the figure as well as in the additional regression findings – represents additional evidence that GASB 34 reduced the amount of debt that phase 1 governments were issuing, further bolstering the findings from the previous sections. In appendix table 3, I also present additional diff-in-diff specifications that use an alternative control group consisting of governments from non-GAAP compliant states. These results also support the main findings.

Figure 7b plots the use of debt among all governments in the sample: phases 1-3. When plotted in this way, with a large y-axis, it is difficult to observe any obvious effect of the treatment. It appears that there is perhaps a slight drop in debt issuance for phase 2 governments in year 2 (the year after implementation), however no other obvious effects are discernible. Table 10 presents panel regression results for all governments. The

[14]These estimates will not be perfectly analagous to the RD estimates since they still represent an average treatment effect over the $50-$100 million range rather than at the $100 threshold, but they should nevertheless provide some basis for comparison.

sample includes observations from years 1998-2005 and only includes governments with a full set of observations. In order to be consistent with the RD and diff-in-diff results, which only use observations from specific parts of the distribution, table 10 presents results for an unscaled measure of new debt; however, appendix table 5 presents results for a scaled measure of new debt that first transforms the distribution so that it is entirely positive and then converts to the log scale.[15] The specifications include government and year fixed effects. A treatment variable ("post-GASB 34 implementation") is coded one in the years following implementation of GASB 34 and zero otherwise. Thus, for phase 1 governments, the treatment variable is coded as a one for the fiscal years beginning after June 15, 2002, and zero otherwise; for phase 2, it is coded as a one for the fiscal year beginning after June 15, 2003, and zero otherwise. For phase 3 governments, it is coded as a one for the fiscal year beginning after June 15, 2004.

The results from Table 10 are uniformly positive and, in four of the six specifications, statistically significant. The estimates suggest that, when looking at the entire population of governments in the U.S., for the *average* municipal borrower the effect of GASB 34 was slightly positive. Notably, the standard errors are approximately one tenth the size of those from the RD estimates, suggesting that the lack of finding at the phase 2/3 threshold may be due to a lack of precision.

5 The Credit Ratings Channel

The results in Table 10 appear consistent with the prior literature on accounting quality: the introduction of accrual accounting reduces information asymmetries and lowers the cost of capital. This in turn should increase debt use for the average borrower. Based on Table 10, it appears that GASB 34 had a slightly positive effect on the use of debt among governments overall. This does not however explain the findings at the higher threshold. How is it possible

[15]These results are necessarily sensitive to the transformation that is used. Alternative transformations, such as subtracting values three standard deviations below the mean and then converting to the log scale, yield similar results.

for GASB 34 to have had a positive effect on debt use overall but a negative effect among larger governments? Recent papers in the finance literature provide some insight.

Bolton, Freixas, and Shapiro (2012) model competition among credit-rating agencies (CRAs) to show how conflicts of interest distort the efficiency of ratings. Specifically, they show that "when an issuer is more important to a CRA, either because it is a repeat issuer or because it has larger issues, the CRA is more prone to inflate that issuer's ratings" (88). This theory has since been validated empirically. He, Qian, and Strahan (2011) study the role of rating agencies in the mortgage-backed securities market and find that "ratings mistakes were correlated with issuer size and market conditions. All three major rating agencies were more optimistic for securities sold by large issuers during the boom years." Kedia, Rajgopal, and Zhou (2014) show that, following Moody's IPO in 2000, "large and frequent issuers of corporate bonds [were] more likely to receive more favorable ratings from Moody's." Similarly, Faltin-Traeger (2009) finds that repeat issuers are more likely to stick with the same CRA if they received a more favorable early rating. Credit ratings may be biased for those segments of the market with more market power, in this case large municipalities that consistently issue long-term debt.

Thus, it appears plausible that inflated ratings explain the negative effects of GASB 34 on large issuers. To the extent that credit ratings and high quality accounting information both decrease the incentives for investor information acquisition, then they should lower the cost of debt. However, if credit rating information is biased for a segment of the market, then the higher quality accounting information should shed light on that bias. In other words, if large issuers were more likely to be inflated, then they should see larger ratings declines once higher quality information was introduced. In this section, I use my sample to look for evidence consistent with this hypothesis. To do so, I use a simple difference-in-difference design, based on credit ratings before and after GASB 34 financial reports were released, to examine whether Phase 1 and Phase 2 governments experienced differential changes in the ratings of their new bond issues.

The results are presented in Table 11. The credit rating variable is based on Moody's underlying rating (without insurance) for primary market bond issues. The ratings are superimposed on a linear scale, so that a top rating (Aaa) = 1, Aa1 = 2, etc. In cases where the Moody's rating is missing, I impute using the S&P or Fitch ratings. The table shows that the phase 1 governments in the sample had slightly better ratings on average prior to GASB 34 (i.e. a lower value on the numeric scale), and slightly worse ratings after. While the phase 2 ratings got slightly better, the phase 1 ratings got worse. The difference in the rating swings between the two groups - 0.8 on the linear scale - is significant at the 5 percent level, suggesting that phase 1 governments did indeed experience a worsening of their credit ratings following the implementation of GASB 34.

This evidence is not definitive. In particular, due to the difficulty of matching the bond data with the Census of Governments, the D-in-D design in this case uses only one year's worth of prior data and consequently relies on the strong assumption that the pre-treatment trends are parallel (St.Clair and Cook, 2015). Moreover, many of the observations in the SDC Platinum database are missing information on credit ratings, and so there is the possibility that the credit rating information is not missing at random. Nevertheless, the results are consistent with other work showing that different levels of informativeness across large and smaller issuers may have led to differential impacts from the introduction of higher quality accounting information.

While this conclusion requires certain assumptions about investor behavior – namely that investors are somewhat naive in their reliance on credit ratings – other findings in the literature appear to support this view. Cornaggia, Cornaggia, and Israelsen (2018) document that, not only do muni bond investors mechanically rely on credit ratings for independent information about credit risk, but they rely on ratings more in poor information environments, as would have been the case prior to GASB 34. With few other sources of information, investors may have placed relatively heavy weight on ratings, causing them to underestimate the credit risk of larger borrowers.

5.1 Alternative Explanations

Are there any other plausible explanations that might account for the difference in findings across large and small governments? One possibility is that larger governments with a greater number of funds might have revealed more proprietary information to contracting parties, for example when negotiating procurement contracts. However, prior to GASB 34, proprietary funds, which are the funds used to account for "business-like" enterprises in government that recover their costs through user fees, were already using accrual accounting in their income statements (called at the time, the "Combined Statement of Revenues, Expenses, and Changes in Fund Equity") and consequently yielded much less new information as a result of GASB 34 than the governmental funds.

Another possible explanation would be that larger governments are more likely to administer their own pension funds, compared with smaller government that generally participate in state-wide programs. However, as noted in section 2, pensions and OPEB were not covered by GASB 34, but by later statements. Hence, this explanation too falls short. Finally, perhaps accounting standards are less strictly enforced for smaller governments. The literature certainly suggests that smaller governments are less likely to comply with GAAP (Khumawala et al, 2014). However, this would not explain why the effects for larger governments were negative. Based on previous findings, one would expect that higher compliance would result in more positive effects.

Higher quality accounting information should only cause debt use to decline and debt costs to rise if there is something biased in the existing public signal. Given recent findings about investors' naive use of credit ratings (Cornaggia, Cornaggia, and Israelsen, 2018) and the greater market power that larger governments possess (He, Qian, and Strahan, 2011), the difference in quality between the credit ratings of large and small governments appears to be the most likely explanation.

6 Discussion and Conclusion

This paper exploits the staggered roll-out of a new financial reporting model to examine the effect that the introduction of accrual accounting had on municipal debt issuance. Unlike previous research on financial accounting standards, much of which relies on cross-sectional regressions and strong identification assumptions, the analysis here uses regression discontinuity methods to more cleanly identify the effects of the change at two points in the revenue distribution. Panel methods also provide estimates of the average treatment effect. The results suggest that, while the average effect of GASB 34 was either negligible or slightly positive insofar as it increased debt usage, the effect on large issuers – those with approximately $100 million in revenues in 1999 – was in the opposite direction, resulting in less use of debt and higher interest costs. Phase 1 governments at the threshold were approximately 20 percentage points less likely to issue long-term debt in the wake of implementation and issued $10 million less on average. Theory and empirical evidence suggest that the negative effects for large issuers were due to inflated credit ratings.

These findings contribute to a large body of literature on accounting quality and disclosure in financial markets. Other than a few studies, much of this work has focused on the average effect. This paper, in contrast, shows that there may be substantial heterogeneity in the effects and that local effects may differ substantially from the average effect if the quality of the existing public information varies across parts of the distribution.

In addition to highlighting the heterogeneity of the effects, this paper also draws a link between credit information and accounting quality. Accrual information and credit ratings are both forms of public information; more accurate ratings as well as higher quality accounting information decrease the incentives for investor information acquisition and improve measures of market quality (Piccolo and Shapiro, 2017). The information contained in credit scores can substitute for the incremental information provided by accrual accounting (Cassar, Ittner, and Cavaluzzo, 2015), and in fact, investors rely on credit ratings the most in the poorest information environments (Cornaggia, Cornaggia, and Israelsen, 2018). The contribution that this paper makes is document that, in such environments, the benefits of accrual accounting depend on the quality of existing information; if credit information is not just noisy but *biased* for a certain segment of the market, and lenders rely on that information, then the introduction of high quality accounting information can actually increase the cost of debt and decrease borrowing.

Several limitations of this study bear noting. Despite using the largest available dataset on U.S. government finances, the findings are hampered in places by a lack of precision, as the data demands of the RD design are substantial. This lack of power as well as a lack of compliance data limited the analysis to an "intent-to-treat" analysis rather than an analysis of the "treatment effect on the treated." The lack of precision is even more evident in the analyses of the cost of debt since this data was collected separately and had to be matched with the data on government borrowing. Future work will ideally be able to build on the findings presented here by investigating a wider variety of outcomes and utilizing a larger sample of borrowing cost data.

References

Ashcraft, Adam, Paul Goldsmith-Pinkham, and James Vickery. 2010. "Credit Ratings and the Mortgage Credit Boom." Federal Reserve Bank of New York Staff Reports, no. 449.

Ashcraft, Adam, Paul Goldsmith-Pinkham, and James Vickery. 2011. "Credit Ratings and Security Prices in the Subprime MBS Market." *American Economic Review: Papers and Proceedings*, 101(3), 115-119.

Baber, William R., and Gore, Angela K. 2008. "Consequences of GAAP Disclosure Regulation: Evidence from Municipal Debt Issues." *The Accounting Review* 83(3), 565-591.

Battistin, Erich, Agar Brugiavini, Enrico Rettore, and Guglielmo Weber. 2009. "The Retirement Consumption Puzzle: Evidence from a Regression Discontinuity Approach." *American Economic Review*, 99(5), 2209-2226.

Benson, Earl D., Barry R. Marks, and K. K. Raman. 1991. "The Effect of Voluntary GAAP Compliance and Financial Disclosure on Governmental Borrowing Costs." *Journal of Accounting, Auditing & Finance* 6(3), 303-319.

Bolton, Patrick, Xavier Freixas, and Joel Shapiro. 2012. "The Credit Ratings Game." *Journal of Finance*. 67(1), 85-111.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82(6), 2295–2326.

Calonico, Sebastian, Matias D. Cattaneo, Max Farrell, and Rocio Titiunik. 2016. "Regression-Discontinuity Designs Using Covariates." Working Paper.

Cassar, Gavin, Christopher D. Ittner, and Ken S. Cavalluzzo. 2015. "Alternative Information Sources and Information Asymmetry Reduction: Evidence from Small Business Debt." *Journal of Accounting and Economics*, 59(2-3), 242-263.

Cornaggia, Jess, Kimberly J. Cornaggia, and Ryan D. Israelsen. 2018. "Credit Ratings and the Cost of Municipal Financing." *Review of Financial Studies*, 31(6), 2038-2079.

Davezies, Laurent, and Thomas Le Barbanchon. 2017. Regression Discontinuity Designs with Continuous Measurement Error in the Running Variable. IZA Discussion Paper No. 10801.

Faltin-Traeger, Oliver. 2009. "Picking the Right Rating Agency: Issuer Choice in the ABS Market." Working paper, Columbia Business School.

Fairchild, L., and T. Koch. 1998. "The Impact of State Disclosure Requirements on

Municipal Yields." *National Tax Journal* 51(4), 733–753.

Governmental Accounting Standards Board (GASB). 1999. "Basic Financial Statements-and Management's Discussion and Analysis-for State and Local Governments." Statement No. 34. Norwalk, CT.

Goldstein, Itay and Liyan Yang. 2017. "Information Disclosure in Financial Markets. *Annual Review of Financial Economics*, 9, 101-125.

Gore, Angela K. 2004. "The Effects of GAAP Regulation and Bond Market Interaction on Local Governmnet Disclosure." *Journal of Accounting and Public Policy* 23, 23-52.

He, Jie, Jun Qian, and Philip E. Strahan. 2011. "Credit Ratings and the Evolution of the Mortgage Backed Securities Market."

Horowitz, Joel L., and Charles F. Manski. 1995. "Identification and Robustness with Contaminated and Corrupted Data." *Econometrica*, 63(2), 281-302.

Hullegie, Patrick, and Tobias J. Klein. 2010. "The Effect of Private Health Insurance on Medical Care Utilization and Self-Assessed Health in Germany." *Health Economics*, 19, 1048-1062.

Kedia, Simi, Shivaram Rajgopal, and Xing Zhou. 2014. "Did Going Public Impair Moody's Credit Ratings?" *Journal of Financial Economics*, 114, 293-315.

Khumawala, Saleha, Justin Marlowe, and Daniel G. Neely. 2014. "Accounting Professionalism and Local Government GAAP Adoption: A National Study." *Journal of Public Budgeting, Accounting, & Financial Management*, 26(2), 292-312.

Lee, David S., and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48, 281-355.

Lemiuex, Thomas and David Milligan. 2008. "Incentive Effects of Social Assistance: A Regression Discontinuity Approach." *Journal of Econometrics*, 142(2), 807-828.

Minnis, Michael. 2011. "The Value of Financial Statement Verification in Debt Financing: Evidence from Private U.S. Firms. *Journal of Accounting Research*, 49 (2), 457–506.

Pei, Zhuan, and Yi Shen. 2016. "The Devil is in the Tails: Regression Discontinuity Design with Measurement Error in the Assignment Variable." Working Paper.

Piccolo, Alessio, and Joel Shapiro. 2017. "Credit Ratings and Market Information." Working Paper.

Pierson, Kawika, Michael L. Hand, and Fred Thompson. 2015. "The Government Finance

Database: A Comprehensive Resource for Quantitative Research in Public Financial Analysis." *Plos One*, 10(6), e0130119.

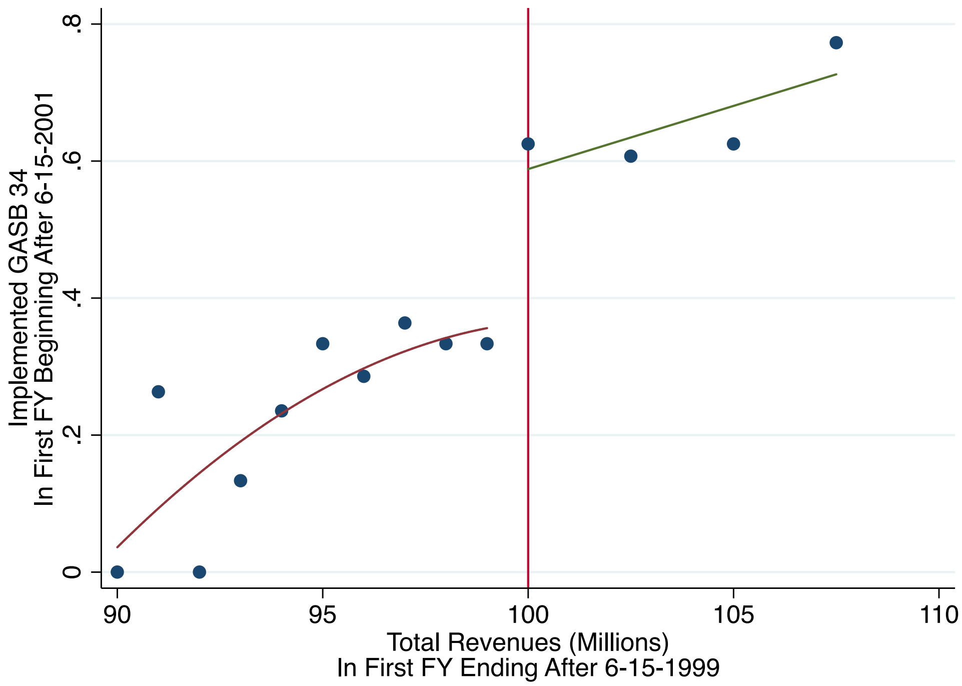
Plummer, Elizabeth, Paul D. Hutchison, and Terry K. Patton. 2007. "GASB No. 34's Governmental Financial Reporting Model: Evidence on Its Information Relevance." *The Accounting Review* 82(1), 205-240.

Reck, Jacqueline L., and Earl R. Wilson. 2014. "The Relative Influence of Fund-Based and Government Wide Financial Information on Municipal Bond Borrowing Costs." *Journal of Government & Nonprofit Accounting* 3, 35-57.

Sekhon, Jasjeet S., and Rocio Titiunik. 2017. "On Interpreting the Regression Discontinuity Design as a Local Experiment." Regression Discontinuity Designs: Theory and Applications (Advances in Econometrics, Volume 38), M. D. Cattaneo and J. C. Escanciano (ed.), Emerald Publishing Limited, 1-28.

St.Clair, Travis, and Thomas D. Cook. 2015. "Difference-in-Differences Methods in Public Finance." *National Tax Journal* 68(2), 319-338.

Figure 1: Compliance with GASB 34 at Phase 1/Phase 2 Threshold



Note. This figure shows graphical evidence of a discontinuity in the implementation of GASB 34. The figure indicates the percentage of governments near the phase 1/phase 2 threshold that implemented GASB 34 in the first fiscal year beginning after June 15, 2001 (fiscal year 2002 for most governments). Data collected by the author. N = 217. The circles represent local sample means.

Figure 2: Use of Debt at Phase 1/Phase 2 Threshold

Figure 2a: Issued Long Term Debt

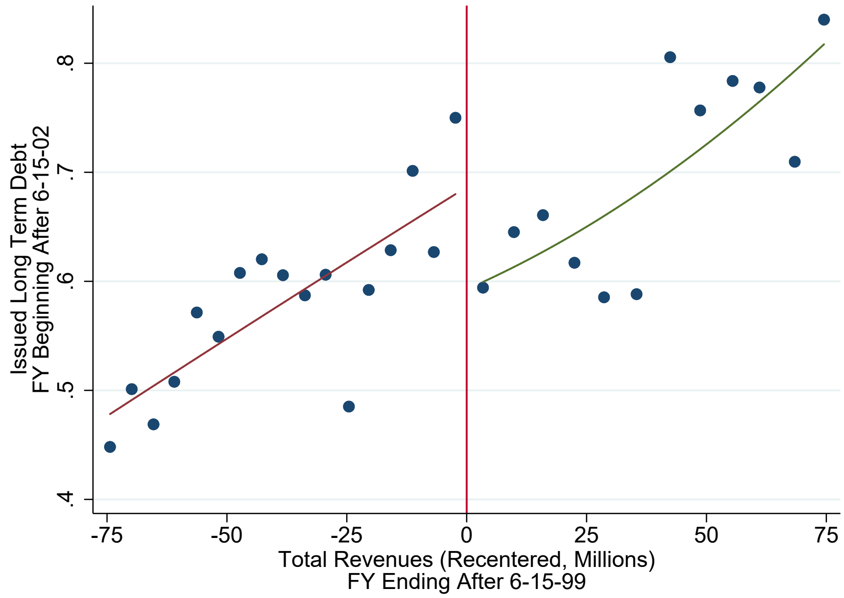
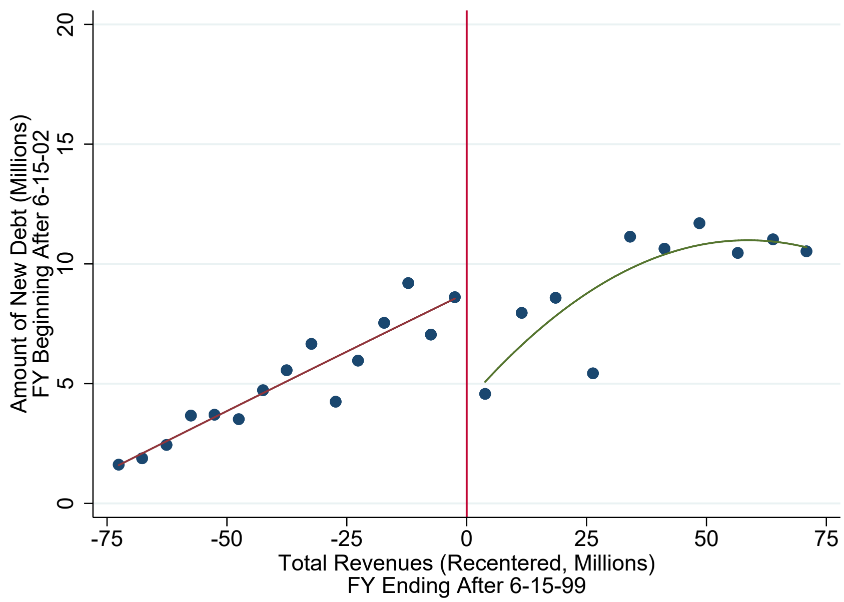
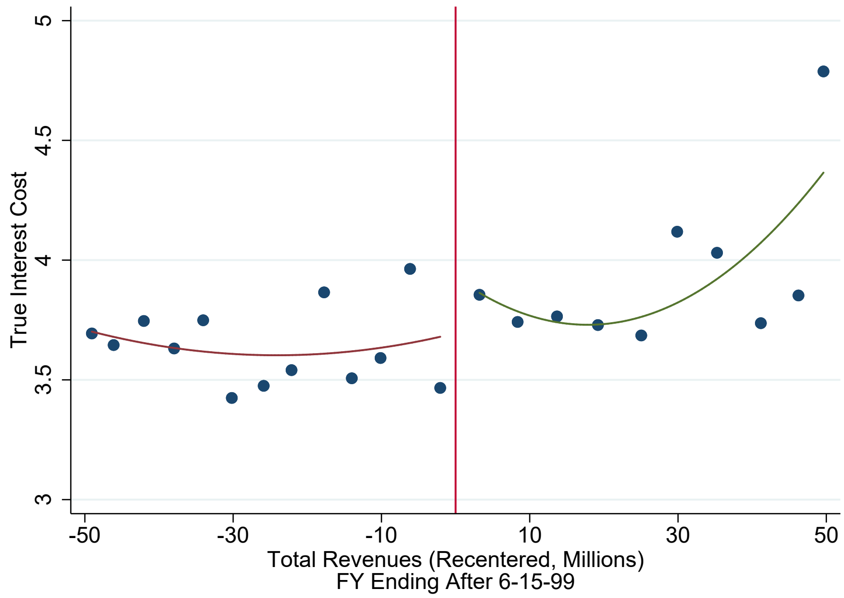


Figure 2b: Amount of New Debt



Note. The figures shows graphical evidence of discontinuities in the use of debt at the Phase 1/Phase 2 threshold. Phase 1 governments just above the threshold issue less debt than Phase 2 governments just below the threshold. The circles represent local sample means.

Figure 3: True Interest Cost at Phase 1/Phase 2 Threshold



Note. The figure shows the cost of debt on both sides of the Phase 1/Phase 2 threshold. Although the data are more sparse, there is slight evidence of a discontinuity in the cost of debt. Phase 1 governments just above the threshold face slightly higher interest rates (as measured by true interest cost) than Phase 2 governments just below the threshold. The circles represent local sample means.

Figure 4: Use of Debt at Phase 2/Phase 3 Threshold

Figure 4a: Issued Long Term Debt

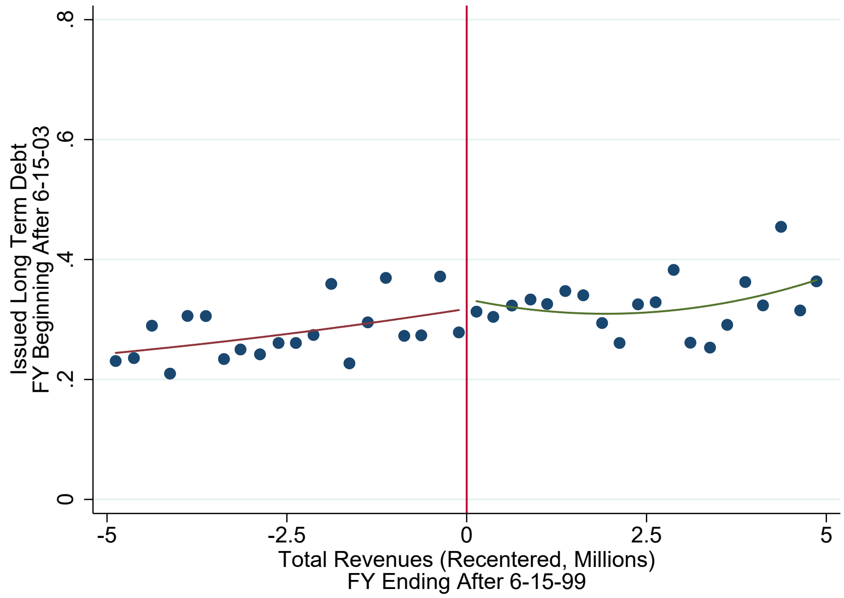
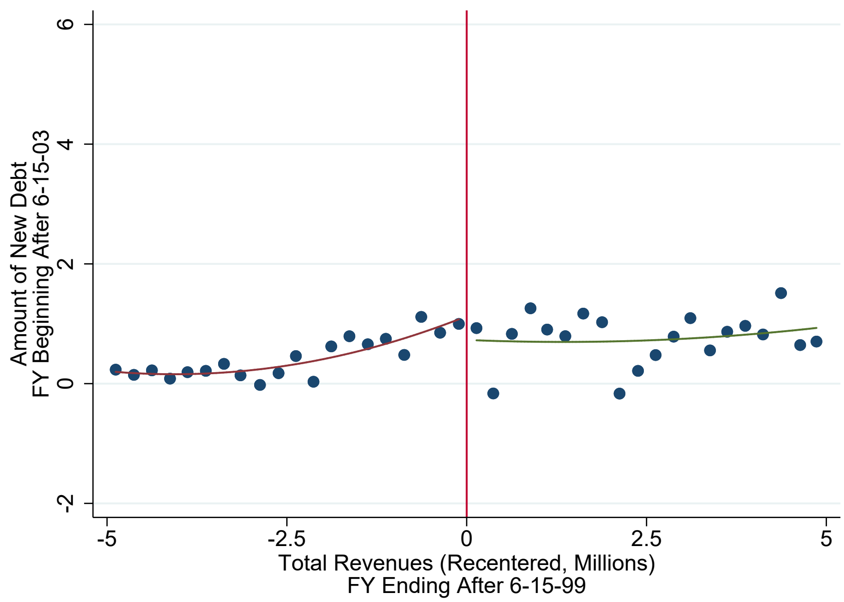
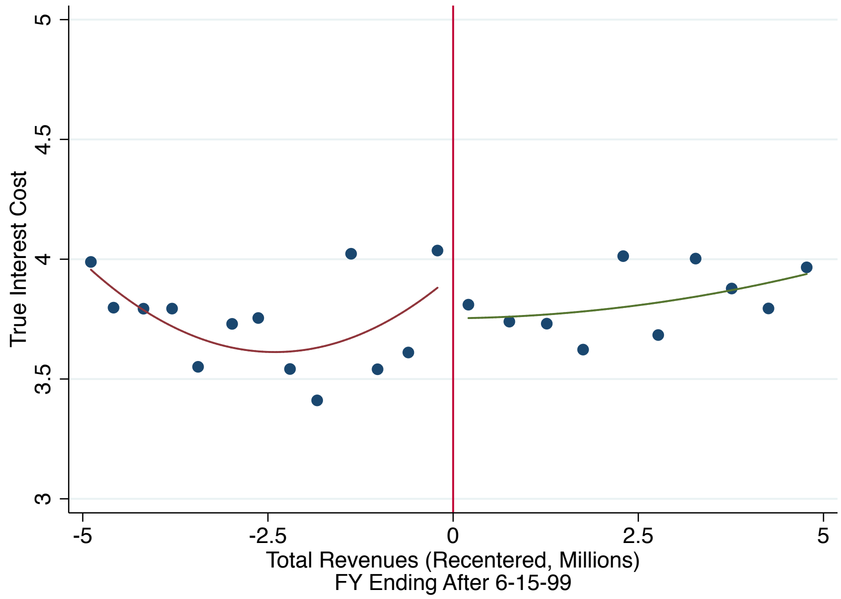


Figure 4b: Amount of New Debt



Note. The figures show the use of debt on both sides of the Phase 2/Phase 3 threshold. There is no evidence of a discontinuity in outcomes. The circles represent local sample means.

Figure 5: True Interest Cost at Phase 2/Phase 3 Threshold



Note. The figures show the cost of debt on both sides of the Phase 2/Phase 3 threshold. There is no evidence of a discontinuity in outcomes. The circles represent local sample means.

Figure 6: Placebo Results (1998) for the Use of Debt at Phase 1 / Phase 2 Threshold

Figure 6a: Placebo - Issued Long Term Debt

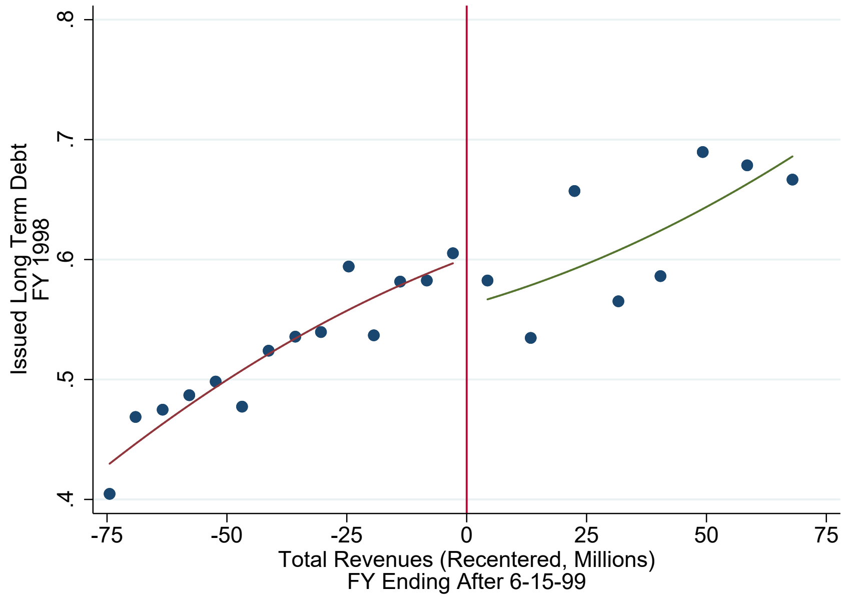
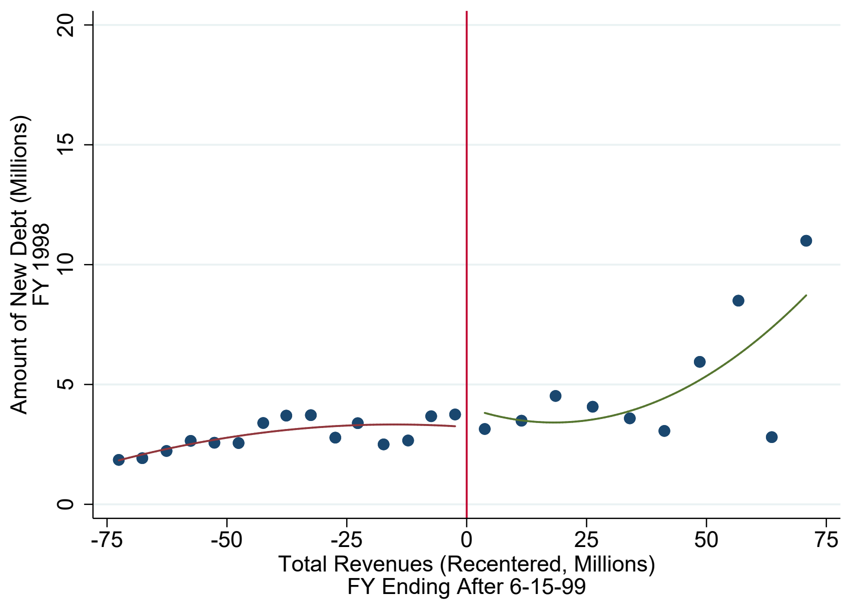


Figure 6b: Placebo - Amount of New Debt



Note. The figures show the use of debt in 1998 on both sides of the Phase 1/Phase 2 threshold. Since GASB 34 was issued in 1999, there should be no evidence of a treatment effect prior to 1999. Although there is a slight discontinuity in Figure 6a, there is no evidence of a discontinuity in Figure 6b, providing some confirmation that the placebo tests do not find evidence of a treatment effect. Table 7 presents the corresponding regression results.

Figure 7: Panel Data Analysis. Proportion of Governments Issuing Long Term Debt Before and After GASB 34 Implementation

Figure 7a: Phase 1 and Phase 2 Govs Near $100 Million Threshold

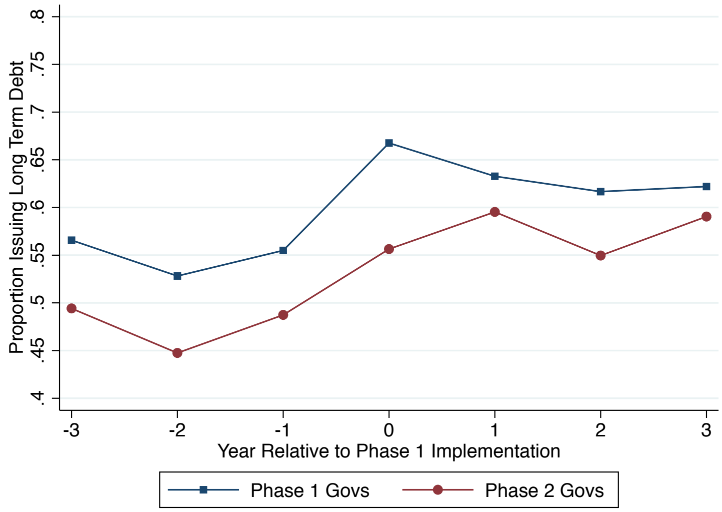
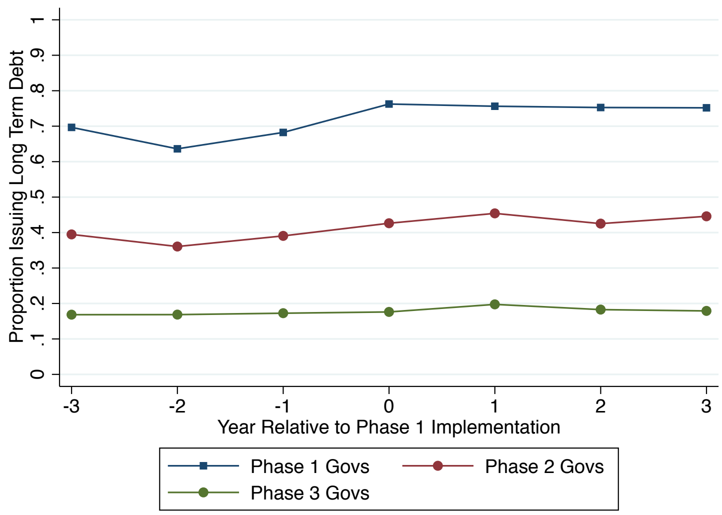


Figure 7b: All Governments



Note: The figures illustrate trends in the use of long-term debt before and after Phase 1 governments were required to implement GASB 34. Year 0 represents the first fiscal year beginning after June 15, 2001. While Figure 4a includes only those governments near the $100 Million Threshold, Figure 4b includes governments of all sizes. For Figure 4a, the sample consists of a panel of governments within $50 million of the $100 million threshold (phase 1 governments with $100-$150 million in total revenues, phase 2 governments with $50-$100 million in total). Figure 4a illustrates that phase 1 governments see a decline in long term debt issuance relative to their prior trend in the year after GASB 34 implementation (year 1). In contrast, when all governments are included in the sample, there is no obvious effect on debt issuance. Both figures exclude governments without a full set of observations for years -3 to +3 (corresponding to fiscal years 1999-2005 for governments with a July 1 fiscal year). N = (4a) 1401 governments, (4b) 13,562 governments, including 1124 phase 1, 5715 phase 2, and 6723 phase 3.

Table 1. Percentage of Governments that Implemented GASB 34 by Distance from $100 Million Revenue Threshold

Distance (Millions)	-10	-7.5	-5	-2.5	+2.5	+5.0	+7.5	+10.0
Total Revenues (Millions)	$90-$92.5	$92.5-$95	$95-$97.5	$97.5-$100	$100-$102.5	$102.5 - $105	$105 - $107.5	$107.5 - $110
Percentage	14%	16%	29%	36%	63%	61%	63%	77%
Observations	37	37	17	28	24	28	24	22

Note: The table shows the data depicted in Figure 1: the number and percentage of governments that implemented GASB 34 in the first fiscal year beginning after June 15, 2001 (fiscal year 2002 for most governments). The first row indicates the distance from the $100 million revenue threshold. Data collected by the author from audited financial statements.

Table 2: Summary Statistics for RD Analyses

Phase 1/2 Threshold	(1)	(2)	(3)	(4)	(5)
FY beginning after 6/15/02	mean	median	sd	min	max
Use of Debt					
Issued Long-Term Debt	0.6	1.0	0.5	0.0	1.0
Amount of New Debt (Millions)	7.4	0.0	20.0	-19.1	102
Total Revenue(99-00) (Millions)	97.4	95.5	11.3	80.0	120
Amount of New Debt(1998) (Millions)	5.7	-0.2	36.0	-96.7	482
Total Taxes (1998) (Millions)	31.4	30.4	19.6	0.0	89.8
Total Cash (1998) (Millions)	73.7	39.5	133.5	0.0	1,445
Survey Year 2004 Indicator	0.3	0.0	0.5	0.0	1.0
Cost of Debt					
True Interest Cost	3.7	3.8	0.8	1.6	5.8
Years to Maturity	16.6	15.4	10.8	1.0	99.0
Callable	0.7	1.0	0.5	0.0	1.0
Phase 2/3 Threshold	(1)	(2)	(3)	(4)	(5)
FY beginning after 6/15/03	mean	median	sd	min	max
Use of Debt					
Issued Long-Term Debt	0.3	0.0	0.5	0.0	1.0
Amount of New Debt (Millions)	0.8	-0.1	5.0	-18.8	78.2
Total Revenue (99-00) (Millions)	9.9	9.8	1.2	8.0	12.0
Amount of New Debt(1998) (Millions)	0.5	0.0	3.6	-33.9	86.9
Total Taxes (1998) (Millions)	2.8	2.4	2.0	0.0	10.4
Total Cash (1998) (Millions)	5.5	2.4	16.3	0.0	208
Survey Year 2005 Indicator	0.2	0.0	0.4	0.0	1.0
Cost of Debt					
True Interest Cost	3.7	3.8	0.7	2.1	4.8
Years to Maturity	14.8	14.2	9.0	1.0	99.0
Callable	0.7	1.0	0.5	0.0	1.0

Note: This table provides summary statistics for governments within a bandwidth of 20 million (N = 488) around the Phase 1/2 threshold (top panel) and a bandwidth of 2 million (N = 1,648)around the Phase 2/3 threshold (bottom panel). Data on the use of debt come from the Census of Governments. The number of observations in the full datasets are 15,189 (first fiscal year after 6-15-02) and 15,889 (first fiscal year after 6-15-03). Data on the cost of debt come from SDC Platinum. Appendix part 2 details the process used to match the census data with debt issues from SDC Platinum in order to produce the cost sample.

Table 3. Use of Debt at Phase 1 / Phase 2 Threshold

	Issued Long Term Debt (Binary)			Amount of New Debt		
	(1)	(2)	(3)	(4)	(5)	(6)
RD Estimate	-0.173*	-0.165	-0.248*	-4,983	-6,036*	-7,831*
	(0.0885)	(0.0902)	(0.126)	(3,064)	(3,078)	(3,805)
N	896	807	884	854	732	953
Bandwidth	Optimal	Optimal	Optimal	Optimal	Optimal	Optimal
Functional Form	Linear	Linear	Quadratic	Linear	Linear	Quadratic
Covariates	No	Yes	Yes	No	Yes	Yes

Note. ** $p < 0.01$, * $p < 0.05$. The table presents regression discontinuity results for the effect of GASB 34 on the use of debt in the first fiscal year beginning after 6-15-02. Specifications 1-3 presents results from a linear probability model where the dependent variable is a binary indicator for whether a government issued long-term debt in the previous year. Specifications 4-6 presents results for the amount of new debt, measured as the sum of new long term debt plus year-ending short-term debt minus the amount of long-term debt retired. The coefficients for specifications 4-6 are in thousands. The bandwidths and standard errors are MSE-optimal and robust respectively based on Calonico, Cattaneo, & Titiunik (2014). Covariates include new debt in 1998, total taxes in 1998, total cash and securities in 1998, and a year dummy indicating the survey year (2003 or 2004). N represents the number of observations within the bandwidth.

Table 4: Cost of Debt at Phase 1 / Phase 2 Threshold

	True Interest Cost		
	(1)	(2)	(3)
RD Estimate	0.253	0.153	0.539**
	(0.518)	(0.143)	(0.203)
N	119	128	124
Bandwidth	Optimal	Optimal	Optimal
Functional Form	Linear	Linear	Quadratic
Covariates	No	Yes	Yes

Note. ** $p < 0.01$, * $p < 0.05$. The table presents regression discontinuity results for the effect of GASB 34 on the true interest costs of government debt issues at the phase 1/2 threshold. See appendix part 2 for a description of how cost data was merged with information on government borrowers. The bandwidths and standard errors are MSE-optimal and robust respectively based on Calonico, Cattaneo, & Titiunik (2014). Covariates include the number of years to maturity, whether the debt issue is callable, and time fixed effects (quarter year intervals). N represents the number of observations within the bandwidth.

Table 5: Use of Debt at Phase 2/ Phase 3 Threshold

	Issued Long Term Debt			Amount of New Debt		
	(1)	(2)	(3)	(4)	(5)	(6)
RD Estimate	0.00249	0.0199	0.0405	-270.7	-300.4	-240.2
	(0.0392)	(0.0405)	(0.0679)	(396.4)	(483.6)	(725.9)
N	3,854	3,496	3,536	4,808	3,481	3,432
Bandwidth	Optimal	Optimal	Optimal	Optimal	Optimal	Optimal
Functional Form	Linear	Linear	Quadratic	Linear	Linear	Quadratic
Covariates	No	Yes	Yes	No	Yes	Yes

Note. ** $p < 0.01$, * $p < 0.05$. The table presents regression discontinuity results for the effect of GASB 34 on the use of debt in the first fiscal year beginning after 6-15-03. Specifications 1-3 presents results from a linear probability model where the dependent variable is a binary indicator for whether a government issued long-term debt in the previous year. Specifications 4-6 presents results for the amount of new debt, measured as the sum of new long term debt plus year-ending short-term debt minus the amount of long-term debt retired. The coefficients for specifications 4-6 are in thousands. The bandwidths and standard errors are MSE-optimal and robust respectively based on Calonico, Cattaneo, & Titiunik (2014). Covariates include new debt in 1998, total taxes in 1998, total cash and securities in 1998, and a year dummy indicating the survey year (2004 or 2005). N represents the number of observations within the bandwidth.

Table 6: Cost of Debt at Phase 2 / Phase 3 Threshold

	True Interest Cost		
	(1)	(2)	(3)
RD Estimate	-0.221	-0.108	-0.222
	(0.377)	(0.134)	(0.161)
N	89	202	202
Bandwidth	Optimal	5,000	5,000
Functional Form	Linear	Linear	Quadratic
Covariates	No	Yes	Yes

Note. ** $p < 0.01$, * $p < 0.05$. The table presents regression discontinuity results for the effect of GASB 34 on the true interest cost for governments at the phase 2/3 threshold. See appendix part 2 for a description of how cost data was merged with information on government borrowers. The bandwidths and standard errors are MSE-optimal and robust respectively based on Calonico, Cattaneo, & Titiunik (2014). Columns 2 and 3 use a default bandwidth of 5000 because there are not enough observations above the threshold to calculate an optimal bandwidth. Covariates include the number of years to maturity, whether the debt issue is callable, and time fixed effects (quarter year intervals). N represents the number of observations within the bandwidth.

Table 7. FY 1998 Placebo Estimates for Use of Debt

Table 7a. Phase 1 / Phase 2 Threshold

	Issued Long Term Debt (Binary)			Amount of New Debt		
	(1)	(2)	(3)	(4)	(5)	(6)
RD Estimate	0.0136	0.00994	0.00328	-755.6	-530.2	-1,280
	(0.0861)	(0.0816)	(0.100)	(1,974)	(1,917)	(2,166)
N	939	988	1,566	895	989	1652
Bandwidth	Optimal	Optimal	Optimal	Optimal	Optimal	Optimal
Functional Form	Linear	Linear	Quadratic	Linear	Linear	Quadratic
Covariates	No	Yes	Yes	No	Yes	Yes

Table 7b. Phase 2 / Phase 3 Threshold

	Issued Long Term Debt (Binary)			Amount of New Debt		
	(1)	(2)	(3)	(4)	(5)	(6)
RD Estimate	-0.0281	-0.0239	0.0205	-189.8	-158.4	-284.4
	(0.0504)	(0.0505)	(0.0837)	(267.9)	(271.0)	(336.8)
N	2,539	2,527	2,527	2,509	2,513	3,595
Bandwidth	Optimal	Optimal	Optimal	Optimal	Optimal	Optimal
Functional Form	Linear	Linear	Quadratic	Linear	Linear	Quadratic
Covariates	No	Yes	Yes	No	Yes	Yes

Note. ** $p < 0.01$, * $p < 0.05$. The tables present regresion discontinuity results for the effect of GASB 34 on the use of debt in 1998. Since GASB 34 was issued in 1999, there should be no evidence of a treatment effect prior to 1999. The tables confirm that there is no evidence of a discontinuity in outcomes. The coefficients for new debt are in thousands. The bandwidths and standard errors are MSE-optimal and robust respectively based on Calonico, Cattaneo, & Titiunik (2014). Covariates include total taxes in 1998 and total cash in 1998. N represents the number of observations within the bandwidth.

Table 8: Summary Statistics for Diff-in-Diff and Panel Analyses

Diff-in-Diff Analysis	(1)	(2)	(3)	(4)	(5)
FYs beginning after 6/15/97 - 6/15/02	mean	median	sd	min	max
Use of Debt					
Issued Long-Term Debt	0.5	1.0	0.5	0.0	1.0
Amount of New Debt (Millions)	3.6	-0.2	13.3	-29	83.5
Total Expenditure (Millions)	90.8	82.9	36.6	0.0	568
Total Taxes (Millions)	30.9	27.7	22.3	0.0	318

Panel Data Analysis	(1)	(2)	(3)	(4)	(5)
1998-2005	mean	median	sd	min	max
Use of Debt					
Issued Long-Term Debt	0.3	0.0	0.5	0.0	1.0
Amount of New Debt (Millions)	1.7	0.0	8.3	-16.6	83.5
Log Expenditures	9.3	9.4	1.8	0.0	19.2
Log Taxes	7.4	7.8	2.9	0.0	18.4

Note: This table provides government-year level summary statistics for the difference-in-difference (top panel) and panel data analyses (bottom panel). The data come from the Census of Governments. The diff-in-diff analysis is limited to governments with total revenues between $50 and $150 million (N = 1,385 governments, 9,695 observations). For the panel data analysis, N = 12,887 governments (1,114 phase 1, 5,435 phase 2, 6,338 phase 3) and 103,096 observations., Both the diff-in-diff and panel data analyses include observations from five years prior to implementation through one year after.

Table 9: Difference-in-Difference Results - Phase 1/Phase 2 Threshold

	Issued Long Term Debt			Amount of New Debt		
	(1)	(2)	(3)	(4)	(5)	(6)
D-in-D	-0.0253	-0.0959**	-0.0998**	-721.7	-3,796**	-3,793**
Estimate	(0.0280)	(0.0343)	(0.0342)	(1,151)	(1,342)	(1,337)
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Group Time Trends	No	Yes	Yes	No	Yes	Yes
Covariates	No	No	Yes	No	No	Yes
Observations	9,695	9,695	9,689	9,495	9,495	9,489

Note. **p $<$ 0.01, * p $<$ 0.05. The table presents difference-in-difference estimates from specifications of the form: $Y_{it} = \alpha + \beta_1 post_t + \beta_2 post * phase1_{it} + \beta_3 X + \gamma_i + \epsilon_{it}$ where $post_{it}$ represents the year after implementation (fiscal year beginning after 6/15/02), $phase1$ represents phase 1 governments (the treatment group), and γ_i represent fixed effects (which subsume the variable $phase1$.). The table presents results for the coefficient of interest, β_2. The sample is limited to governments with total revenues between \$50 and \$150 million and only includes governments with a full set of observations for the five years prior to implementation through the one year after (year -5 to +1, corresponding approximately to 1997-2003). The treatment effect is measured for the year following implementation, as in Tables 3 and 4. The covariates are total expenditures and total taxes. Standard errors are clustered by government.

Table 10: Panel Estimation Results - All Governments

	Issued Long Term Debt			Amount of New Debt		
	(1)	(2)	(3)	(4)	(5)	(6)
Post-GASB 34	0.0254**	0.00790	0.00625	452.1**	389.5*	370.2*
Implementation	(0.00332)	(0.00648)	(0.00646)	(76.27)	(169.5)	(169.0)
Gov Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	No	Yes	Yes	No	Yes	Yes
Covariates	No	No	Yes	No	No	Yes
Observations	103,096	103,096	103,096	100,578	100,578	100,578

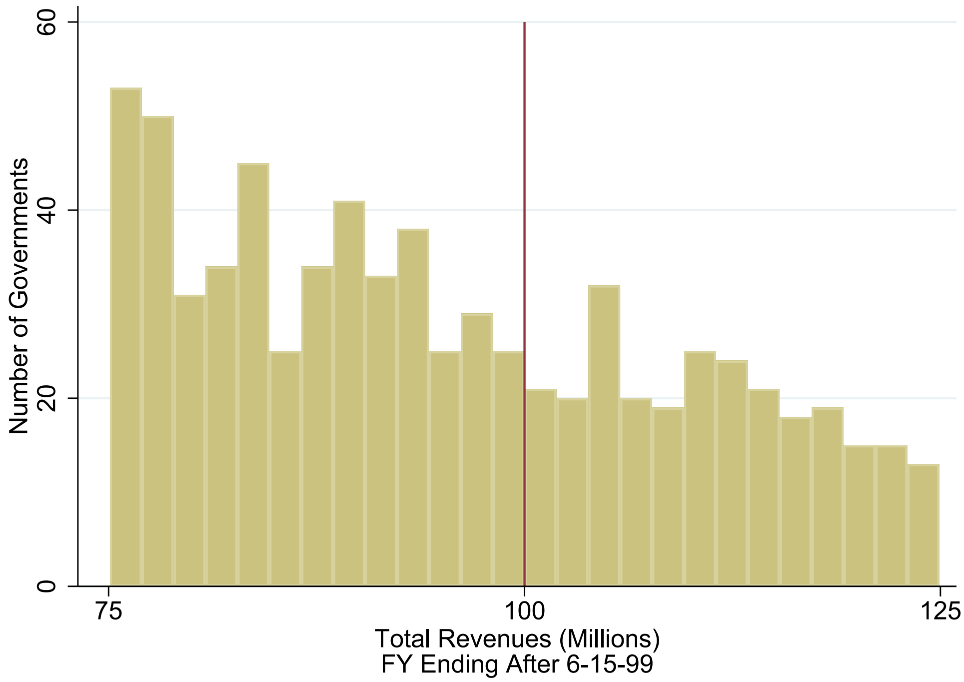
Note. **p $<$ 0.01, * p $<$ 0.05. The table presents panel estimation results for specifications of the form: $Y_{it} = \alpha + \beta_1 post_{it} + \beta_2 X + \gamma_i + \epsilon_{it}$ where $post_{it}$ represents the year following GASB 34 implementation (the first FY beginning after 6-15-02, 6-15-03, and 6-15-04 for Phase 1, 2, and 3 governments respectively). The table presents results for the coefficient of interest, β_1. The variation in $post_{it}$ (across types of government and over time) allows for both government and year fixed effects. The sample includes observations from 1998-2005 and only includes governments with a full set of observations. The covariates are log total expenditures and log taxes (recoded to zero when the untransformed variable is equal to zero). Standard errors are clustered by government.

Table 11: Credit Ratings Before and After GASB 34

	Before	After	Difference
Treatment Group (Phase 1)	4.2	4.6	0.5
	[45]	[96]	
Comparison Group (Phase 2)	4.6	4.3	-0.3
	[91]	[216]	
D-in-D Estimate			0.77*
(SE)			(0.39) [448]

Note. * $p < 0.05$. The rating variable is based on Moody's underlying (without insurance) ratings (Aaa = 1, Aa1 = 2, etc.) but is imputed based on S&P and Fitch ratings where the Moody's rating is missing. A positive coefficient indicates a worsening credit rating. The brackets indicate the number of observations. The standard error for the D-in-D estimate is in parentheses and is clustered by issuer. The "before" period consists of the first fiscal year beginning after 6/15/2001. The sample consists of primary market data from SDC Platinum's municipal database for the period 6/15/2002-5/31/2005 that has been matched to the Census of Governments. Unlike the process described in Appendix part 2, which matches governments to one debt issue in the post-treatment period, in this case governments are matched to multiple debt issues. The sample is restricted to governments for whom a "pre-treatment" observation exists in the data, i.e. governments that issued debt in the first fiscal year beginning after 6/15/2001.

Appendix Figure 1: Density Plot of Governments



Note. The figure provides a density plot of governments to check for manipulation of the running variable. Since GASB issued the standard in June of 1999, and implementation was based on revenues in the fiscal year ending after June 15, 1999 there would have been little opportunity for governments to manipulate their total revenues. The figure confirms that there is no obvious bunching to the left of the $100 million revenue threshold.

Appendix Table 1. Additional RD Specifications for the Use and Cost of Debt at the Phase 1/Phase 2 Threshold

Appendix Table 1a. Use of Debt

	Issued Long Term Debt			Amount of New Debt		
	(1)	(2)	(3)	(4)	(5)	(6)
RD Estimate	-0.408*	-0.219+	-0.120+	-12,385*	-9,110*	-3,428
	(0.191)	(0.132)	(0.056)	(5,160)	(4111)	(2,743)
N	298	615	190	291	601	186
Bandwidth	12,500	25,000	Optimal	12,500	25,000	Optimal
Functional Form	Linear	Linear	Local Randomization	Linear	Linear	Local Randomization
Covariates	Yes	Yes	Yes	Yes	Yes	Yes

Appendix Table 1b. Cost of Debt

	True Interest Cost		
	(1)	(2)	(3))
RD Estimate	0.383*	0.104	0.115
	(0.176)	(0.146)	(0.202)
N	125	240	68
Bandwidth	12,500	25,000	Optimal
Functional Form	Linear	Linear	Local Randomization
Covariates	Yes	Yes	Yes

Note. **p < 0.01, *p < 0.05, + p <0.10. The tables are identical to Tables 3 and 4 respectively in the main text except they includes results from specifications using alternate bandwidths and also for a specification using the local randomization assumption. The coefficients for specifications 4-6 in Table 1a are in thousands, as are all the alternative bandwidths. The standard errors are robust based on Calonico, Cattaneo, & Titiunik (2014). The covariates for the use of debt include new debt in 1998, total taxes in 1998, total cash and securities in 1998, and an indicator variable for the survey year (2003 or 2004). The covariates for the cost of debt include the number of years to maturity, whether the debt issue is callable, and time fixed effects (quarter year intervals). N represents the number of observations within the bandwidth.

Appendix Table 2. Use of Debt at Phase 1/ Phase 2 Threshold - Results from Additional Years

Appendix Table 2a. Issued Long-Term Debt

	Year t-2 (1)	Year t - 1 (2)	Year t (3)	Year t + 1 (4)	Year t + 2 (5)
RD Estimate	-0.117 (0.0900)	-0.0788 (0.0899)	-0.0262 (0.0903)	-0.165+ (0.0902)	-0.0554 (0.0784)
N	736	730	885	807	1297
Bandwidth	Optimal	Optimal	Optimal	Optimal	Optimal

Appendix Table 2b. Amount of New Debt

	Year t - 2 (1)	Year t - 1 (2)	Year t (3)	Year t + 1 (4)	Year t + 2 (5)
RD Estimate	-1,470 (2,938)	-2,124 (3,187)	-2,428 (3,826)	-6,036* (3,078)	-3,280 (3,668)
N	549	496	667	732	747
Bandwidth	Optimal	Optimal	Optimal	Optimal	Optimal

Note. ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$. These tables look at the effect of the treatment at different points in time, from two years before Phase 1 governments were required to implement GASB 34 to two years after. Although Phase 1 governments were required to implement GASB 34 in the first fiscal year after June 15, 2001 (2002 for most governments), the standard was issued in 1999, and thus there is the possibility of an early treatment effect. The column labeled Year t+1 represents the year after implementation; this is the focus of the main analysis and corresponds to the third column of Tables 6 and 7. By year t+2, both Phase 1 and Phase 2 governments had implemented the standard and presumably released financial reports; hence, any significant effect in this year would be due to a different in treatment *duration* rather than simply exposure to the treatment. None of the specifications other than the main estimates - those in year t+1 - are significant, suggesting that there is no evidence of an early treatment effect. It also suggests that once Phase 2 governments also implemented GASB 34 and released financial statements (year t+2), there was no longer any difference in debt issuance. All specifications are linear in the running variable and include covariates. Bandwidths and the coefficients for new debt are in thousands. The bandwidths and standard errors are MSE-optimal and robust respectively based on Calonico, Cattaneo, & Titiunik (2014).

Appendix Table 3: Fuzzy RD Estimates for Use of Debt at Phase 1 / Phase 2 Threshold

	Issued Long Term Debt			Amount of New Debt		
	(1)	(2)	(3)	(4)	(5)	(6)
RD Estimate	-0.794	-0.746	-0.147	-24,266	-28,905	2,588
	(1.453)	(1.917)	(1.069)	(45,157)	(63,300)	(43,438)
N	181	180	180	177	177	177
Functional Form	Linear	Linear	Quadratic	Linear	Linear	Quadratic
Covariates	No	Yes	Yes	No	Yes	Yes

Note. ** $p < 0.01$, * $p < 0.05$. The table presents fuzzy regression discontinuity results ("treatment effect on the treated") at the Phase 1/ Phase 2 threshold for the effect of GASB 34 implementation on the use of debt. All estimates use a default bandwidth of \$10 million because compliance data was only collected within a \$10 million window of the threshold (\$90-\$110 million). Although the estimates are very imprecise and not statistically significant due to the small bandwidth and measurement error in the running variable, five of the six show the same sign as the coefficients from sharp RD estimation. The coefficients for new debt are in thousands. Standard errors are robust based on Calonico, Cattaneo, & Titiunik (2014).

Appendix Table 4: Difference-in-Difference Results Using Governments from Non-GAAP Compliant States as An Alternative Control Group

	Issued Long Term Debt			Amount of New Debt		
	(1)	(2)	(3)	(4)	(5)	(6)
D-in-D	-0.0912	-0.239**	-0.239**	-2,454	-6,701+	-6,528+
Estimate	(0.0639)	(0.0838)	(0.0839)	(3,068)	(3,747)	(3,804)
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Group Time Trends	No	Yes	Yes	No	Yes	Yes
Covariates	No	No	Yes	No	No	Yes
Observations	2,975	2,975	2,975	2,913	2,913	2,913

Note. **p < 0.01, * p < 0.05, + p < 0.10. The table reproduces the estimates from Table 8 except that it uses an alternate control group consisting of governments from non GAAP-compliant states. Thus, the sample consists of the same treatment group as Table 8 along with a comparison group that includes school districts in Indiana, Missouri, Washington, and West Virginia, as well as local governments in New Jersey and Indiana with total revenues between $100 and $150 million.

Appendix Table 5: Difference-in-Difference and Panel Results for Log New Debt Issued

	Diff-in-Diff			Panel Analysis		
	(1)	(2)	(3)	(4)	(5)	(6)
Estimate	-0.0295	-0.108*	-0.108*	0.00638*	0.00268	0.00197
	(0.0389)	(0.0433)	(0.0431)	(0.00305)	(0.00680)	(0.00679)
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Trend/FE	No	Yes	Yes	No	Yes	Yes
Covariates	No	No	Yes	No	No	Yes
Observations	9,494	9,494	9,494	100,577	100,577	100,577

Note. **p < 0.01, * p < 0.05. The table reproduces the estimates from columns 4-6 in Table 9 and columns 4-6 in Table 10 except it uses a transformed measure of new debt. The transformation first makes the entire distribution positive by subtracting the minimum value and then converts to the log scale.

Appendix Part 2: Process for Matching Census Data with Debt Issues

In this paragraph, I outline the process I used for matching the data from the Census of Governments with primary market data from SDC Platinum's municipal database. First, I performed searches in the SDC Platinum database for tax exempt issues. For the Phase 1/Phase 2 analysis, I collected primary market information for the period 6/15/2002 - 5/31/2005. For the Phase 2/Phase 3 analysis, I performed a separate search that yielded primary market information for the period 6/15/2003 - 5/31/2006. These periods cover the first two years following implementation for Phase 1 governments and Phase 2 governments respectively (for the full range of fiscal-year-end dates). Using these data sources, I matched each governments in the Census file with the corresponding debt issues based on the name of the issuer. In some cases, if a government did not issue any debt in the two years after the start of the fiscal year beginning after 6/15/2002 (Phase 1/2 threshold) or 6/15/2003 (Phase 2/3 threshold), then there was no match. In other cases, there were multiple debt issues that corresponded to a government. In that case, I matched the government to a debt issue using the following criteria. First, I looked for debt issues with complete information on true interest cost (no missing observations). Next, I looked for debt issues closest to 3/30/2003 (for the Phase 1/2 analysis) and 3/30/2004 (for the Phase 2/3 analysis). These dates fall nine months after implementation. Finally, I looked for debt issues with a time to maturity closest to ten years. If a municipality issued several bonds with identical maturities on the same day, I selected the largest of the issues.