ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Dharmapala, Dhammika

Working Paper

Estimating the Compliance Costs of Securities Regulation: A Bunching Analysis of Sarbanes-Oxley Section 404(b)

CESifo Working Paper, No. 6180

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Dharmapala, Dhammika (2016) : Estimating the Compliance Costs of Securities Regulation: A Bunching Analysis of Sarbanes-Oxley Section 404(b), CESifo Working Paper, No. 6180, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/149267

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

The international platform of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute

CESifo Working Papers

www.cesifo.org/wp

Estimating the Compliance Costs of Securities Regulation: A Bunching Analysis of Sarbanes-Oxley Section 404(b)

Dhammika Dharmapala

CESIFO WORKING PAPER NO. 6180
CATEGORY 11: INDUSTRIAL ORGANISATION
NOVEMBER 2016

An electronic version of the paper may be downloaded
- *from the SSRN website: www.SSRN.com*
- *from the RePEc website: www.RePEc.org*
- *from the CESifo website: www.CESifo-group.org/wp*

ISSN 2364-1428

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 6180*

Estimating the Compliance Costs of Securities Regulation: A Bunching Analysis of Sarbanes-Oxley Section 404(b)

Abstract

An extensive literature has analyzed the Sarbanes-Oxley (SOX) legislation enacted in 2002, but its social welfare consequences remain controversial. One of the most significant provisions of SOX – auditor attestation of internal controls under Section 404(b) – has been applied only to firms that have a public float (i.e. the market value of shares held by non-insiders) of at least $75 million. Public float is not reported in standard databases; this paper "scrapes" public float data from 10-K filings, resulting in about 161,000 observations at the firm-year level for the universe of reporting entities for fiscal years 1993-2015. The analysis compares the number of public float observations immediately below the $75 million threshold to a smooth counterfactual density estimated using a flexible polynomial fitted to the observed density away from the threshold. In the pre-SOX period, there is no detectable divergence between the actual and counterfactual density around $75 million. Since the enactment of SOX, there is substantial and statistically significant "bunching" below the threshold, exhibiting a bimodal pattern. The magnitude of bunching implies that firms around the threshold reduce their public float by about $1.7 million. This estimate is consistent with SOX 404(b) imposing a net compliance cost of about $4-$6 million (in present value terms) on firms in the region of the threshold. The results are robust to considering various alternative explanations for bunching, including the possibility that firms' insiders avoid crossing the threshold in order to continue extracting private benefits of control. The paper also explores potential financial and economic consequences of bunching, in particular on the debt financing and investment choices of firms close to threshold.

JEL-Codes: G380, K220.

Keywords: securities regulation, Sarbanes-Oxley, public float, compliance costs, bunching analysis.

Dhammika Dharmapala
University of Chicago Law School
USA - 60637 Chicago IL
dharmap@uchicago.edu

October 2016

I am grateful to Daniel Marcin (formerly of the Coase-Sandor Institute for Law and Economics at the University of Chicago Law School) for writing the Python code used to collect the data on public float. I also thank Miguel Almunia, Michael Carlos Best and Ben Marx for kindly sharing their Stata code for implementing the analysis of bunching, and Ofer Eldar, Feng Gao, Michael Guttentag, Todd Henderson, Birgit Hüsecken, Prasad Krishnamurthy, John Landon-Lane, Christian Leuz, Anup Malani, Ben Marx, Holger Spamann, Juan Carlos Suarez-Serrato, David Weber, Laura Wellman, Eleanor Wilking, participants at the American Law and Economics annual meetings and the International Institute of Public Finance annual Congress, and workshop participants at the University of Chicago and Rutgers University for helpful conversations and comments. I also acknowledge the financial support of the Lee and Brena Freeman Faculty Research Fund at the University of Chicago Law School. Any remaining errors or omissions are, of course, my own.

1) Introduction

An extensive literature across law, accounting, economics and finance has analyzed the compliance costs of securities regulation.[1] In the US context, recent literature in this area has focused especially on the Sarbanes-Oxley (hereafter, SOX) legislation. Enacted in 2002 in the wake of accounting scandals in the early 2000s (such as those involving Enron and Worldcom), SOX introduced an array of new measures, including requirements of enhanced internal controls with respect to firms' financial disclosures. Firms' executives were required to certify financial statements, and external auditors were to attest to the quality of these internal controls. Coates and Srinivasan (2014) provide a comprehensive recent review and assessment of the scholarly literature on the impact of SOX, spanning several academic disciplines. While substantial progress has been made in understanding the effects of SOX, they conclude that research on its net social welfare consequences remains inconclusive.

This paper brings both a novel dataset and a new empirical approach to bear on this important question. The empirical strategy exploits the behavior of firms close to the threshold for the application of SOX rules in order to shed light on the magnitude of the regulatory burdens. Many of the most significant provisions of SOX – in particular, auditor attestation of internal controls under Section 404(b) – have been applied only to firms at or above a threshold of $75 million of "public float" (i.e. the market value of shares held by non-insiders).[2] Although it is therefore a crucial concept for determining regulatory obligations, public float is not reported in standard financial databases. The previous literature has hand-collected public float data from firms' annual 10-K filings with the Securities Exchange Commission (SEC) for various subsamples of firms (e.g. Gao, Wu and Zimmerman, 2009: Iliev, 2010; Dharmapala and Khanna, 2016; Weber and Yang, 2016). This paper constructs a much larger dataset on public float by using Python code to "scrape" this information from firms' 10-K filings.[3] Using this method, we collect public float information for the universe of reporting entities for fiscal years 1993-2015. This period spans the SOX legislation and several subsequent changes in its implementation. The

[1] The empirical literature on securities regulation goes back to Stigler (1964). A particularly influential strand of research has used cross-country analysis (e.g. La Porta, Lopez de Silanes, and Shleifer, 2006).
[2] Insiders (or "affiliates") typically include officers, directors and large blockholders. Public float is also sometimes referred to as the "free" float.
[3] These filings are available through the SEC website at http://www.sec.gov/edgar/searchedgar/companysearch.html

resulting dataset contains nearly 161,000 observations at the firm-year level, although the analysis focuses on a subset of these that are relatively close to the $75 million threshold.

This paper applies to this new dataset an empirical technique that has not previously been used in the literature on securities regulation. It draws on a growing literature in economics analyzing "bunching" around bright-line tax and regulatory thresholds (e.g. Saez, 2010; Chetty et al., 2011; Kleven and Waseem, 2013; Almunia and Lopez-Rodriguez, 2015; Marx, 2015; for a general overview see Slemrod (2013), and for a theoretical analysis of such thresholds see Dharmapala, Slemrod and Wilson (2011)). In a setting such as ours that is characterized by a size-based threshold, a bunching analysis examines the divergence between the number of firms around the threshold and the counterfactual density (i.e. the number of firms that would be expected to be found around the threshold, absent the regulation). The latter is calculated by fitting a flexible polynomial function to the observed density of firms, excluding an interval close to the threshold. This approach imposes assumptions that the behavior of firms far from the threshold is not affected by its presence, and that the density of firms around the threshold would be smooth absent the threshold.

In the SOX context, if internal control attestation (and the associated regulatory enforcement) makes firms' financial disclosures more credible to investors, then firms may bunch just above $75 million. If the compliance costs of the regulation exceed the credibility (and other) benefits, then firms will bunch below the threshold.[4] Thus, bunching provides a direct source of evidence on firms' perceptions of the net value of the regulation that is difficult to obtain using other types of data or empirical approaches. The estimated magnitude of bunching allows inferences to be made about the extent to which firms are willing to change their public float in response to the regulation. The statistical significance of bunching is determined by computing standard errors using a bootstrapping procedure.

The SOX legislation enacted in 2002 did not include an exemption for smaller firms. However, on June 5, 2003, the SEC announced a phased-in implementation schedule under which firms with public float of at least $75 million in 2002 or thereafter (known as "accelerated filers" or AFs) were required to comply with SOX Section 404 (which required internal control

[4] There may be other explanations for bunching below the threshold, for instance the possibility that insiders may lose private benefits of control due to the regulations. These are discussed later in the paper (see in particular Section 5.6). In addition, there are also other regulatory consequences of crossing the threshold, such as the acceleration of filing deadlines. These issues are also discussed later in the paper (see e.g. Section 3).

mechanisms) starting in 2004. On the other hand, firms with public float below $75 million (known as "nonaccelerated filers" or NAFs) were granted a temporary delay in the date of expected compliance.[5] In 2007, all firms became subject to SOX Section 404(a), which required managers to report on the effectiveness of internal controls. However, NAFs continued to receive temporary exemption from SOX 404(b) – which required auditor attestation of internal controls – and this exemption was extended annually. In 2010, Congress enacted the Dodd-Frank Act, which (among other things) made the exemption of NAFs from SOX 404(b) permanent. Thus, NAFs have been exempt (either on an ostensibly temporary basis or permanently) from SOX 404(b) throughout the period that SOX has been in operation.

As discussed in the prior literature, firms have available a variety of mechanisms through which public float might potentially be "managed," such as insider purchases of stock or disclosures that reduce stock price (Gao, Wu and Zimmerman, 2009; Nondorf, Singer and You, 2012; Gao, 2016). Of course, each of these strategies entails some type of cost. As shown in a simple theoretical framework developed in Section 4 below, firms weigh these costs against the net costs of SOX 404(b) in determining what public float to report.

The empirical analysis aggregates observations of public float at the firm-year level into "bins" of equal width; this bin-level dataset measures the number of firm-year observations in each bin. The reported results use a bin width of $1 million, and focus on those bins ranging from $50 million to $150 million of public float. The counterfactual number of firms in bins around the $75 million threshold is estimated by fitting a flexible fifth-order polynomial function of public float to the observed density of firms, excluding an interval – $66-$83 million of public float – close to the threshold.

As 2003 was the first year in which firms could have adjusted their public float in response to SOX (Iliev, 2010), we divide the sample into pre-SOX (1993-2002) and SOX (2003-2015) periods of approximately equal duration. In the pre-SOX period (1993-2002), there is no detectable divergence between the actual and counterfactual density around $75 million.[6] After the enactment of SOX (2003-2015) there is a substantial "excess" number of firms below the $75 million threshold, along with a substantial number of "missing" firms immediately above the threshold,

[5] SEC release 33-8238, available at: http://www.sec.gov/rules/final/33-8238.htm

[6] There is also no evidence in the pre-SOX period of bunching around $57 million, an inflation-adjusted equivalent of the nominal $75 million threshold.

relative to a smooth counterfactual density. In contrast to the patterns observed in other bunching studies, the bunching region is bimodal – bunching is evident both immediately below $75 million and (to a greater extent) some distance below it (between $69 and $71 million). Bunching some distance from the threshold is explicable bearing in mind that although firms can manage their public float, they do not have full control over fluctuations in their stock price on the day that public float is determined (the last business day of the second fiscal quarter); they may thus choose a public float some distance from the threshold in order to provide a buffer in the event of an unexpected increase in stock price.

The basic analysis shows bunching below the threshold, implying that (at least for firms of this size) the costs perceived by firms from SOX 404(b) exceed the benefits.[7] The magnitude of the excess mass (which consists of 257 excess observations, relative to a counterfactual number of 1334 observations in the range $66-$75 million) implies that firms around the threshold reduce their public float by about $1.7 million on average (that is, averaged over all firm-years in the region, including those firms that are for various reasons indifferent to the threshold). This estimate is statistically significant, using bootstrapped standard errors.

The estimated missing mass above the threshold, however, is substantially smaller. To address this, we exploit the fact that we observe both the pre-SOX and SOX periods. In particular, we use the observed pre-SOX density away from the threshold to estimate the counterfactual density (rather than using the SOX period counterfactual density, as in the standard bunching methodology). This takes account of the possibility that SOX (or other factors that were different over 2003-2015) may have changed the density of firms in regions away from the threshold. This alternative approach leads to quite similar conclusions. However, it enables us to more satisfactorily reconcile the observed amount of bunching below the threshold with the observed missing region above the threshold.

We also address a number of potential alternative explanations. It is possible that negative market reactions upon crossing the threshold (as found in the prior literature) may increase the number of public float observations just below the threshold. However, the result is robust to modifying the threshold to account for this possibility. It is possible that some firms crossed the

[7] Bunching is most pronounced during the earlier years of the SOX period. However, it appears to persist over time to the end of the period, although firm conclusions are difficult to draw because of the limited number of observations within each subperiod.

threshold in the past and became AFs (even though their public float is now below the threshold). However, the result is robust to excluding such firms. We also seek to address the possibility that firms' insiders may avoid crossing the threshold to retain private benefits of control rather than to avoid compliance costs. The basic result is robust to omitting firms where private benefits may be particularly high (those that have dual-class stock or a high entrenchment index, from Bebchuk, Cohen and Ferrell (2009)).

Using a set of additional assumptions, it is possible to derive a rough estimate of the magnitude of the net compliance costs implied by the magnitude of bunching. The baseline estimate that firms reduce their public float by about $1.7 million entails a net compliance cost of approximately $4-$6 million for firms in the region of the threshold. This estimate is net of any benefits that SOX 404(b) may confer on these firms, and is in present value terms (over the time horizon – typically likely to be a few years – over which a firm may expect to remain below the threshold in the future). This estimate is broadly similar to those in the prior literature using other approaches.

The paper also provides some preliminary evidence on the financial and economic consequences of bunching behavior. We identify a set of firms that repeatedly report public float in the bunching region. In the SOX period, such firms appear to have higher debt-to-asset ratios (suggesting substitution from equity to debt financing) and lower investment-to-asset ratios (suggesting that bunching may exacerbate financial constraints). Further analysis of these consequences is left for future research.

This paper proceeds as follows. Section 2 reviews the relevant literature. Section 3 provides some background on SOX and its implementation, focusing on the role of the $75 million threshold. Section 4 introduces the bunching approach, and develops a simple theoretical model of firms' choice of public float and bunching behavior. Section 5 describes the data, empirical analysis, and results. Section 6 concludes.

2) Related Literature

There is a large empirical literature on the impact of SOX across several academic disciplines, as reviewed in Coates and Srinivasan (2014). For instance, Litvak (2007) analyzes market reactions to political events that increased the likelihood of SOX being enacted. While all US firms were expected to become subject to the new legislation if it were enacted, Litvak (2007)

compares a treatment group of foreign firms cross-listed in the US and subject to US securities law with a control group of firms in the same foreign country that were not cross-listed (or were cross-listed but not subject to US securities law). The pattern of market reactions implies that SOX was expected to reduce firm value on balance. Similarly, Zhang (2007) finds a decline in value around SOX-related events for US firms, relative to non-US firms. Chhaochharia and Grinstein (2007) also analyze market reactions to SOX, using as a control group those firms that already complied with the major provisions of SOX prior to its enactment. They find positive market reactions for large firms, but negative reactions for smaller firms, suggesting that compliance costs outweigh benefits for the latter. Kamar, Karaca-Mandic and Talley (2009) find that the enactment of SOX increased the propensity of small US firms to leave the public market through acquisition by private acquirers, using as a control the propensity of similar firms in non-US markets to exit in similar fashion.[8]

The literature in finance, accounting, economics and law that uses market reactions to SOX-related events provides powerful evidence of negative effects on smaller firms. However, Leuz (2007) and others argue that contemporaneous events that affected all US firms, or that may have affected treatment and control groups of firms differently, make it difficult to reach firm conclusions about the costs and benefits of SOX from event studies. In contrast, this paper does not rely on market reactions but rather on firms' responses to the threshold as a means of inferring their perceptions of the regulatory burden of SOX.

The most closely related prior paper is Iliev (2010), which also uses the $75 million threshold as a central element of a quasi-experimental research design.[9] However, Iliev (2010) uses a quite different approach: a regression discontinuity (RD) design that compares the value of firms just above the $75 million threshold for the application of SOX Section 404 in 2004 with the value of firms just below this threshold, hand-collecting public float for a sample of 301 firms. Using this RD approach, Iliev (2010) finds that SOX Section 404 reduced firm value. Iliev (2010) notes incidentally that there is some apparent bunching below the threshold. For Iliev's (2010) RD

[8] This paper is also related, more broadly, to quasi-experimental studies of the impact of securities law. For instance, Greenstone, Oyer, and Vissing-Jorgensen (2006) and Ferrell (2007) analyze the effects of the 1964 amendments that extended the disclosure requirements of US securities law to a subset of firms trading over-the-counter (OTC). Bushee and Leuz (2005) study the further extension of disclosure requirements in 1999 to the small firms that trade on the OTC Bulletin Board. This paper is also related, albeit more distantly, to quasi-experimental studies of broader corporate governance reforms in various countries (e.g. Black, Jang and Kim, 2006; Dharmapala and Khanna, 2013).
[9] Chhaochharia, Ott and Vig (2011) also analyze the $75 million threshold, but their focus is on its impact on mergers and acquisitions that would result in the threshold being crossed.

design, any manipulation of public float is a source of potential bias. To address this problem, Iliev (2010) instruments for 2004 values of public float using 2002 values (which could not have been manipulated in response to the SOX threshold). In contrast, the empirical approach in this paper is premised on the ability of firms to manipulate public float, and relies on the extent of this manipulation to provide insight into the compliance costs of the regulation.

There is a small literature in accounting that focuses directly on firms' management of public float. Gao, Wu and Zimmermann (2009) construct a treatment group of 806 firms that reported being NAFs in their 10-K filings with the SEC during the 2003-2005 period with a control group of 485 firms that reported being AFs and had market capitalization below $150 million over the same period. They find evidence that NAFs (when compared to the control group and to their own behavior in the pre-SOX period) engaged in various actions to remain below the $75 million threshold. These included reducing investment, increasing payout, and reducing the number of shares held by outsiders. They also find that NAFs disclose negative news and report lower accounting earnings in the second fiscal quarter (when public float is calculated).

Nondorf, Singer and You (2012) analyze a sample of 257 firms that have public float around the $75 million threshold. They find evidence of changes in ownership by insiders, and of the use of earnings management in financial reporting during the second fiscal quarter, that are consistent with attempts to remain below the threshold. The likelihood of firms remaining below the threshold is positively related to a measure of CEO power and negatively related to measures of monitoring, such as being audited by a major accounting firm. Gao (2016) explores in more detail how firms use their discretion in reporting public float, using a sample of 716 firms that had market capitalization between $75 million and $150 million over 2003-2006. Financial and ownership data for these firms is used to construct a benchmark public float, which is then compared to the reported public float (hand-collected from the SEC filings). Gao (2016) finds that firms that are predicted to face higher SOX 404 compliance costs count more shares as being "affiliated" in order to report a lower public float.

The accounting literature reviewed above thus finds evidence of firms manipulating public float in an apparent attempt to remain below the $75 million threshold. However, no previous study has used a formal bunching analysis to address this issue, and none has collected the universe of public float reports. There is an extensive accounting literature (beginning with Burgstahler and Dichev (1997)) that studies the density of reported earnings around relevant thresholds (such as

zero or analysts' forecasts). However, this literature does not formally estimate a counterfactual density or derive an estimate of the magnitude of manipulation.[10]

3) The Sarbanes-Oxley Legislation and the Role of the $75 Million Threshold

As discussed earlier, SOX was enacted in response to various accounting scandals in the early years of this century. The legislation contained a wide variety of different provisions. However, the central elements for the purposes of this paper are contained in Section 404. In particular, Section 404(a) requires management to provide an annual assessment of the issuer's internal controls over financial reporting.[11] Section 404(b) – the most important provision for our purposes – requires that the registered public accounting firm that audits the company's financial statements must provide an attestation of the firm's internal controls over financial reporting. The latter is generally thought to be the most costly SOX provision, especially for small firms, entailing additional audit fees as well as other types of costs (e.g. Alexander et al., 2013). It is thus on Section 404(b) that most of the empirical literature surveyed in Section 2 above has focused.

The SOX legislation did not include an exemption for smaller firms, and was initially expected to apply to all issuers on the US market. However, on June 5, 2003, the SEC announced a phased-in implementation schedule under which smaller firms would be permitted a longer period in which to move towards compliance with Section 404. Specifically, AFs (firms with public float of at least $75 million in 2002 or thereafter) were required to comply with SOX Section 404 starting in 2004. NAFs (firms with public float below $75 million) were granted a temporary delay in the date of expected compliance, initially until 2005.[12] The division of issuers into AFs and NAFs had been made earlier by the SEC in September 2002 in relation to new rules relating

[10] A number of other approaches have been used to analyze the consequences of SOX 404(b). For instance, Ge, Koester and McVay (2016) use a prediction model to compute the probability that firms that are exempt from 404(b) misreport the effectiveness of their internal controls.

[11] The SEC rules state that: "The internal control report must include: a statement of management's responsibility for establishing and maintaining adequate internal control over financial reporting for the company; management's assessment of the effectiveness of the company's internal control over financial reporting as of the end of the company's most recent fiscal year; a statement identifying the framework used by management to evaluate the effectiveness of the company's internal control over financial reporting" - see http://www.sec.gov/rules/final/33-8238.htm#ia

[12] SEC release 33-8238, available at: http://www.sec.gov/rules/final/33-8238.htm

to deadlines for filing 10-K and related forms with the SEC. AFs, as the term suggests, were subject to an earlier filing deadline than were NAFs.[13]

Public float is defined as the "aggregate worldwide market value of the voting and non-voting common equity held by its non-affiliates . . . as of the last business day of the issuer's most recently completed second fiscal quarter . . ."[14] Affiliates typically include managers, directors, and large blockholders (typically those owning 10% or more of the firm's shares, though sometimes a 5% threshold is used).[15] Thus, public float reflects the value of the firm that is held by outside shareholders.[16] Once a firm becomes an AF as a result of crossing the $75 million threshold, it remains one in the future unless its public float falls below $50 million.[17]

In 2005, the SEC extended the exemption from Section 404 for NAFs by an additional year. This extension was repeated in 2006.[18] In 2007, all firms became subject to SOX Section 404(a), which required managers to report on the effectiveness of internal controls. However, NAFs continued to receive temporary exemption from 404(b) (which required auditor attestation of internal controls). Also in 2007, the Public Company Accounting Oversight Board (PCAOB) introduced Accounting Standard 5 (AS5), with the apparent intention of lowering the compliance costs of Section 404(b) (Gao, 2016). The temporary exemption from Section 404(b) for NAFs was further extended in 2008 and 2009. In 2010, Congress enacted the Dodd-Frank Act, which (among other things) made the exemption of NAFs from SOX 404(b) permanent.

Thus, NAFs have been exempt (either on an ostensibly temporary basis or permanently) from SOX 404(b) throughout the period that SOX has been in operation. As the first SEC announcement of the $75 million threshold was on June 3, 2003 and public float is calculated as of the last business day of the second fiscal quarter (i.e. typically the last business day in June),

---

[13] Firms with over $700 million of public float – known as "large accelerated filers" (LAFs) – were subject to even earlier deadlines; however, the SOX 404 treatment of AFs and LAFs is very similar, so we do not emphasize the distinction between LAFs and AFs in this paper.

[14] 17 CFR 240.12b-2.

[15] An affiliate is defined as "a person that directly, or indirectly through one or more intermediaries, controls, or is controlled by, or is under common control with" the issuer (17 CFR 240.12b-2).

[16] Even if it meets the public float threshold, a firm is not defined as an AF until it has been required to file disclosures with the SEC for a year. Firms that have not satisfied this requirement and are close to the $75 million threshold will thus not face immediate regulatory consequences from crossing the threshold. This is unlikely to be of much importance in firms' choices, however, as such a firm would anticipate becoming subject to SOX 404(b) in the near future.

[17] This rule has applied since 2005; a stricter rule, requiring public float to fall below $25 million and certain other conditions to be met, applied before 2005.

[18] See e.g. Table 1 in Gao, Wu and Zimmerman (2009) and Figure 1 in Ge, Koester and McVay (2016) for a detailed timeline.

2003 was the first year in which firms could have adjusted their public float in response to SOX (Iliev, 2010). We thus divide the sample into a pre-SOX period (1993-2002) during which there was no opportunity nor incentive for manipulation in response to SOX, and a SOX period (2003-2015), during which firms in the region of the threshold may have used the various strategies discussed in Section 2 to manipulate their public float.

There are certain categories of firms that do not face any incentives to remain below the threshold. Until 2007, firms wishing to undertake seasoned equity offerings (SEOs) were permitted to use a simplified form (Form S-3) only if they had public float exceeding $75 million (e.g. Gao, 2016). This creates an incentive for firms undertaking SEOs to bunch above the $75 million threshold, and leads to a potential bias against the results we find. Also until 2007, AFs that are foreign private issuers (FPIs, a category of foreign companies defined by the SEC on the basis of the degree of US share ownership and business contacts) were temporarily exempt from SOX 404(b) and would have been indifferent to the threshold. This creates a potential bias against finding bunching under the threshold during those years. Some firms not subject to SOX 404(b) voluntarily complied with its provisions (e.g. Ge, Koester, and McVay, 2016) and so would be indifferent to the threshold, also creating a bias against our results.

In 2011 and 2012, various proposals to substantially increase the threshold for SOX 404(b) exemption were discussed in Congress (e.g. Ge, Koester, and McVay, 2016). The Jumpstart Our Business Startups (JOBS) Act, enacted on April 5, 2012, relaxed disclosure and compliance obligations for a new category of firms, known as “emerging growth companies” (EGCs), defined on the basis of revenue (and certain other characteristics) rather than public float, for a period of 5 years from their initial issuance of securities (e.g. Dharmapala and Khanna, 2016). Among other things, EGCs were permitted an exemption from SOX 404(b) for that 5-year period, even if they are AFs on the basis of their public float. As these firms are indifferent to the threshold, this also creates a potential bias against finding bunching under the threshold during those years.

The history of SOX 404 implementation is thus rather complex. For the purposes of this paper, the crucial point is that the $75 million threshold on which we focus has continued to be important throughout the 2003-2015 period. Three caveats are worth noting, however. First, AFs are subject (as described above) to accelerated deadlines for filings, and the timing of those rule changes does not permit us to easily distinguish between the impact of Section 404 and that of the filing deadlines. However, the filing deadlines appear to be relatively minor in their impact, and

not much ink has been spilt in practitioner or academic discussion of their effect (in stark contrast to the situation for SOX 404(b)). Second, "smaller reporting companies" are defined similarly to NAFs in terms of having public float below $75 million,[19] and have been eligible for scaled disclosure requirements since 2008. Thus, over the period since 2008, it is difficult to distinguish between SOX 404(b) effects and those of scaled disclosure. Third, until 2007 firms that became AFs thereby became subject to both Sections 404(a) and 404(b). Thus, it is difficult during the period until 2007 to distinguish between firms that reduced their public float in order to avoid Section 404(b) and those that did so to avoid Section 404(a) (or were seeking to avoid both provisions). However, we generally refer to the effects we find as attributable to Section 404(b) because the NAF exemption applied only for 404(b) for the majority of the SOX sample period. Moreover, most of the academic and practitioner literature has focused on 404(b) rather than 404(a).

4) A Simple Framework for Analyzing Firms' Bunching Behavior

4.1) An Overview of the Bunching Approach

This paper applies to the setting described in Section 3 an empirical technique that draws on a growing literature in economics analyzing "bunching" around tax and regulatory thresholds. Bunching analysis was originally applied to the study of responses to taxation (e.g. Saez, 2010; Chetty et al., 2011). However, it has increasingly been applied to the analysis of the consequences of regulation, exploiting bright-line thresholds – often based on firm size – that are commonly used to determine the applicability of various types of regulations. For instance, Almunia and Lopez-Rodriguez (2015) analyze the responses of Spanish firms to a threshold at €6 million of revenue above which they are subjected to greater scrutiny by the tax authorities. Marx (2015) studies the responses of US charities to an income threshold above which they face more onerous reporting requirements in order to establish their tax-exempt status. In the terminology of Slemrod (2013), these types of thresholds represent "notches" at which there is a discrete jump in regulatory obligations.

The basic idea underlying this approach can be illustrated with the aid of Figure 1. This graph represents in stylized form the distribution of firms on either side of a size-related threshold

[19] If public float cannot be calculated, a firm can also qualify as a smaller reporting company if its revenue is below $50 million

(such as the $75 million public float threshold in our setting). A bunching analysis uses data on firms' reported size to group firms into size "bins" of equal width. The analysis focuses on the number of firms in each bin (i.e. the density of firms). The aim is to estimate the counterfactual density (i.e. the number of firms that would be expected to be found around the threshold), by fitting a high-order polynomial to the data on the number of firms in each bin, excluding an interval close to the regulatory threshold. In Figure 1, the observed density of firms in regions away from the threshold is shown as a solid line. The dashed line represents the counterfactual density over the excluded interval. Importantly, the bunching approach imposes the assumption that the density of firms over the excluded interval would be smooth absent the regulatory threshold.

A bunching analysis then compares the observed density of firms over the excluded interval with the counterfactual density. In a typical scenario in which firms bunch below the threshold, the actual density will exceed the counterfactual density below the threshold, creating a bunching region with an "excess" number of firms. In Figure 1, the area (or "mass") of this region is denoted by *B*. This approach can determine whether there is an unexpectedly large number of firms in bins immediately below the threshold. However, it cannot pinpoint *which* of the firms that are observed to be in the bunching region are "bunchers" and which firms would have located there even in the counterfactual scenario without the threshold. While bunching typically occurs below the threshold, in principle beneficial regulations may induce bunching above the threshold. If there is bunching below the threshold, this will typically be accompanied by a "missing" region on the other side of the threshold. The area (or mass) of the missing region is denoted by *H* in Figure 1.

The determination of the excluded interval is of course of crucial importance in this methodology. There is no standard approach to making this determination. Often, the lower bound of the excluded region will be readily apparent upon visual inspection of a histogram representing the actual density of firms. However, inferring how far beyond the threshold the bunching firms would have located (absent the threshold) is critical to understanding how much distortion is created by the regulation. The excluded interval – $66-$83 million of public float – is chosen using an iterative process based on the divergence from the counterfactual density at various candidate locations close to the threshold.[20]

[20] Kleven and Waseem (2013) develop an iterative procedure that determines the upper limit of the excluded region by a process of repeatedly increasing this limit until the area of *H* converges to that of *B* (see also Almunia and Lopez-Rodriguez, 2015). This imposes the assumption that firms that find themselves just above the threshold do not exit the dataset in substantial numbers. We do not impose the assumption that $B = H$ in our baseline analysis because, for

The estimated magnitude of bunching allows inferences to be made about the extent to which firms are willing to change their reported size in order to avoid the regulation. It may also be possible to infer the welfare consequences of the regulation in particular contexts. However, it is important to emphasize that what these studies aim to do is not to highlight distortions arising from the particular choice of a bright-line threshold, but rather to use the existence of this threshold as a source of quasi-experimental variation to draw wider inferences about the costs and benefits of the regulation. Indeed, this literature does not take a stand on whether bright-line thresholds are optimal as a normative matter. In theory, it is possible that in the presence of administrative costs, an optimally-designed tax or regulatory system may include bright-line thresholds, notwithstanding the distortionary behavior that they induce immediately around the threshold (for a formal model, see e.g. Dharmapala, Slemrod and Wilson, 2011).

4.2) A Simple Model of Firms' Bunching Behavior near the $75 Million Threshold

Consider a firm with exogenous fundamental value *V* that is controlled by an insider who owns a fraction (1 - α) of its stock. The insider has exogenous personal wealth *W* (in addition to her ownership of the firm's stock). The insider's payoff, denoted U_I, can thus be expressed as:

$$U_I = W + (1 - \alpha)V \tag{1}$$

In our very simple setting, public float (denoted *PF*) can be expressed as: $PF = \alpha V$; α (the fraction of the firm owned by outsiders) is the ratio of public float to market value (i.e. *PF/V*). The ownership structure of the firm is encapsulated by α. We assume that there is a continuum of firms that are identical, apart from their exogenously given fundamental value *V*; thus, the only source of heterogeneity among firms is variation in *V*.

It is assumed that the insider has a certain amount of discretion in the choice of α (for instance, by buying or selling more or less of the firm's stock). This choice potentially affects the insider's payoff and firm value by modifying the fundamental value *V*. For instance, it is possible that the choice of α affects the firm's governance, the incentives of the insider to monitor the firm's operations, or the firm's access to external capital. Absent the regulatory threshold introduced below, there is assumed to be an optimal choice of α, denoted α^*, that is (at least locally) independent of *V*. While the existence and uniformity of α^* is simply assumed here, it is possible

instance, it is possible that firms that would counterfactually be just above the threshold may go private and therefore not appear in the data. However, one of our specifications - using pre-SOX data to compute the counterfactual density - generates approximately equal values of *B* and *H* (as discussed in Section 5.4 below).

to derive these characteristics from a more general formulation of insider payoffs than is used in Equation (1).[21]

Deviating from this optimal choice α^* entails two types of costs. The first type is borne exclusively by the insider; this represents, for instance, increased risk-bearing by the insider due to reduced portfolio diversification when inside ownership is high (i.e. when α is low). This is captured by a loss function $K_I(\alpha^* - \alpha)$. Second, there are costs to the firm that are borne *pro rata* by all shareholders. For instance, the value of the firm will fall below its (potential) fundamental value V if α is too low and the firm's access to external capital is limited. This is captured by a separate loss function $K_F(\alpha^* - \alpha)$, which is borne by the insider in proportion to her ownership stake. It is assumed that both $K_I(\alpha^* - \alpha)$ and $K_F(\alpha^* - \alpha)$ are increasing and convex, and that $K_I(0) = K_F(0) = 0$.

Now, suppose that a regulatory threshold $\overline{PF}$ is introduced, defined in terms of public float at $PF = \overline{PF}$. Firms with $PF \geq \overline{PF}$ are subject to an additional set of regulatory obligations, such as those entailed by SOX 404(b) (for which $\overline{PF} = 75$, where public float is measured in millions of dollars). Suppose that these regulatory obligations entail a net compliance cost of C; this is the cost of complying with the regulation, net of any benefits that are conferred on the firm by being subject to the regulation. For instance, SOX 404(b) entails higher audit fees, the potential diversion of managerial effort into compliance activity, liability risk, and other possible costs. On the other hand, internal control attestation may make firms' financial disclosures more credible to investors, and firms not subject to the regulation may not be able to replicate the same level of credibility through voluntary mechanisms (for instance, because sanctions are less severe). Our definition of C is the sum of these various costs, minus the benefits of being subject to the regulatory regime. For simplicity, it is assumed that C is a fixed cost (independent of V).

The simple model we present here assumes that insiders can manage PF only by changing the ownership structure (i.e. α). In reality, there may be other mechanisms to achieve this aim, such as disclosures that reduce stock price or reclassification of the insider status of blockholders (Gao, Wu and Zimmerman, 2009; Nondorf, Singer and You, 2012; Gao, 2016). The focus here on the potential manipulation of α is purely illustrative; the empirical analysis does not hinge on this

[21] Suppose that we define U_I more generally as: $U_I = u(W - (1-\alpha)V) + (1-\alpha)V$, where u is an increasing, concave function. Then, the first-order condition with respect to α is $Vu'(\alpha^*) = V$, and the optimal choice α^* is defined by $u'(\alpha^*) = 1$. This implies that α^* is independent of V.

being the only means of managing public float. The basic results would hold regardless of how public float is managed, as long as alternative methods of managing public float have costs analogous to $K_I(\alpha^* - \alpha)$ and $K_F(\alpha^* - \alpha)$. The model also does not take account of any potential change in insiders' private benefits of control as a result of being subject to SOX 404(b); however, an extension to the model is Section 5.6.3 addresses this issue.

Let $\bar{V} = \overline{PF}/\alpha^*$ be the fundamental value of a firm that just has a sufficiently large (undistorted) public float to satisfy the regulatory threshold. Clearly, firms with $V < \bar{V}$ are unaffected by the introduction of the threshold, and the insiders of these firms will choose α^* and the same *PF* as in the absence of the threshold. Consider a firm with *V* at or slightly above $\bar{V}$: i.e. $V = \bar{V} + \epsilon$, where ϵ is small. The payoff to the insider of this firm from choosing the (undistorted) ownership structure α^* is:

$$U_I = W + (1 - \alpha^*)(\bar{V} + \epsilon - C) \tag{2}$$

Alternatively, the insider of this firm could change the ownership structure by slightly reducing outside investors' stake to $\hat{\alpha} < \alpha^*$ (where $\hat{\alpha}$ is the value of α that is just low enough to reduce *PF* below $\overline{PF}$). For values of $\hat{\alpha}$ very close to α^*, the impact of such a change on the loss functions $K_I(\alpha^* - \alpha)$ and $K_F(\alpha^* - \alpha)$ will be negligible, as they are approximately zero in the neighborhood of zero. Thus, for a sufficiently small ϵ, the payoff to this insider from changing the ownership structure is approximated by:

$$U_I \approx W + (1 - \hat{\alpha})(\bar{V} + \epsilon) - (\alpha^* - \hat{\alpha})(\bar{V} + \epsilon) = W + (1 - \alpha^*)(\bar{V} + \epsilon) \tag{3}$$

If $C > 0$ (i.e. the net compliance cost is positive) and ϵ is sufficiently small, then the insider will be better off managing public float downwards. This behavior will generate bunching, as some firms with $V \geq \bar{V}$ report public float $PF < \overline{PF}$.

For firms with *V* further above $\bar{V}$, the insider's payoff from choosing to bunch is:

$$U_I = W - (\alpha^* - \hat{\alpha})\big(V - K_F(\alpha^* - \hat{\alpha})\big) - K_I(\alpha^* - \hat{\alpha}) + (1 - \hat{\alpha})\big(V - K_F(\alpha^* - \hat{\alpha})\big) \tag{4}$$

Here, the second term represents the direct personal cost to the insider of acquiring additional shares of the firm, while the third term represents indirect costs to the insider of changes to the ownership structure (such as from a loss of personal portfolio diversification). Equation (4) can be simplified as follows:

$$U_I = W + (1 - \hat{\alpha})\big(V - K_F(\alpha^* - \hat{\alpha})\big) - K_I(\alpha^* - \hat{\alpha}) \tag{5}$$

In choosing whether to bunch, insiders will compare the payoff in Equation (5) with the payoff (implied by Equation (2) above) of $U_I = W + (1 - \alpha^*)(V - C)$ from not bunching.

For some fundamental value $V_m > \bar{V}$, the insider of a firm with fundamental value V_m will be indifferent between bunching (by choosing an ownership structure $\alpha_m = \overline{PF}/V_m$) and choosing the undistorted ownership structure α^*. This value V_m can be characterized by equating the payoffs from bunching and not bunching:

$$W + (1 - \alpha^*)(V_m - C) = W + (1 - \alpha_m)\big(V_m - K_F(\alpha^* - \alpha_m)\big) - K_I(\alpha^* - \alpha_m) \quad (6)$$

Let $PF_m = \alpha^*(V_m - C)$ be the undistorted public float that corresponds to V_m. Then, rearranging Equation (6) yields the following characterization of V_m:

$$V_m = \left(\frac{1}{\alpha^*}\right)\big(C + PF_m - K_F(\alpha^* - \alpha_m) - K_I(\alpha^* - \alpha_m)\big) + K_F(\alpha^* - \alpha_m) \quad (7)$$

Equation (7) implies that the larger the net compliance cost C of the regulation, the larger is V_m, and hence the larger the number of firms whose insiders find it optimal to bunch.

In this simple framework, all firms with fundamental value from $\bar{V}$ to V_m will bunch – i.e. will reduce their outside ownership to the (minimal) extent required to remain below the threshold. These firms will report public float below $\overline{PF}$. There will be an excess mass of firms just below the threshold $\overline{PF}$, as those firms that would counterfactually locate in this region are joined by the bunching firms. Conversely, there will be a missing region above the threshold, with no firms reporting public float in the range $\overline{PF}$ to PF_m. From PF_m onwards, firms will report undistorted values of their public float. Thus, firms with fundamental value $V > V_m$ will not change their ownership structure due to the regulatory threshold, although they will of course bear the net compliance cost C.

This pattern of outcomes is illustrated in Figure 2. Here, $\Delta PF = PF_m - \overline{PF}$ represents the amount (in millions of dollars) by which the marginal buncher (the firm whose insider is indifferent between bunching and not bunching) reduces its public float in response to the SOX 404(b) threshold. Let $f_0(PF)$ be the counterfactual density of firms defined over the possible values of public float. Assuming that $B = H$, the magnitude of the excess mass (i.e. the area of B in Figure 2) can in general be defined as:

$$B = \int_{\overline{PF}}^{PF_m} f_0(PF)dPF \quad (8)$$

Assuming that $B \approx H$ and that $f_0(PF)$ is approximately flat in the region around the \$75 million threshold (as shown in Figure 2), it is possible to use the following approximation for the magnitude of the excess mass:

$$B \approx H \approx \Delta PF * f_0(75) \tag{9}$$

where $f_0(75)$ is the counterfactual density evaluated at the threshold. Rearranging Equation (9) results in a simple approximation for ΔPF:

$$\Delta PF \approx \frac{B}{f_0(75)} \equiv \hat{b} \tag{10}$$

Here, $\hat{b}$ is the "bunching ratio" that is commonly estimated in bunching studies (e.g. Chetty et al., 2011; Almunia and Lopez-Rodriguez, 2015; Marx, 2015). It is the ratio of the excess mass (*B*) to the height of the counterfactual density at the threshold (i.e. $f_0(75)$). This ratio is a sufficient statistic for estimating the extent to which firms reduce their public float in response to the \$75 million threshold (or, more generally, for the extent to which firms respond to a size-based threshold by reducing their reported size).

There are a number of important caveats to be noted. As we show below (in common with most applications of the bunching approach), there is not a completely missing region above the threshold (such as the region *H* shown in Figure 2). Rather, there is only a partially missing region, with many firms being insensitive to the threshold for various reasons. We do not impose the restriction that *B* = *H* in the basic analysis below because it is possible that firms just above the threshold may go private and therefore not appear in the data (however, one of our specifications - using pre-SOX data to compute the counterfactual density - generates approximately equal values of *B* and *H*; see Section 5.4 below). The estimated counterfactual density is declining rather than flat. Thus, it should be borne in mind that the analysis described above relies on a number of simplifications; nonetheless, the approximation given above by $\hat{b}$ remains valuable in interpreting the magnitude of the observed bunching behavior (as shown in Section 5 below).

The amount ΔPF by which firms reduce public float is an important behavioral parameter, for which the bunching ratio $\hat{b}$ is a sufficient statistic under certain assumptions. Much of the SOX literature has focused on the magnitude of the net compliance costs, and thus it would also be of great interest to estimate *C*. Unfortunately, however, $\hat{b}$ is not a sufficient statistic for *C*. To see this, we can derive an expression for *C* from Equation (6):

$$C = K_F(\alpha^* - \alpha_m) + K_I(\alpha^* - \alpha_m) + (\alpha^* - \alpha_m)\big(V_m - K_F(\alpha^* - \alpha_m)\big) - \Delta PF \tag{11}$$

As is apparent from Equation (11), C depends not only on ΔPF (and hence on $\hat{b}$), but also on various other unobservable terms, such as the loss functions K_F and K_I. The net compliance cost C thus cannot be directly inferred from the observed magnitude of bunching. In Section 5.7 below, we use a different approach – based on combining the estimated $\hat{b}$ with extrapolations of the reduced-form relationships among relevant variables – to provide a rough estimate of C.

5) Empirical Analysis

5.1) Data

As is evident from the account in Section 3, public float is a crucial concept in determining firms' obligations under SOX. Nonetheless, research using public float has been limited because it is not reported in standard archival financial databases such as Compustat and CRSP. The previous literature has hand-collected the value of public float from firms' annual 10-K filings with the SEC for various subsamples of firms (e.g. Gao, Wu and Zimmerman, 2009: Iliev, 2010; Dharmapala and Khanna, 2016; Weber and Yang, 2016). This paper constructs a much larger dataset on public float by using Python code to "scrape" information on public float from firms' 10-K filings.[22] The filings are accessed through the SEC's Electronic Data Gathering and Retrieval (EDGAR) system.[23]

The data-gathering process is facilitated by the relatively uniform nature of public float reporting. On the first page of the 10-K form, a reporting entity states its public float (determined as of the last day of the second fiscal quarter), as well as checking one of a number of checkboxes specifying its filing status. A typical example of the language used when reporting public float is the following:[24]

> "The aggregate market value of common stock held by non-affiliates of the registrant based on the closing price of the registrant's common stock as reported on the NASDAQ Global Market on June 28, 2013, was $234,272,491. Shares of voting and non-voting stock held by executive officers, directors and holders of more than 5% of the outstanding stock have been excluded from this calculation because such persons or institutions may be deemed affiliates. This determination of affiliate status is not a conclusive determination for other purposes."

[22] I am very grateful to Daniel Marcin (formerly of the Coase-Sandor Institute for Law and Economics at the University of Chicago Law School) for writing this code.
[23] See http://www.sec.gov/edgar/searchedgar/companysearch.html
[24] This example is from the 2013 10-K form for Brightcove, Inc.

The Python code automatically extracts the number that follows the phrase: "The aggregate market value . . ." – in this instance, $234,272,491.

Using this method, we collect public float information for the universe of US reporting entities for fiscal years 1993-2015. This period spans the SOX legislation and the various changes in its implementation that were outlined in Section 3. It extends back as far as is feasible in terms of the online availability of these filings. In addition, four checkboxes are listed on the first page of the 10-K form to indicate the issuer's filing status. A firm must report being an accelerated filer, a large accelerated filer, a nonaccelerated filer, or a smaller reporting company. The automated process also collects this data on firms' filing status from the checkboxes. However, the coverage of the checkbox data is substantially more limited than that of the numerical public float data.[25]

The full dataset on public float contains 160,988 observations at the firm-year level on 23,719 distinct reporting entities over 1993-2015. Firms are identified by the Central Index Key (CIK) number that is assigned by the SEC to filing entities. The number of reporting entities may seem large in relation to the number of publicly-traded companies on the US stock market. The full dataset may include, for instance, very small entities that become subject to SEC reporting requirements by issuing debt securities.[26] While such issuers are potentially subject to many elements of SOX (e.g. Bartlett, 2009), a more narrowly focused dataset is used in the empirical analysis below. In particular, it omits all firm-years with less than $50 million of public float. The reporting entities that are used in the analysis are thus more likely to resemble the conventional publicly-traded companies that have been the focus of discussion in relation to SOX.

Descriptive statistics for public float are provided in Table 1. As our interest is in the behavior of firms that are relatively close to the $75 million threshold, we focus only on issuers with public float in the range of $50-$150 million. There are 21,848 firm-year observations on 7,338 issuers in this range.[27] Of these observations, 10,247 observations are from the pre-SOX period and 11,601 from the SOX period. The mean and dispersion of public float across the pre-SOX and SOX periods is remarkably similar, as shown in Table 1.

[25] For instance, there are 88,793 observations on filing status from the checkbox data, compared to nearly 161,000 observations on public float.

[26] The dataset does not include registered investment companies regulated by the Investment Company Act of 1940, for which SOX rules are not relevant, or other entities that do not report public float to the SEC.

[27] In most cases, the scraping procedure obtains one value of public float for each firm-year. However, firms in some situations file amended 10-K forms, and the automated data collection process collects data from both the original and amended forms. The dataset has been cleaned to ensure that there is only one observation per firm-year, using the latest filing by a firm for a given filing year.

This dataset is very large in comparison to those used in prior studies of SOX. Indeed, it constitutes the universe of public float reports available online, and includes in particular the entire universe of post-SOX public float reports. Nonetheless, it is important to bear in mind that it is relatively small in relation to the size that is ideal for the implementation of bunching analysis. This sample size limitation affects aspects of the analysis below (for instance, necessitating relatively large bin widths).

The data on public float collected here can be linked (using the CIK number) to standard databases such as Compustat. It is thus possible to compute α (the ratio of public float to market capitalization) by combining the public float variable summarized in Table 1 with the Compustat variable MKVALT. Table 2 reports descriptive statistics for α for the approximately 67,000 firm-year observations for which the data can be successfully linked.[28] If firms reduce their public float in order to remain below the $75 million threshold, it would be expected that α would on average be smaller for these firms. The descriptive statistics in Table 2 are consistent with this expectation. Among firm-years for which public float falls within the region $66-$83 million (the excluded region that is derived in Section 5.3 below), those that report public float of $66-$75 million during the SOX period (some of which may potentially be bunchers) have a mean α of 0.61. In contrast, pre-SOX firm-years within this region and those that are just above the threshold ($75-$83 million) have a mean α of 0.63. However, when this difference in α is tested within a panel regression framework (with firm fixed effects, year effects, and an extensive set of controls), it is not robustly statistically significant.

We also analyze the importance of the threshold by regressing the future growth rate of public float on a set of relevant variables at the firm-year level (with firm fixed effects and year effects). Firm-years for which public float is close to the threshold (specifically, in the $66-$75 million range) experience slower growth in public float in the subsequent year (relative to that for firm-years in the $50-$66 million range). However, the difference in this difference across the pre-SOX and SOX periods is not statistically significant. A possible reason is that (as shown in Section 5.3 below) excess firm-years constitute only about a sixth of the observed firm-years within the bunching region. Thus, any distinctive characteristics of bunching firms would be highly attenuated in a panel regression framework.

[28] For some firm-years, public float exceeds the market capitalization reported in Compustat. These apparently anomalous observations are omitted from the sample for which descriptive statistics are reported in Table 2.

Using the longitudinal structure of the public float data, it is possible to calculate the probability that a firm that reports public float close to $75 million (specifically, in the $66-$75 million range) in year *t* subsequently crosses the threshold. There is a substantial probability (of about 0.37) that a firm in this situation crosses $75 million in the following year. However, as shown in Figure 3, this "hazard" is smaller in the SOX period than in the pre-SOX period. By year (t + 5), there is a 0.61 probability of crossing $75 million in the pre-SOX period, and a 0.52 probability of doing so in the SOX period. While there may be alternative explanations, this difference is consistent with firms seeking to avoid crossing the regulatory threshold after the enactment of SOX.

5.2) Graphical Evidence of Bunching

The analysis aggregates the firm-year-level data described in Section 5.1 to a bin-level dataset that measures the number of firms in each bin of public float (i.e. firm density). As entities with zero or very low values of public float and larger firms with public float that substantially exceeds the threshold are unlikely to be of much relevance, we restrict attention to values of public float in the range of $50 million to $150 million. We start at $50 million because the density appears substantially different below this level. The range is extended to $150 million (rather than the more natural $50-$100 million) in order to provide a sufficient number of bin-level observations to meaningfully estimate the counterfactual density using a fifth-order polynomial, and to be able to compute bootstrapped standard errors for the bunching parameter using 200 replications; in any event, the density over $100-$150 million does not appear substantially different from that for the preceding bins.

We divide the $50-$150 million range into bins with a uniform width of $1 million for the baseline analysis. For instance, "bin 75" includes observations of public float that range from $74 million up to (but not including) $75 million. This definition ensures that no bin crosses the regulatory threshold.[29] This bin width is narrower than would be ideal, creating a substantial degree of noisiness in the observed density. However, it allows for a sufficient number of bin-level observations to implement the analysis, while the noise may be expected to create a bias against any findings. Observations of public float are pooled across years within the pre-SOX period to calculate a variable N_i^{PreSOX} (the number of firm-year-level observations of public float that are

[29] Recall that the definition of accelerated filers requires an issuer to have ". . . an aggregate worldwide market value of the voting and non-voting common equity held by its non-affiliates of $75 million or more" (17 CFR 240.12b-2).

found within bin *i* over the pre-SOX period (1993-2002)). Similarly, observations of public float are pooled across years within the SOX period to calculate a variable N_i^{SOX} (the number of firm-year-level observations of public float that are found within bin *i* over the SOX period (2003-2015)).

In comparison to many other statistical approaches, bunching analysis lends itself to "the ocular proof" in that bunching can often be readily apparent through visual inspection. Figure 4 is a histogram representing the number of firm-year observations of public float during the pre-SOX period using bins of width $5 million. In general, the density declines as public float rises – i.e. there are fewer firm-years with larger values of public float, even in our restricted range. However, $75 million of public float (indicated by the vertical red line) does not appear to be associated with any particular departure from the general pattern. Figure 5 shows the pre-SOX density using bins of width $1 million. With this narrower bin width, the patterns are considerably noisier, with substantial variation in density across bins. However, there still appears to be no bunching around the $75 million level. In other words, Figures 4 and 5 do not suggest that prior to SOX $75 million was especially salient or important as a focus for public float reporting.

In contrast, Figure 6 - a histogram representing the number of firm-year observations of public float during the SOX period, using the $1 million bins defined above - reveals strong visual evidence of bunching below the $75 million threshold. In particular, there appears to be an unusually large number of firm-years in the range of $73-$75 million of public float. What may appear anomalous is that the bunching region (corresponding to *B* in Figure 1) is bimodal – in addition to the excess mass of firms over $73-$75 million, there is another, substantially larger, excess mass in the range $69-$71 million. As this is some distance from the threshold, its connection to SOX may appear questionable. Note, however, that there is no excess mass in that region in the pre-SOX period (see Figure 5). There also appears to be no other regulatory threshold around $71 million of public float that could account for the pattern in the SOX era.[30] It is possible that the $69-$71 million region is economically salient in the SOX period (though not the pre-SOX period), but this appears highly unlikely, and has not been remarked upon in the accounting or finance literatures.

[30] Public float is a concept used primarily in securities law, and we are unaware of any securities law provision that applies at the $71 million threshold.

The distinctive bimodal pattern of bunching seen in Figure 6 may be more explicable when it is borne in mind that while firms have a certain amount of discretion in determining public float, reported public float also depends on stock price. Having chosen a level of insider ownership, the firm remains subject to unexpected increases in share price on the last day of the second quarter that drive its public float above the $75 million threshold. This possibility may induce many bunching firms to locate some distance below the threshold, rather than immediately below it. For example, a firm that chooses public float of $70 million allows itself a buffer of about a 7% unexpected stock price increase before it crosses the threshold. While we observe realized rather than expected public floats, if the probability of an unexpected stock price increase is small and firms are risk-averse about the possibility of crossing the threshold, realized public floats among bunching firms may be mostly well below the threshold.[31]

Formally, the impact of stock price volatility can be viewed as adding a stochastic term to reported public float. An implication is that firms locating around $70 million should have higher stock price volatility than those locating around $74 million. Using the Compustat variable OPTVOL, it appears that volatility for firms that repeatedly locate in the bunching region and are found in the $66-71 million range is higher than for firms found in the $71-75 million range. However, the limited number of observations makes it difficult to reach definite conclusions.

5.3) Basic Empirical Analysis

While Figure 6 provides strong visual evidence of bunching, our methodology is based on the formal estimation of a counterfactual density and a bunching parameter $\hat{b}$, and on obtaining standard errors for $\hat{b}$ through bootstrapping. The regression specification used to estimate the counterfactual density can be represented as follows:

$$N_i^{SOX} = \alpha + \sum_{j=1}^{5} [\beta_j (Mid_i)^j] + \gamma_i I_i + \epsilon_i \tag{12}$$

Here, *i* indexes bins (where, for example bin 55 is the bin that includes public float observations that fall in the range from $54 million to just under $55 million). N_i^{SOX} is the number of firm-year-level observations of public float that are found within bin *i* over the SOX period (for instance,

[31] Risk-aversion may be a particularly reasonable assumption because of the asymmetric nature of the threshold – a firm that crosses it remains subject to SOX 404(b) even if its public float subsequently falls below $75 million (unless its public float subsequently falls below $50 million; a stricter rule, requiring public float to fall below $25 million and certain other conditions to be met, applied before 2005).

N_{55}^{SOX} represents the number of observations of public float that fall in the range from \$54 million to just under \$55 million over the SOX period). Mid_i is the midpoint of the range of public floats included within bin i – for instance, for bin 55 (i.e. i = 55), Mid_{55} = 54.5 (where public float is measured in millions of US \$). The summation over the values of j represents a flexible fifth-order polynomial that is used to estimate the counterfactual density.

The estimation of the counterfactual density excludes an interval around the \$75 million threshold. This excluded interval is denoted by [L, U], where L and U are the bins that represent the lower and upper bounds of the excluded interval. Its exclusion is accomplished in Equation (12) by adding a series of indicator variables I_i = 1 if $i \in [L, U]$ and 0 otherwise, that remove the influence of the bins in the excluded interval from the estimation of the counterfactual density. In choosing the excluded interval [L, U], we begin with the visual inspection of Figure 6 to determine reasonable lower and upper bounds, and then follow an iterative process of including indicators for the bins in and around that interval until we exhaust the bins that exhibit a significantly different observed density relative to the counterfactual density. Based on this iterative process, we use an excluded interval ranging from \$66-\$83 million.

Let the predicted values from Equation (12) – i.e. the estimated counterfactual density – be denoted by $\widehat{N}_i^{SOX}$; then, the excess mass B can be estimated as follows:

$$\hat{B} = \sum_{i=L}^{75} (N_i^{SOX} - \widehat{N}_i^{SOX}) \tag{13}$$

The magnitude of the missing mass H can be estimated as follows:

$$\widehat{H} = \sum_{i=76}^{U} (\widehat{N}_i^{SOX} - N_i^{SOX}) \tag{14}$$

The height of the counterfactual density $f_0(75)$ can be estimated by averaging the values of $\widehat{N}_i^{SOX}$ over the 9 bins (representing public float of \$66-\$75 million) within the excluded region and under the threshold:[32]

$$\hat{f}_0(75) = \frac{\sum_{i=L}^{75} \widehat{N}_i^{SOX}}{9} \tag{15}$$

Finally, the bunching ratio defined in Equation (10) can be estimated as:

[32] This approach smooths the estimated height of the counterfactual density over the bunching region, and thus takes account of the fact that the counterfactual density in this region is not flat but rather is declining (and so is not necessarily well-approximated by simply evaluating f_0 at the threshold).

$$\hat{b} = \frac{\hat{B}}{\hat{f}_0(75)} \tag{16}$$

Table 3 reports the estimates of $\hat{B}$, $\hat{H}$ and $\hat{b}$. As we use the universe of public float data rather than a sample, there is arguably no sampling error and conventional standard errors are thus inappropriate. We compute bootstrapped standard errors using 200 replications (the standard number of replications in this literature – see e.g. Almunia and Lopez-Rodriguez (2015)). The first row of Table 3 reports these parameter estimates for the pre-SOX period (1993-2002), which involves using N_i^{PreSOX} as the dependent variable in Equation (12), with the subsequent equations defined analogously. As would be expected from Figures 4 and 5, the estimated $\hat{b}$ is relatively small and statistically insignificant; indeed, it has a negative sign, suggesting that there are slightly fewer firm-year observations immediately below $75 million than would be expected. Even so, it is possible that bunching under the $75 million threshold may be manifested in the pre-SOX period at a value of public float that corresponds not to the nominal $75 million level, but to its inflation-adjusted equivalent. Thus, we use the inflation data provided by the Bureau of Labor Statistics[33] to adjust the threshold for the rate of inflation from the midpoint of the pre-SOX period to the midpoint of the SOX period. When implementing a similar analysis using the inflation-adjusted threshold of $57 million, the estimated $\hat{b}$ remains small and statistically insignificant. This casts doubt on the possibility that there is some economic reason underlying bunching under the $75 million threshold during the SOX period.[34]

The second row of Table 3 – which presents our baseline results – reports $\hat{B}$, $\hat{H}$ and $\hat{b}$ for the SOX period (2003-2015). There are approximately 257 excess firm-years below the threshold; the counterfactual number of firm-year observations in the interval from $66-$75 million is 1334, and the observed number is 1591. Hence, about one sixth of the observations in this range represent "bunchers" (though, as discussed previously, it is not possible to identify *which* of these observations are bunchers and which are not). The estimated $\hat{b}$ is approximately 1.7, and is statistically significant at the 5% confidence level (using bootstrapped standard errors as described above). Thus, the excess mass below the threshold is sufficiently unusual (relative to variations in the observed density elsewhere in the distribution) to be well outside the expected range of random

[33] See http://www.bls.gov/data/inflation_calculator.htm.

[34] Ideally, we would also take account of inflation (and the changing economic meaning of the nominally fixed $75 million threshold) within the SOX period. However, the limited number of observations for each year within the SOX period precludes this.

variation. This result strongly vindicates the visual evidence from Figure 6 of substantial bunching below the $75 million threshold, and implies that firms around the threshold reduce their public float by about $1.7 million on average (that is, averaged over all firm-years in the region of the threshold, including those firms that are for various reasons indifferent to the threshold).

While there are approximately 257 excess firm-years below the threshold, the missing mass above the threshold consists only of about 69 firm-years, thus implying that $B > H$. This is difficult to explain, as it would require substantial entry onto the public markets in the SOX period by firms just above the threshold. We thus turn next to the issue of how to reconcile the estimated values of B and H.

5.4) A "Pre-Reform Counterfactual" Approach

Equation (12) implies that the counterfactual density is estimated using data for the SOX period. This conforms to standard bunching analysis, in which a maintained assumption is that the threshold does not affect the behavior of firms that are located far from it. Given that we observe pre-SOX data, however, it is also possible to use the density of firms in the pre-SOX period to control for the density of firms in the SOX period (in contrast to many prior bunching studies, in which the regulatory threshold does not change over time).[35] This alternative approach recognizes that it is possible to imagine two alternative counterfactuals – one in which the threshold does not exist because *no* firms (at least in the $0-150 million range) are subject to 404(b), and one where the threshold does not exist because *all* firms are subject to 404(b). Arguably, the former is more relevant for policy purposes, given that the exemption for smaller firms has been made permanent and that policy debate has focused on increasing the threshold for exemption. The pre-SOX counterfactual density provides us with a glimpse into the situation where no firms are subject to 404(b). In such a scenario, there may be more observations of public float above the excluded interval than in the actual SOX period density. If, as in Equation (12), we use the SOX period to derive the counterfactual density, then the counterfactual density may arguably be biased downward (relative to the "true" counterfactual); in turn, this may lead to H being underestimated.

This point is illustrated in Figure 7, which shows the same observed density below the threshold as in Figure 1. Above the threshold (where the regulation applies), however, the observed density in Figure 7 lies somewhat below the Figure 1 density. In Figure 1, the estimated

[35] For instance, the €6 million revenue threshold used by Almunia and Lopez-Rodriguez (2015) is fixed throughout their sample period.

counterfactual density in the excluded interval is given by the upper dotted line. If we instead observe the actual density when firms are subject to SOX 404(b), then the estimated counterfactual density - as given by the thick dashed line in Figure 7 - will be lower. This implies that the estimated magnitude of *H* will be smaller (and the estimated magnitude of *B* will be larger) than if we were able to observe the counterfactual density in the absence of SOX 404(b). Potentially, this may account for the baseline estimate we obtained (in Table 3, Row 2) in which $B > H$.

Figure 8 provides some evidence that this scenario may apply to our setting. It compares the actual SOX period density to the counterfactual density computed using pre-SOX data. The scale of the actual and counterfactual densities is somewhat different, as there are slightly more SOX-period observations in our data. What is most notable, however, is that there is a substantial difference in how well the pre-SOX counterfactual density matches the actual SOX period density above versus below the excluded interval. In particular, the entire density in the SOX period appears to be dampened relative to the pre-SOX counterfactual, not only immediately above the threshold but also everywhere above the excluded interval. This would be consistent with a general tendency for larger firms (those that would be subject to SOX 404(b)) to go private or to stay out of the public capital markets in the 2003-2015 period.[36] It should be emphasized that this phenomenon cannot necessarily be claimed to be a causal effect of SOX: it is possible that other factors operative in 2003-2015 may be responsible. Even so, Figure 8 suggests that the pre-SOX counterfactual density constitutes an important alternative control.

The "prereform counterfactual" approach can be implemented as follows. Define:

$$N_i^{PreC} = \begin{cases} N_i^{SOX} & \text{if } i \in [L, U] \\ N_i^{PreSOX} + k & \text{otherwise} \end{cases} \tag{17}$$

Here, k is an arbitrary constant that scales the pre-SOX density to correct for the fact that our dataset has a larger number of observations on average in the SOX period than in the pre-SOX period. We then estimate the counterfactual density as follows:

[36] In the literature, much attention has been paid to the exit of firms subject to SOX from the market (e.g. Bartlett, 2009; Kamar, Karaca-Mandic and Talley, 2009) or from reporting entity status (e.g. Leuz, Triantis and Wang, 2008). It is not clear whether it is reasonable to expect the measured missing mass to be substantially affected by the exit of firms in our setting. This depends on whether going-private transactions would be clustered just above the threshold. Insiders need less financing to take a $75 million firm private than to take a $150 million firm private; however, it is also less costly to reduce the former's public float below the threshold. Thus, it is not clear that exit would disproportionately affect firms just above the threshold.

$$N_i^{PreC} = \alpha + \sum_{j=1}^{5} [\beta_j (Mid_i)^j] + \gamma_i I_i + \epsilon_i \tag{18}$$

where the variables are as defined earlier. We choose the constant *k* in Equation (17) based on an iterative procedure that converges when we approximate the condition that *B* = *H* (this turns out to be when *K* = 18.5).

The bunching parameters that we estimate using this approach are reported in Row 3 of Table 2. In most respects, the results are fairly close to those in Row 2. The estimated excess mass is 182 observations; the estimated $\hat{b}$ is approximately 1.2, and is statistically significant at the 10% confidence level (using bootstrapped standard errors as described above). However, the estimated area *H* of the missing region is now substantially larger, at about 180 observations. Thus, the specification in Equation (18) enables us to more satisfactorily reconcile the observed amount of bunching below the threshold with the observed missing region above the threshold.

Alternatively, it is possible to use a multiplicative constant *k* to adjust for the fact that our dataset has a larger number of observations on average in the SOX period than in the pre-SOX period. This involves defining:

$$N_i^{PreC} = \begin{cases} N_i^{SOX} & \text{if } i \in [L, U] \\ kN_i^{PreSOX} & \text{otherwise} \end{cases} \tag{19}$$

The most natural choice for *k* is arguably the ratio of the aggregate number of observations in the SOX period to the aggregate number of observations in the pre-SOX period: i.e. *k* = 11,601/10,247 = 1.13. Using this value of *k* leads to very similar results - the estimated $\hat{b}$ is approximately 1.2, and is statistically significant at the 10% confidence level. However, it is somewhat less successful at reconciling the observed amount of bunching below the threshold with the observed missing region above the threshold - the estimated *B* (187 observations) is slightly larger than the estimated *H* (161 observations).

Overall, it appears that the pre-reform counterfactual approach yields results that are more satisfactory in some respects, but are in essence quite similar to our baseline results. Thus, we continue to use the baseline results (and especially the estimate that $\hat{b} \approx 1.7$) in our calculations below.

5.5) Examining Subperiods within the SOX Period

The results in Table 3, Row 2 use the entire SOX period (2003-2015). There is some evidence that SOX compliance costs were particularly large in the early years of implementation

(e.g. Grundfest and Bocher, 2007). Thus, firms' apparent aversion to crossing the $75 million threshold may have arisen due to high initial costs of SOX 404 compliance, and may have disappeared over time as accounting and legal practitioners became more familiar with the regime and compliance costs fell. Unfortunately, it is difficult to estimate the magnitude of bunching for different subperiods within the SOX period, because of the limited number of observations. Estimated bunching parameters are sizeable but not statistically significant, and even the histograms are too noisy to permit definite conclusions.

However, some general patterns are apparent. Bunching appears visible in 2003, the first year in which firms could have adjusted their public float in response to the SEC's announcement of the threshold. It is most pronounced in the years 2004-2006, consistent with the idea that costs were particularly high at the beginning of implementation (which started in 2004). In addition, over this period, remaining below the threshold exempted firms from 404(a) as well as 404(b). The prevalence of bunching during this period also suggests that the overall finding cannot be attributed primarily to the scaled disclosures available to smaller reporting companies from 2008. Substantial bunching remains apparent in the years 2007-2009, despite the reduced compliance costs thought to be associated with Accounting Standard 5 (AS5), and despite the fact that SOX 404(a) could no longer be avoided by bunching. Bunching is also apparent in the years 2010-2015, when the exemption from SOX 404(b) was made permanent. There is no indication that bunching increased in response to the permanent exemption, but it also does not appear that bunching is in the process of disappearing.

5.6) Alternative Explanations for Bunching

5.6.1) The Impact on Public Float of Market Reactions to Crossing the Threshold

The prior literature has in many instances found a negative market reaction to the imposition of SOX 404(b). For instance, Iliev (2010, p. 1190) finds about a 4% abnormal return for small firms that received a temporary exemption from compliance, relative to other firms. While this finding is broadly consistent with our results on bunching below the threshold, it raises a potential problem for our analysis. A firm that crosses the threshold will potentially experience a decline in market value, and hence in public float. Consider a firm initially below the threshold whose insiders are indifferent to the consequences of the $75 million threshold and intend to set public float at precisely $75 million. The decline in price upon crossing the threshold may result in public float remaining below the threshold. For instance, a 4% decline in value would cause

public float to end up at $72 million, despite the insider's intention to set it at $75 million. It is thus possible that some of the observations that appear to be "bunching" below the threshold actually represent firms that sought to locate *above* the threshold.

This possibility may create a bias towards finding bunching below the threshold. To address this potential bias, we redefine the threshold in a manner calibrated to the approximately 4% effect in Iliev (2010, p. 1190). We treat all firms that locate in the range $72-$75 million as having intended to cross the threshold, and the threshold as being $72 million rather than $75 million. This entails that the "true" bunchers can only be found among firms below $72 million of public float, and creates a strong bias against our finding. However, as shown in the fourth row of Table 3, the results are quite robust. The estimated bunching ratio is about 1.3, and is of borderline statistical significance. This provides some reassurance that the basic result is not driven by market reactions to firms' crossing the $75 million threshold.

5.6.2) The Role of Accelerated Filers with Public Float Below the Threshold

It was noted in Section 3 that once a firm passes the $75 million threshold, it remains an accelerated filer unless its public float subsequently falls below $50 million. Thus, an AF will retain this status absent a quite large future reduction in public float. Firms that have previously crossed the threshold and become AFs will thus in most circumstances be indifferent to the threshold in their future decisions about reporting public float. For instance, a firm that is already an AF, currently has public float of $70 million, and can freely choose next year's public float anywhere in the range $73 million to $77 million should be equally likely to locate just below or just above the threshold. When such firms locate just below the threshold, it is not an indication that the regulation is costly; conversely, when such firms locate just above the threshold, it is not an indication that the regulation is beneficial.

Taken by itself, the existence of firms that are indifferent to the threshold because they are already AFs should create a bias against finding bunching below the threshold. However, it is not impossible that due to random variation a disproportionate number of such firms may appear below the threshold. We seek to address this possibility by using firms' responses to the checkboxes in the 10-K filing that identify their filing status (in particular, whether they report being an AF). Using the checkbox data, it is possible to identify firm-year observations in year t in which the firm had public float below $75 million in year ($t$ – 1) but nonetheless reported being an AF in year (t – 1). Such firms are indifferent to crossing the threshold in year t, and so are in a similar

position to our hypothetical firm that is choosing its public float in the range \$73-\$77 million. We exclude 514 firm-year observations that fall into this category, and implement the same bunching analysis as before. The results, shown in the fifth row of Table 3, are quite robust. The estimated bunching ratio is about 1.6, and is statistically significant.

As previously noted, however, the coverage of the checkbox data is limited relative to the numerical public float data. Thus, an alternative approach to implementing this test is to use data on past values of public float to infer whether firms became AFs in the past. In particular, we compute the maximum past value of public float reported by a firm since fiscal year 2002, and use this to infer which firms are likely to already be AFs. Using the same criteria outlined above for the checkbox data, we identify a larger number of observations – 2736 firm-years – for which we would expect the firm to be insensitive to the threshold. Omitting these observations tends to strengthen the result: $\hat{b} \approx 2.3$, and is statistically significant at the 1% level. This provides some reassurance that firms that are already AFs are not biasing the basic result.

5.6.3) The Role of Private Benefits of Control

One of the key features of the model in Section 4 is that decisions about the reporting of public float are made by firms' insiders. While the model focuses on the role of compliance costs, insiders' perceived cost of SOX 404(b) may also involve the potential loss of private benefits of control. To extend the model to the case where private benefits of control are affected by SOX 404(b), assume that the insider's payoff when the regulation does not exist is:

$$U_I = W + PB + (1 - \alpha^*)(V - PB) \tag{20}$$

PB represents (pecuniary) private benefits extracted by the insider. Once the regulation and threshold are introduced, an insider whose firm crosses the threshold is assumed to lose the ability to extract *PB*. The insider's payoff when bunching is now:

$$U_I = W + PB + (1 - \hat{\alpha})\big(V - PB - K_F(\alpha^* - \hat{\alpha})\big) - K_I(\alpha^* - \hat{\alpha}) \tag{21}$$

The critical value of the firm above which insiders will not bunch (V_m) is thus now:

$$V_m = \left(\frac{1}{\alpha^*}\right)\big(C + PF_m - K_F(\alpha^* - \alpha_m) - K_I(\alpha^* - \alpha_m)\big) + K_F(\alpha^* - \alpha_m) + PB \tag{22}$$

It follows that the larger is *PB* (the private benefits lost by crossing the threshold), the larger is V_m and hence the larger is the magnitude of bunching.

In principle, it is thus possible that the loss of insiders' private benefits of control, rather than the compliance costs, may be a reason for bunching below the threshold. However, the past

literature does not suggest that private benefits of control are a significant factor in decisions to avoid SOX 404(b). For example, Gao, Wu and Zimmermann (2009) find no evidence that the NAFs that engage in strategies to remain below the threshold have weaker governance, as measured by standard indices. In addition, *a priori* considerations suggest that private benefits of control are likely to be determined more strongly by state corporate law than by Federal securities regulation.

Nonetheless, we address concerns about bunching being motivated by private benefits by using the entrenchment index (E-index) constructed by Bebchuk, Cohen and Ferrell (2009). The E-index measures how many of six important corporate governance rules apply to a firm, with higher values of the index being associated with worse governance. They also identify firms with dual-class stock, where (as in firms with a high E-index) private benefits of control are likely to be high. Unfortunately, there is limited coverage of the E-index for smaller firms. However, there are 314 nonmissing observations of the E-index for firms with less than $75 million in public float during the SOX period. This appears more than sufficient to generate the observed bunching, if it is indeed driven by private benefits of control. Thus, we implement our bunching analysis while omitting all observations on firms with an E-index of 3 or higher, as well as omitting all firms with dual-class stock (as identified in the Bebchuk, Cohen and Ferrell (2009) data). As shown in the sixth row of Table 2, the results are robust to omitting these firms. The estimated bunching ratio is about 1.7 and is statistically significant. However, the limited availability of governance data for smaller firms should be borne in mind as a caveat.

5.7) Inferring the Magnitude of Net Compliance Costs

It was noted in Section 4 that $\hat{b}$ is a sufficient statistic for the amount by which firms are willing to reduce public float, but not for the net compliance cost *C*. Nonetheless, *C* is a highly policy-relevant parameter that has been the subject of intense academic and public discussion. Thus, we seek in this subsection to draw out the implications of our approach for understanding the compliance cost of SOX 404(b). It is important to emphasize, however, that this exercise requires additional information and assumptions beyond the basic bunching analysis that we have conducted so far.

An implication of the model in Section 4 is that firms with fundamental value of V_m or higher will choose not to bunch. The public float of the marginal nonbunching firm was thus

defined earlier as $PF_m = \alpha^*(V_m - C)$. Rearranging this equation, and noting that $PF_m = \overline{PF} + \Delta PF$, yields:

$$C = V_m - \frac{75 + \Delta PF}{\alpha^*} \tag{23}$$

Of course, we cannot observe V_m (which is the counterfactual market value of the marginal nonbunching firm, were it not subject to SOX 404(b)). However, it is possible to derive a rough estimate based on the reduced-form relationship between market value (using the Compustat variable MKVALT) and public float. This approach is illustrated in Figure 9, which shows observations of market value plotted against public float. The relationship between market value and public float can be estimated for firm-year observations with public float in the range \$50-\$66 million – i.e. below the excluded interval. Extrapolating this relationship yields an estimate of the counterfactual market value within the excluded interval, which can be used to obtain the predicted market value at PF_m. Because it uses data solely from below the excluded interval, this estimate of the counterfactual V_m is in principle free of the effect of SOX 404(b), and is also unaffected by possible bunching behavior within the excluded interval. However, it requires the assumption that the reduced-form relationship between market value and public float would counterfactually be the same within the excluded interval as just below that interval.

This approach can be implemented using the following simple regression:

$$M_{it} = \beta_0 + \beta_1 PF_{it} + \epsilon_i \tag{24}$$

M_{it} is the market capitalization (Compustat variable MKVALT) of firm *i* in year *t*. PF_{it} is the public float of firm *i* in year *t*. This regression is estimated using firm-year observations with public float in the range \$50-\$66 million. The results are shown in column 1 of Table 4. As our estimate of ΔPF is 1.7, we evaluate the counterfactual market value (the predicted value from Equation (24), denoted $\widehat{M}_{it}$) at public float of $PF_m = \$76.7$ million. This yields $\widehat{M}_{it} \approx \$132$ million.

An augmented version of Equation (24), with control variables and firm and year fixed effects, is reported in Column 2 of Table 4. This specification is:

$$M_{it} = \beta_1 PF_{it} + \mathbf{X}_{it}\gamma + \mu_i + \delta_t + \epsilon_{it} \tag{25}$$

As before, M_{it} is the market capitalization and PF_{it} is the public float of firm *i* in year *t*. $\mathbf{X}_{it}$ includes controls for assets and revenue (Compustat variables AT and REVT, respectively), μ_i is a firm fixed effect, and δ_t is a year effect. This regression is also estimated using firm-year observations with public float in the range \$50-\$66 million. Evaluating the predicted value $\widehat{M}_{it}$

from this regression at a public float of $PF_m = \$76.7$ million yields a counterfactual market value of approximately \$134 million.

We use a similar approach to estimate the counterfactual α, with specifications analogous to Equations (24) and (25). The dependent variable is α_{it}, the ratio of public float to market capitalization (Compustat variable MKVALT) for firm *i* in year *t*. As α should be undistorted by bunching behavior above the excluded interval, we use observations with public float in the range \$83-\$150 million to estimate these regressions. The results are shown in Columns 3 and 4 of Table 4. Both these specifications result in a predicted value of α at $PF_m = \$76.7$ million of about 0.6. Using instead observations with public float in the range \$50-\$66 million to estimate these regressions also leads to a predicted value of α at $PF_m = \$76.7$ million of about 0.6.

We combine these estimates and use our bunching estimate of $\Delta PF \approx 1.7$ in Equation (23). This yields a rough estimate of $C \approx 132 - (76.7/0.6) \approx \$4$ million, using the predicted V_m from Equation (24), or $C \approx 134 - (76.7/0.6) \approx \$6$ million, using the predicted V_m from Equation (25). This estimated net compliance cost of \$4-\$6 million applies to firms in the region of the threshold, and is net of any benefits that SOX 404(b) may confer on these firms. It represents about 3%-5% of the value of a typical firm in the region of the threshold. It is thus quite comparable to the approximately 4% market reaction estimated by Iliev (2010, p. 1190). Importantly, the estimate is in present value terms, over the time horizon over which a firm may expect to remain below the threshold in the future. This horizon is likely to be typically a few years; until 2010, the exemption for smaller firms was temporary, and even after the exemption was made permanent in 2010, many firms would expect to grow sufficiently to eventually cross the threshold. As shown by the hazard estimates in Figure 3, there is a substantial probability that a firm below the threshold will cross it within a few years. However, given substantial compliance costs, even merely postponing compliance may be quite valuable in present value terms.[37]

5.8) The Financial and Economic Consequences of Bunching: A Preliminary Exploration

Most of the analysis in this paper has focused on documenting the existence of bunching, and drawing inferences from its magnitude. However, it is also of interest to consider whether bunching (and hence regulatory thresholds in securities law) may affect various dimensions of firm

[37] Chhaochharia, Otto and Vig (2011, fn. 1) illustrate this point by assuming annual compliance costs of \$0.7 million and a 5% discount rate, in which case postponing compliance for 5 years has a \$3 million present value. As their example suggests, a compliance cost of \$4-\$6 million (taking account of indirect as well as direct costs) is not dramatically different from the numbers that have been discussed in the past literature.

behavior. In particular, we provide some preliminary evidence in this subsection on possible consequences for firms' financing choices and for firms' real economic activity (in particular, investment). There is inadequate space to analyze these issues fully here, and further exploration is left for future research.

Issuing equity to outside shareholders increases firms' public float, potentially causing them to cross the threshold for SOX 404(b). On the other hand, issuing debt enables a firm to finance its projects without increasing its public float. Weber and Yang (2016) find evidence that NAFs tend to issue more debt and less equity, relative to AFs, consistent with seeking to avoid crossing the $75 million threshold. To explore this issue in the context of our dataset, we compute the ratio of debt in current liabilities (Compustat variable DCL, representing short-term debt) to total assets (Compustat variable AT) for firm-years close to the threshold (i.e. in the excluded interval). We focus on a subset of observations that we term "repeat bunchers" – firm-year observations in year t that are in the bunching region ($66-$75 million), and for which the firm was also in the bunching region in at least one of year (t – 1) or year (t + 1). As shown in Table 5, the short-term debt ratio is higher for repeat bunchers, compared to other firm-year observations in the excluded interval. A similar, though less pronounced, pattern holds for total debt (the sum of long-term and short-term debt, using Compustat variables DCL and DLTT).

Avoiding increases in public float entails constraints on the growth of outside ownership and equity financing. Thus, bunching behavior may (notwithstanding the availability of debt financing) exacerbate financial constraints for the generally small firms that are close to the threshold. This, in turn, may limit the ability of bunching firms to invest. To explore this possibility, we compute the ratio of capital expenditures (Compustat variable CAPX) to total assets (Compustat variable AT) for firm-years in the excluded interval. As shown in Table 5, this ratio is lower for repeat bunchers, compared to other firm-year observations in the excluded interval, suggesting that bunching firms may be less able to invest.

One approach - very much in the spirit of the bunching methodology - to investigating these questions is to ask whether the debt or investment ratios exhibit unusual patterns for public float bins around the threshold.[38] For each public float bin, we compute the mean ratio of total debt

[38] For instance, Diamond and Persson (2016) develop a formal methodology for analyzing the consequences of bunching for outcome variables other than the variable that is potentially manipulated (which, in our setting, is public float).

(the sum of long-term and short-term debt, using Compustat variables DCL and DLTT) to total assets (Compustat variable AT) for firm-years in that bin. We then construct a counterfactual distribution of this debt ratio by fitting a fifth-order polynomial function of public float to these debt ratios (excluding the $66-$83 million region). Figure 10 plots these debt ratios and the smooth counterfactual derived from the fitted polynomial. Note that this figure shows only the interval $50-$100 million of public float, as the debt ratio distribution becomes quite different above $100 million. It appears that the mean debt ratio seems generally higher than expected for public float bins below the threshold, and is particularly high just below the threshold. While Figure 10 is only illustrative, this pattern suggests that debt ratios may be affected by the regulatory threshold.

A similar approach can be used for the ratio of capital expenditures (Compustat variable CAPX) to total assets, as shown in Figure 11. Here, the bins just below the threshold mostly have lower investment ratios than would be expected based on the counterfactual. However, this is also true of the bins immediately above the threshold. Interpreting this pattern is difficult; however, it may be consistent with a scenario in which both bunching behavior and the compliance costs of crossing the threshold tend to constrain investment activity. This preliminary evidence suggests the possibility that firms' bunching responses to regulatory thresholds may have implications for their financial and economic choices and constraints. This is a potentially important channel through which securities law may affect firm behavior and economic outcomes, but we leave further analysis for future research.

6) Conclusion

This paper brings both a new dataset and a new empirical approach to bear on longstanding but unresolved questions regarding the consequences of the Sarbanes-Oxley legislation. We collect public float data for the universe of reporting entities for fiscal years 1993-2015 by "scraping" this information from 10-K filings. We apply to this new dataset (consisting of about 161,000 observations at the firm-year level) an approach that has not previously been used in the literature on securities regulation. In particular, we document substantial and statistically significant bunching under the $75 million regulatory threshold during the SOX period (2003-2015). Firms in the region of the threshold appear to reduce their public float by about $1.7 million, a response that is consistent with about a $4-$6 million net compliance cost in present value terms. The paper

also presents some preliminary evidence on potential financial and economic consequences of bunching.

Nothing in this paper establishes directly that the observed bunching behavior is motivated by the regulatory threshold. However, there is no evidence of bunching at this threshold (or at an economically equivalent inflation-adjusted value) in the pre-SOX period (1993-2002). Moreover, the result is robust to considering various alternative explanations for bunching. Thus, this paper provides a novel form of evidence on one of the central issues in the study of securities regulation. More generally, its empirical approach is potentially of wide applicability in the empirical analysis of company law and securities law in various countries where bright-line thresholds are used to determine which firms are subject to particular legal rules.

References

Alexander, C. R., S.W. Bauguess, G. Bernile, Y.H.A. Lee, and J. Marietta-Westberg (2013) "The Economic Effects of SOX Section 404 Compliance: A Corporate Insider Perspective" *Journal of Accounting and Economics*, 56, 267-290.

Almunia, M., and López-Rodríguez, D. (2015) "Under the Radar: The Effects of Monitoring Firms on Tax Compliance" Warwick Economics Research Paper Series, No. 1070.

Bartlett III, R. P. (2009) "Going Private but Staying Public: Reexamining the Effect of Sarbanes-Oxley on Firms' Going-Private Decisions" *University of Chicago Law Review*, 76, 7-44.

Bebchuk, L., Cohen, A. and Ferrell, A. (2009) "What matters in corporate governance?" *Review of Financial Studies*, 22(2), pp.783-827.

Black, B. S., H. Jang and W. Kim (2006) "Does Corporate Governance Affect Firms' Market Values? Evidence from Korea" *Journal of Law, Economics, & Organization*, 22, 366-413.

Bushee, B. J., and C. Leuz (2005) "Economic Consequences of SEC Disclosure Regulation: Evidence from the OTC Bulletin Board," *Journal of Accounting and Economics*, 39, 233–264.

Burgstahler, D. and I. Dichev (1997) "Earnings management to avoid earnings decreases and losses" *Journal of Accounting and Economics*, 24(1), 99-126.

Chhaochharia, V. and Y. Grinstein (2007) "Corporate Governance and Firm Value: The Impact of the 2002 Governance Rules" *Journal of Finance*, 62, 1789-1825.

Chhaochharia, V., Otto, C. A., and Vig, V. (2011) "The Unintended Effects of the Sarbanes–

Oxley Act" *Journal of Institutional and Theoretical Economics*, 167(1), 149-164.

Coates IV, J. C. and S. Srinivasan (2014) "SOX After Ten Years: A Multidisciplinary Review" *Accounting Horizons*, 28, 627-671.

Chetty, R., Friedman, J. N., Olsen, T., & Pistaferri, L. (2011) "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records" *Quarterly Journal of Economics*, 126(2), 749.

Dharmapala, D. and V. S. Khanna (2013) "Corporate Governance, Enforcement and Firm Value: Evidence from India" *Journal of Law, Economics, & Organization*, 29, 1056-1084.

Dharmapala, D. and V. S. Khanna (2016) "The Costs and Benefits of Mandatory Securities Regulation: Evidence from Market Reactions to the JOBS Act of 2012" *Journal of Law, Finance, & Accounting*, 1, 139-186.

Dharmapala, D., Slemrod, J., and Wilson, J. D. (2011) "Tax policy and the missing middle: Optimal tax remittance with firm-level administrative costs" *Journal of Public Economics*, 95(9), 1036-1047.

Diamond, R. and P. Persson (2016) "The Long-term Consequences of Teacher Discretion in Grading of High-stakes Tests" Working Paper.

Ferrell, A. (2007) "Mandatory Disclosure and Stock Returns: Evidence from the Over-the-Counter Market" *Journal of Legal Studies*, 36, 215-251.

Gao, F. (2016) "To Comply or Not to Comply: Understanding the Discretion in Reporting Public Float and SEC Regulations" *Contemporary Accounting Research*, 33(3), 1075-1100.

Gao, F., Wu, J. S., & Zimmerman, J. (2009) "Unintended Consequences of Granting Small Firms Exemptions from Securities Regulation: Evidence from the Sarbanes-Oxley Act" *Journal of Accounting Research*, 47(2), 459-506.

Ge, W., A. Koester, and S. McVay (2016) "The costs and benefits of Section 404(b) exemption: Evidence from small firms' internal control disclosures" Working Paper.

Greenstone, M., P. Oyer and A. Vissing-Jorgensen (2006) "Mandated Disclosure, Stock Returns, and the 1964 Securities Acts Amendments" *Quarterly Journal of Economics*, 121, 399-460.

Grundfest, J. A. and S. E. Bocher (2007) "Fixing 404" *Michigan Law Review*, 105, 1643-1676.

Iliev, P. (2010) "The Effect of SOX Section 404: Costs, Earnings Quality, and Stock Prices" *Journal of Finance*, 65, 1163-1196.

Kamar, E., P. Karaca-Mandic, and E. Talley (2009) "Going-Private Decisions and the Sarbanes-

Oxley Act of 2002: A Cross-Country Analysis" *Journal of Law, Economics, & Organization*, 25, 107-133.

Kleven, H. J., and Waseem, M. (2013) "Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan" *Quarterly Journal of Economics*, 128(2), 669-723.

La Porta, R., F. Lopez de Silanes and A. Shleifer (2006) "What Works in Securities Laws?" *Journal of Finance*, 61, 1-32.

Leuz, C. (2007) "Was the Sarbanes–Oxley Act of 2002 really this costly? A discussion of evidence from event returns and going-private decisions" *Journal of Accounting and Economics*, 44(1), 146-165.

Leuz, C., Triantis, A., & Wang, T. Y. (2008) "Why do firms go dark? Causes and economic consequences of voluntary SEC deregistrations" *Journal of Accounting and Economics*, 45(2), 181-208.

Litvak, K. (2007) "The Effect of the Sarbanes-Oxley Act on Non-US Companies Cross-Listed in the US" *Journal of Corporate Finance*, 13, 195-228.

Marx, B. M. (2015) "Dynamic Bunching Estimation and the Cost of Reporting Regulations for Charities" Working Paper.

Nondorf, M. E., Singer, Z., and You, H. (2012) "A study of firms surrounding the threshold of Sarbanes–Oxley Section 404 compliance" *Advances in Accounting*, 28(1), 96-110.

Saez, E. (2010) "Do taxpayers bunch at kink points?" *American Economic Journal: Economic Policy*, 2, 180-212.

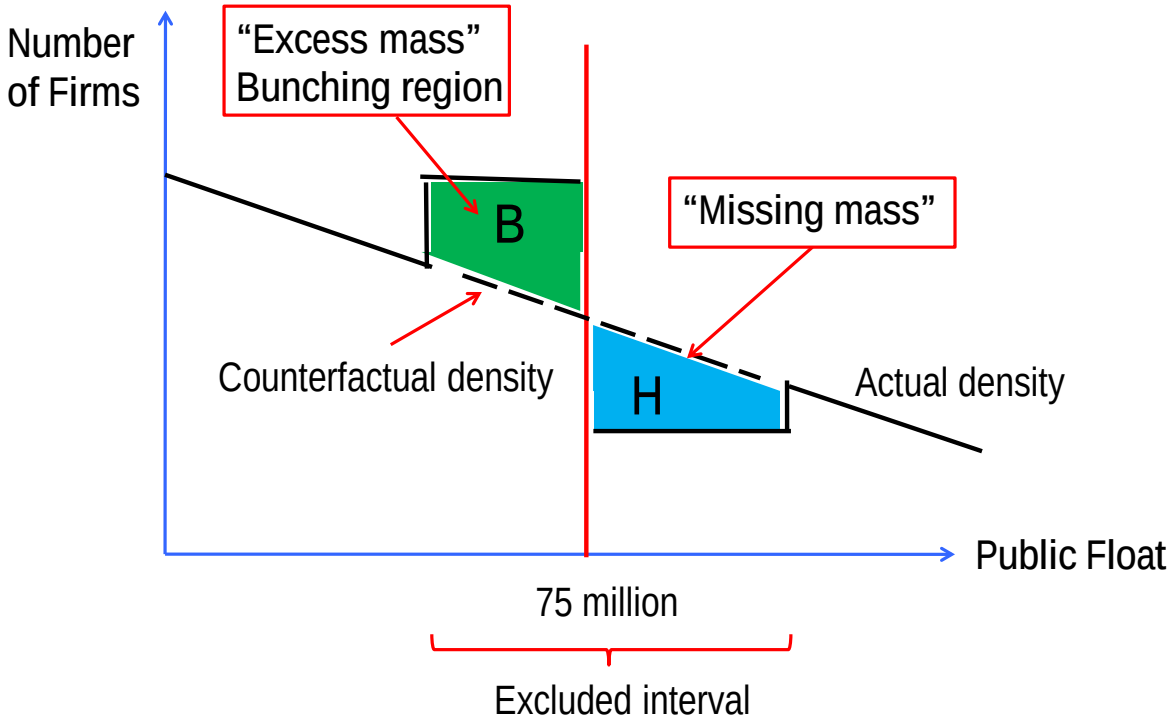
Slemrod, J. (2013) "Buenas notches: lines and notches in tax system design" *eJournal of Tax Research*, 11 (3), 259-283.

Stigler, G. (1964) "Public Regulation of the Securities Markets," *Journal of Business*, 37, 117–142.

Weber, D. and Y. S. Yang (2016) "The debt-equity choice when regulatory thresholds are based on equity values: Evidence from SOX 404" Working Paper.

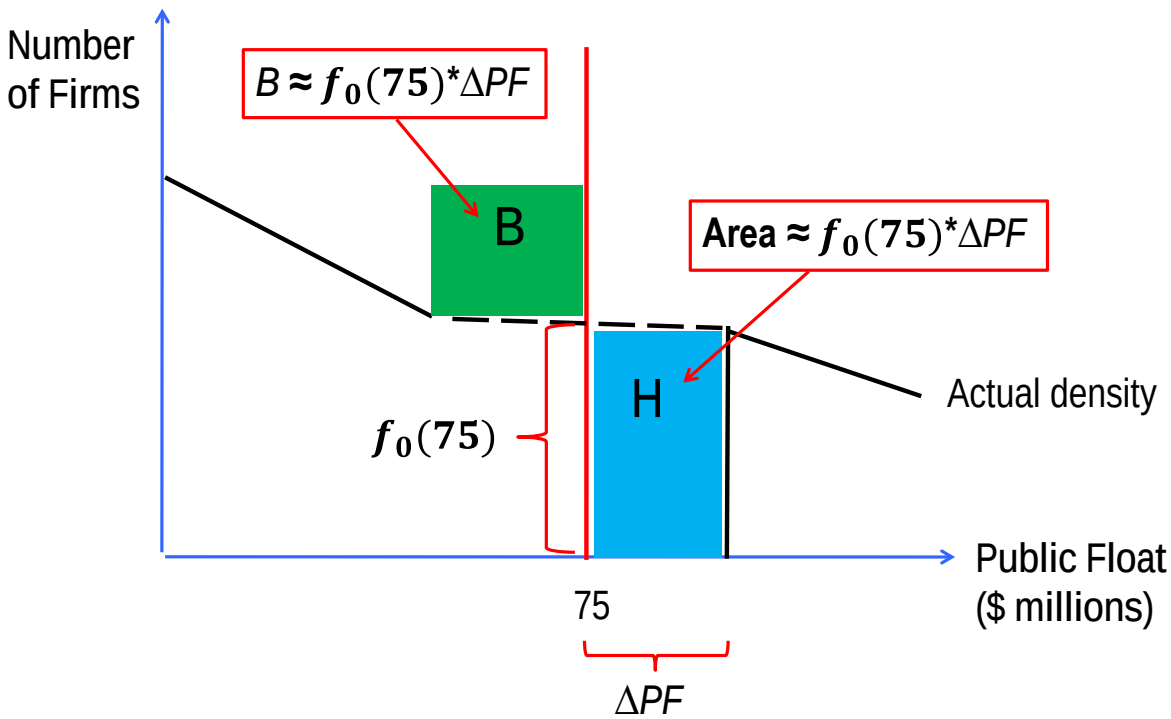
Zhang, I. X. (2007) "Economic consequences of the Sarbanes–Oxley Act of 2002" *Journal of Accounting and Economics*, 44(1), 74-115.

Figure 1: A Stylized Illustration of Bunching Analysis



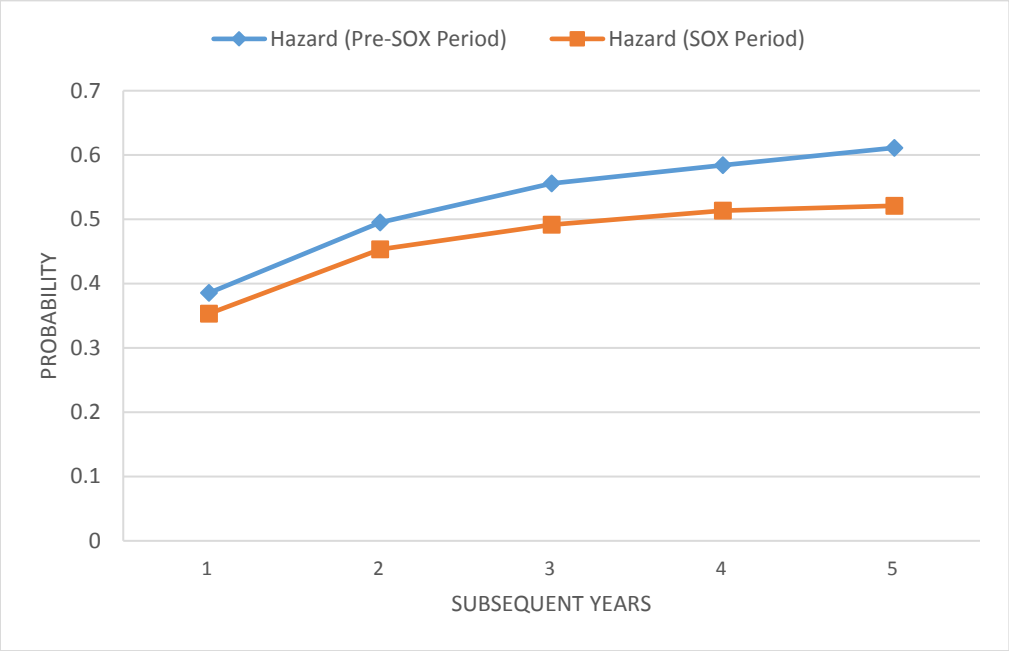
Notes: See the text for an explanation. B is the area of the bunching region, and H is the area of the missing region.

Figure 2: Estimating the Magnitude of Bunching



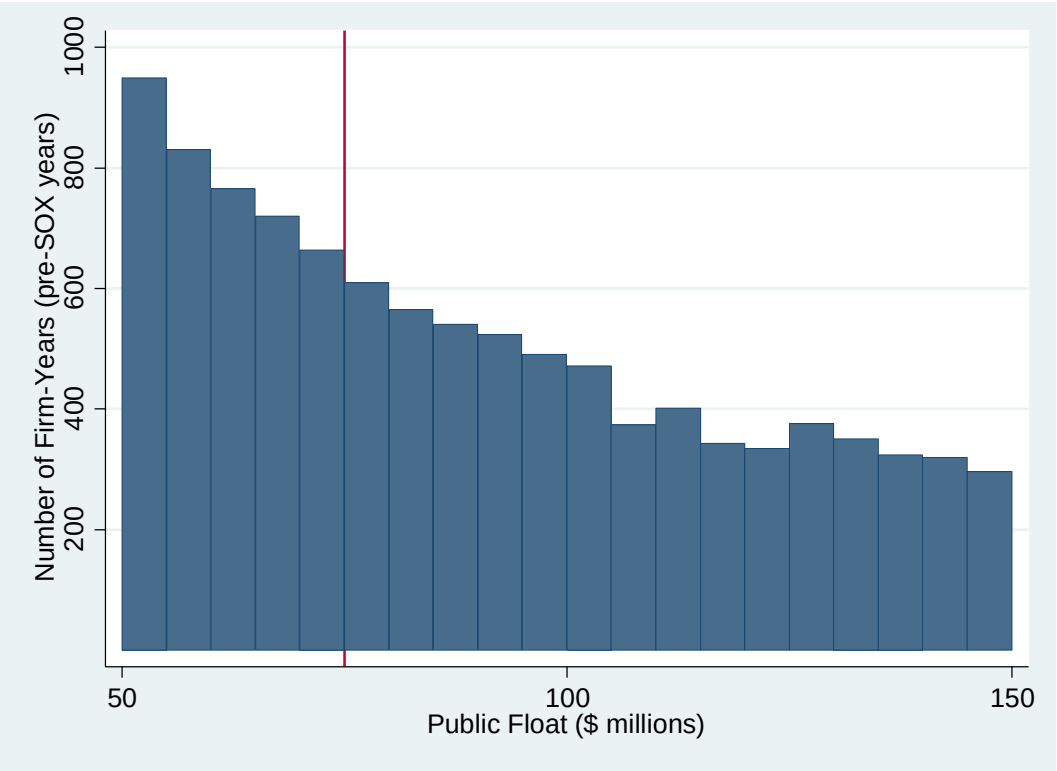
Notes: See the text for an explanation. B is the area of the bunching region, and H is the area of the missing region.

Figure 3: The Probability of Firms with Public Float in the Bunching Region in Year *t* Crossing the $75 Million Threshold in Subsequent Years, in the Pre-SOX and SOX Periods



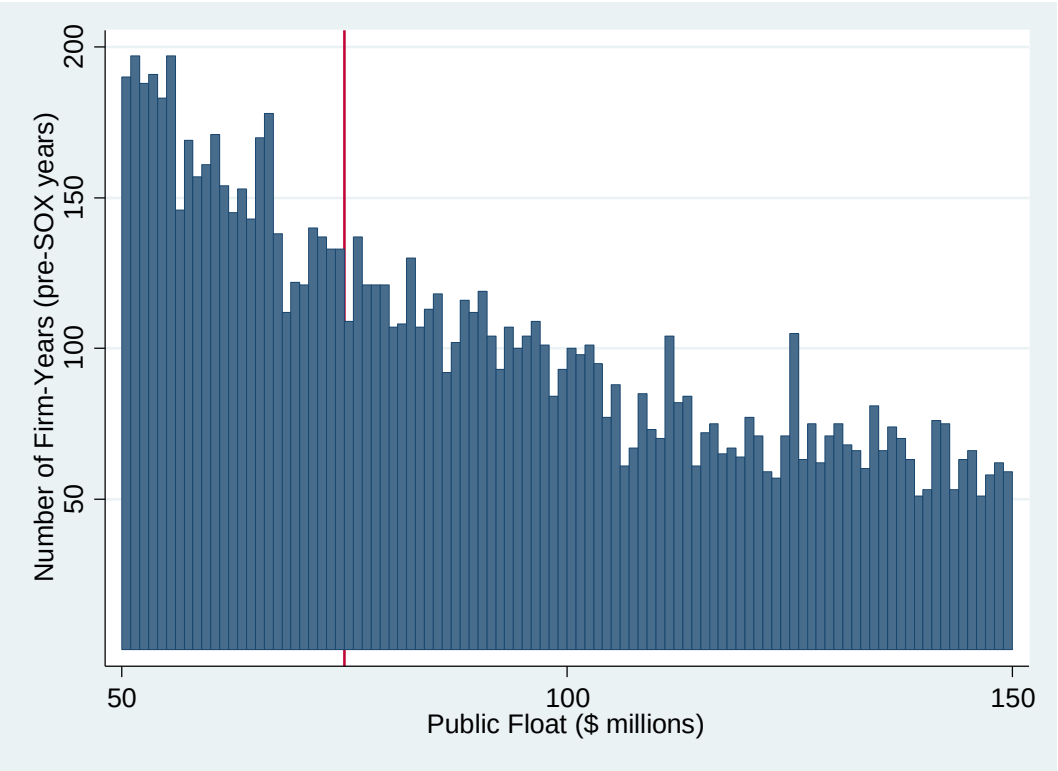
Note: This graph shows the probability that a firm that reports public float in the region $66 million to $75 million in year *t* has reported public float of $75 million or more at least once by year (*t* + *n*), where *n* = 1, 2, 3, 4, and 5 and is shown on the horizontal axis. The "hazard" of crossing the threshold is shown separately for the case where year *t* is in the pre-SOX period (1993-2002) and where it is in the SOX period (2003-2015).

Figure 4: The Density of Firms in the Pre-SOX period (1993-2002) using $5 million Bins



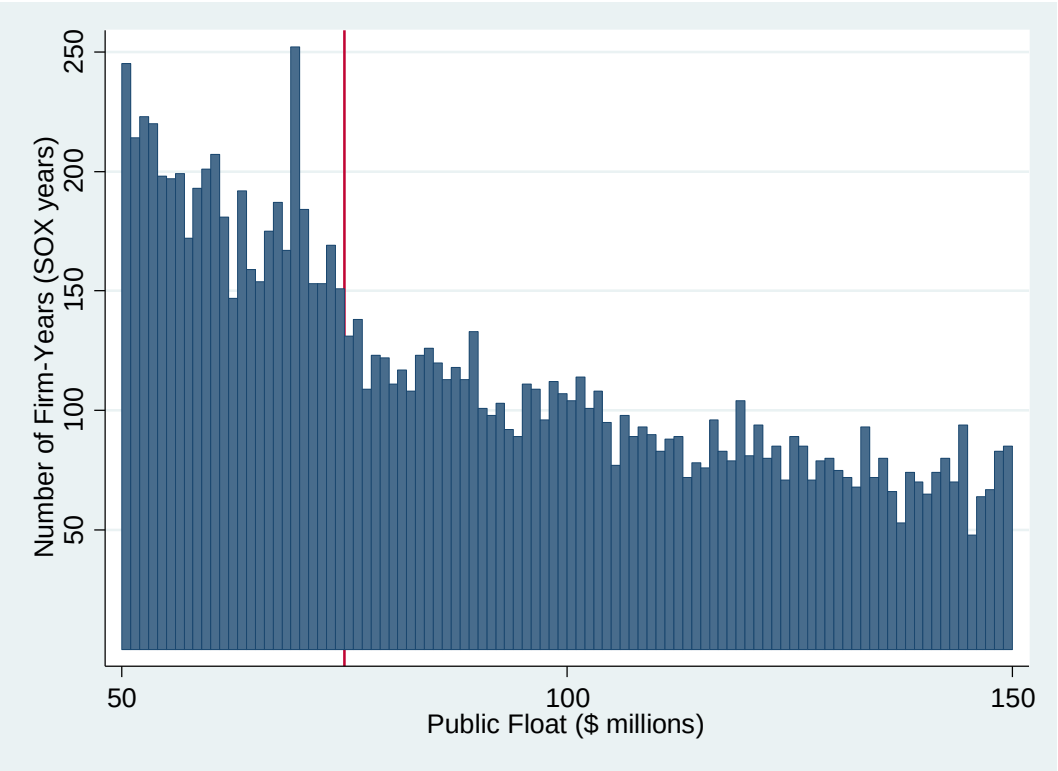
Note: This graph represents the number of firm-year level observations of public float during the pre-SOX period (1993-2002) within each of the bins representing public float observations in the range $50 million to $150 million. The vertical red line represents public float of $75 million. The bin width is $5 million, and the underlying number of public float observations is 10,247.

Figure 5: The Density of Firms in the Pre-SOX period (1993-2002) using $1 million Bins



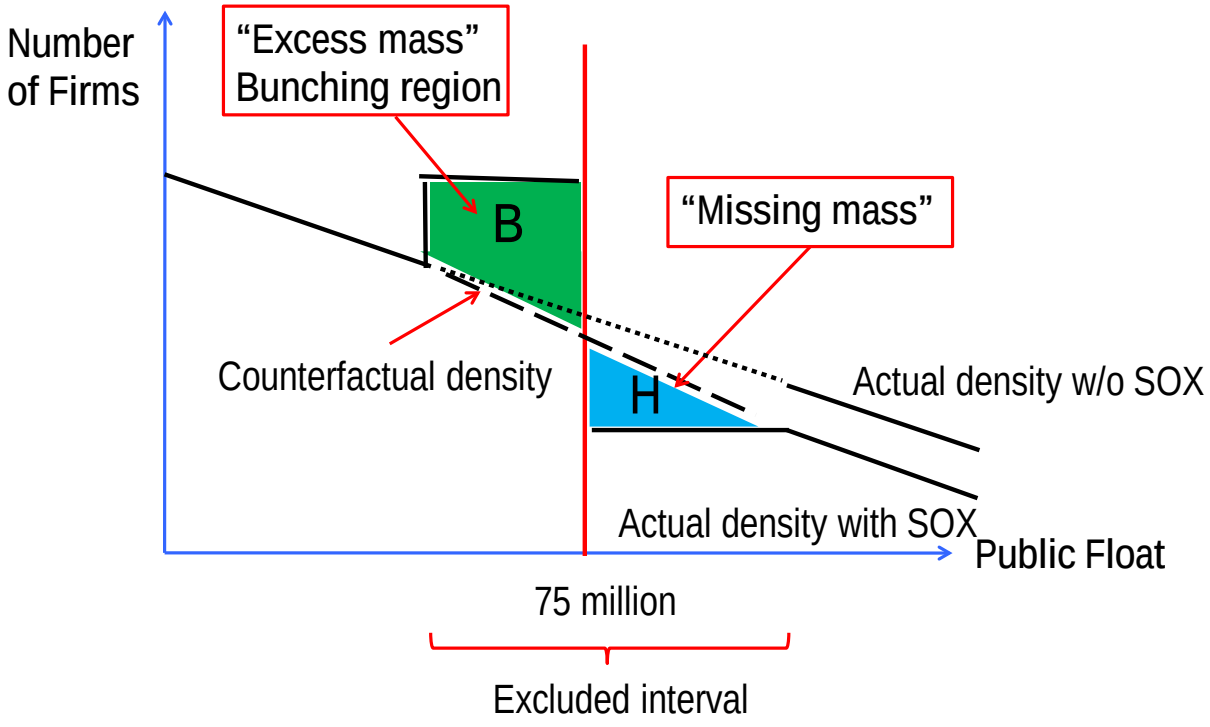
Note: This graph represents the number of firm-year level observations of public float during the pre-SOX period (1993-2002) within each of the bins representing public float observations in the range $50 million to $150 million. The vertical red line represents public float of $75 million. The bin width is $1 million, and the underlying number of public float observations is 10,247.

Figure 6: The Density of Firms in the SOX period (2003-2015)



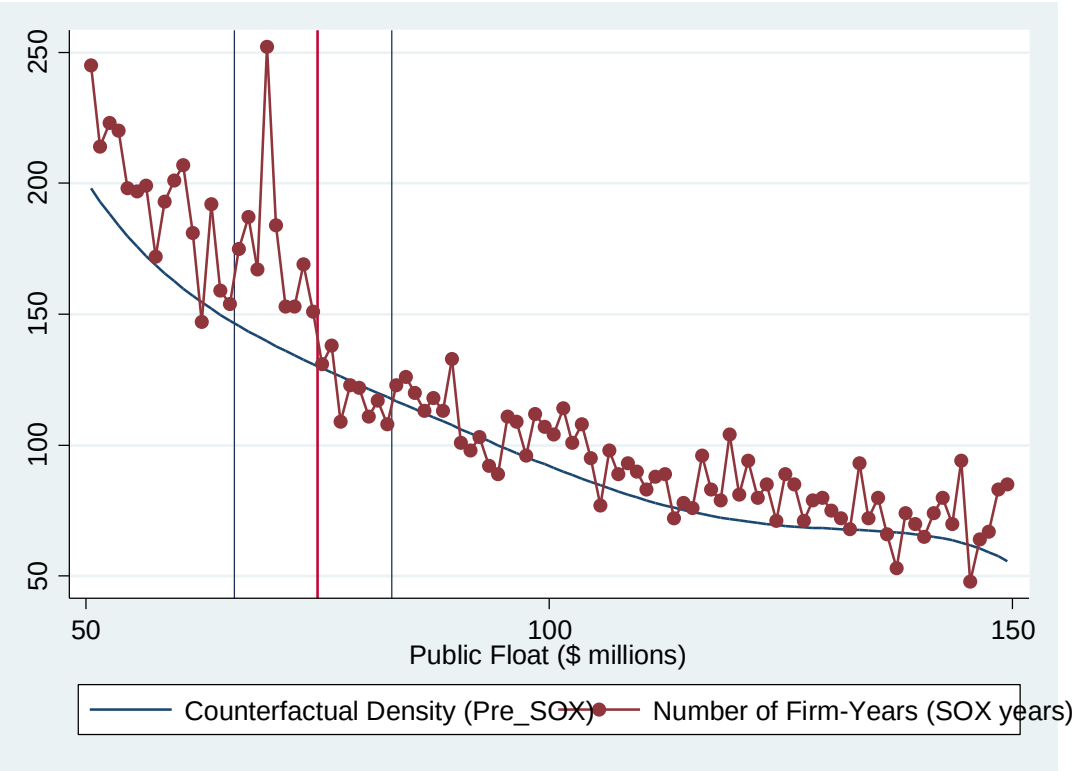
Note: This graph represents the number of firm-year level observations of public float during the SOX period (2003-2013) within each of the bins representing public float observations in the range $50 million to $150 million. The vertical red line represent public float of $75 million. The bin width is $1 million, and the underlying number of public float observations is 11,601.

Figure 7: An Illustration of Alternative Counterfactual Scenarios (with and without SOX)



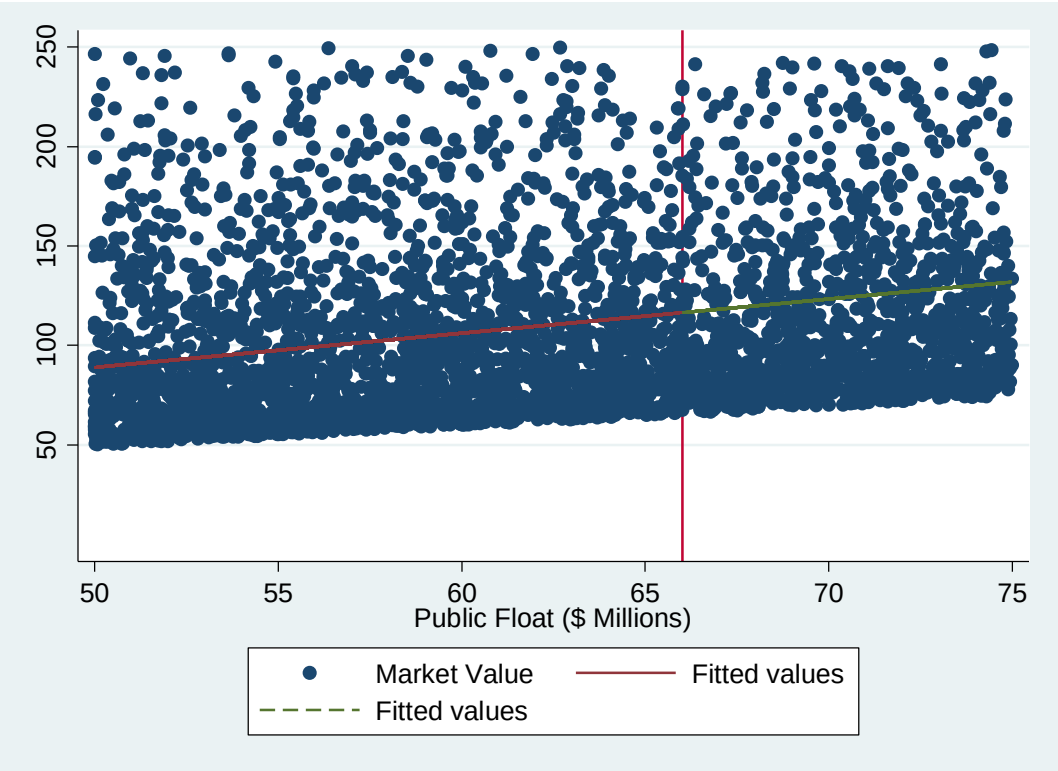
Notes: See the text for an explanation. B is the area of the bunching region, and H is the area of the missing region.

Figure 8: The SOX Period Density Compared to the Pre-SOX Counterfactual Density



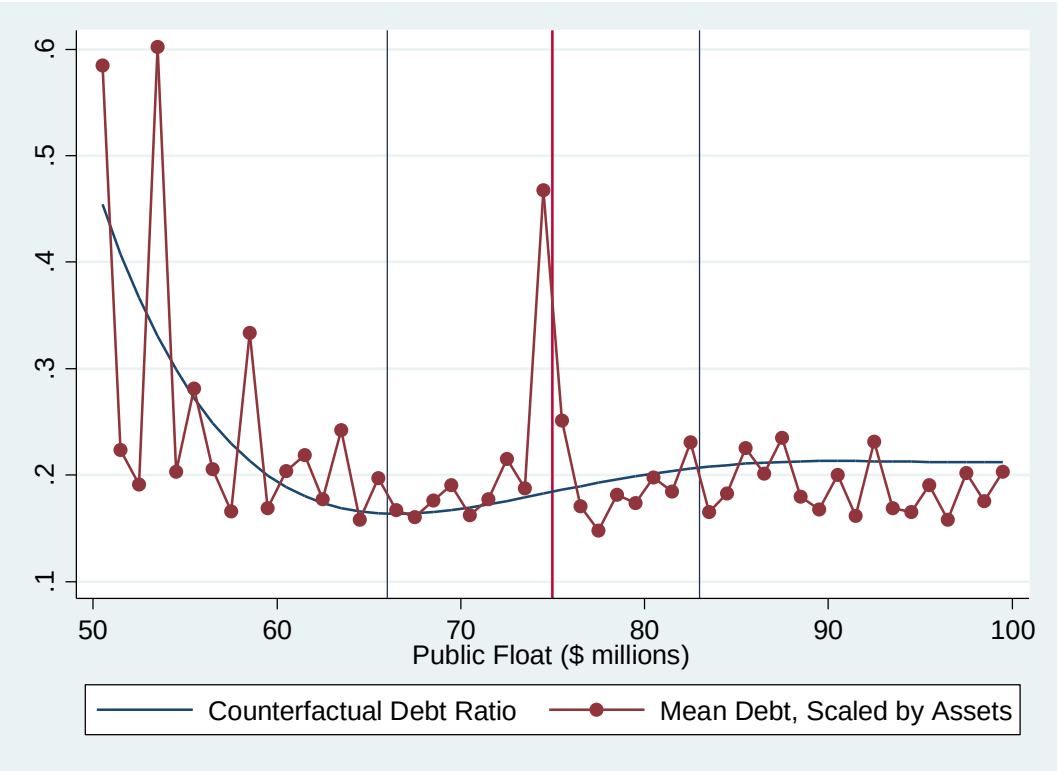
Note: This graph shows a connected line plot of the number of firm-year level observations of public float during the SOX period (2003-2015) within each of the bins representing public float observations in the range $50 million to $150 million. It also shows the counterfactual density, estimated (excluding bins in the range $66 million to $83 million) using the number of firm-year level observations of public float during the pre-SOX period (1993-2002). The vertical red line represent public float of $75 million. The vertical blue lines show the excluded interval (bins in the range $66 million to $83 million). The bin width is $1 million; the underlying number of public float observations during the SOX period (2003-2015) is 11,601 and during the pre-SOX period (1993-2002) is 10,247.

Figure 9: The Relationship between Market Value and Public Float



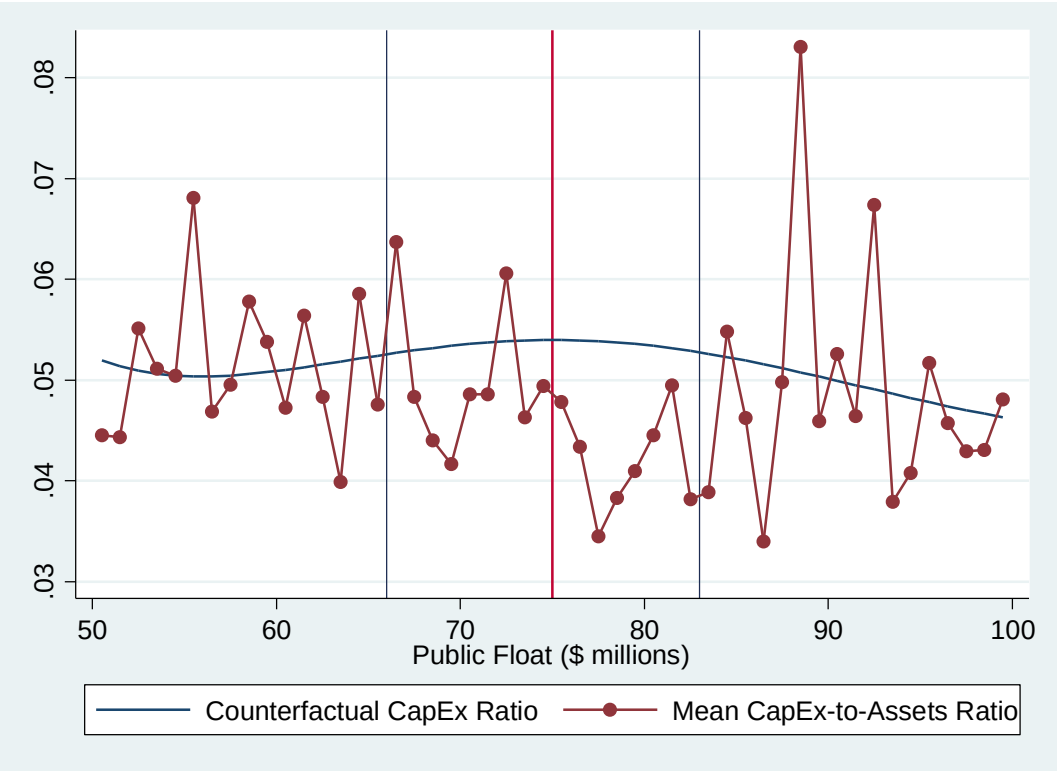
Note: This graph is a scatterplot of observations of market value (Compustat variable MKVALT), plotted against public float. Market value is shown on the vertical axis, and public float on the horizontal axis; each is measured in millions of dollars. Observations for which reported public float exceeds market value are omitted. The vertical red line represents the lower limit of the bunching region ($66 million). The red fitted line represents the estimated linear relationship between public float and market value when public float takes on values from $50-$66 million. The green fitted line is an extrapolation of this relationship to the bunching region.

Figure 10: The Ratio of Total Debt to Total Assets, by Public Float Bin



Note: This graph plots the mean of the ratio of total debt - the sum of debt in current liabilities (Compustat variable DCL) and long-term debt (Compustat variable DLTT) - to total assets (Compustat variable AT), calculated for each bin defined by public float. The counterfactual debt ratio is computed by fitting a fifth-order polynomial function of public float to the mean debt ratios (excluding the $66-$83 million region).

Figure 11: The Ratio of Capital Expenditures to Total Assets, by Public Float Bin



Note: This graph plots the mean of the ratio of capital expenditures (Compustat variable CAPX) to total assets (Compustat variable AT), calculated for each bin defined by public float. The counterfactual capital expenditure ratio is computed by fitting a fifth-order polynomial function of public float to the mean capital expenditure ratios (excluding the $66-$83 million region).

Table 1: Descriptive Statistics for Public Float

Sample	Mean	Median	Standard Deviation	Number of Observations	Number of Firms
Full sample	51,179.3	46.70	$1.99*10^7$	160,988	23,719
Pre-SOX (1993-2002) sample	89.94	85.09	28.45	10,247	4,665
with public float $50-150					
million					
SOX (2003-2015) sample with	89.63	84.10	28.74	11,601	4,222
public float $50-150 million					

Note: Public float is defined as the "aggregate worldwide market value of the voting and non-voting common equity held by its non-affiliates . . . as of the last business day of the issuer's most recently completed second fiscal quarter . . ." (17 CFR 240.12b-2). Public float is obtained using the web scraping approach described in the text, and is measured in millions of US $.

Table 2: Descriptive Statistics for the Ratio of Public Float to Market Value

Sample	Mean	Median	Standard Deviation	Number of Observations	Number of Firms
Full sample	0.64	0.61	0.26	66,998	12,364
Subsample with public float $66-83 million and not subject to SOX 404(b)	0.63	0.67	0.23	1,377	1,101
Subsample with public float $66-75 million in the SOX period (2003-2015)	0.61	0.64	0.24	891	713

Note: Public float and market value are measured in millions of US $. Public float is obtained using the web scraping approach described in the text. Market value is obtained from Compustat (Compustat item MKVALT). The descriptive statistics are calculated excluding observations for which the ratio exceeds one.

Table 3: Estimates of Bunching Parameters

		Bunching parameter ($\widehat{b}$)	**Excess Mass ($\widehat{B}$)**	**Missing Mass ($\widehat{H}$)**	**Number of Observations (N)**
(1)	Pre-SOX Period (1993-2002)	-0.21	-28.45	37.08	10,247
		(0.38)	(53.12)	(32.05)	
(2)	SOX Period (2003-2015):	1.73**	256.66**	68.65**	11,601
	Full Sample	(0.78)	(116.51)	(30.61)	
(3)	SOX Period (2003-2015):	1.16*	182.04*	180.08***	11,601
	Using Pre-SOX Counterfactual	(0.64)	(100.54)	(60.65)	
(4)	SOX Period (2003-2015):	1.34*	204.20*	16.19	11,601
	Using Threshold = $72M	(0.74)	(111.66)	(51.19)	
(5)	SOX Period (2003-2015):	1.60**	225.14**	73.98**	11,087
	Excluding firms that were	(0.78)	(109.36)	(32.37)	
	AFs with public float < $75M				
	in the previous year				
(6)	SOX Period (2003-2015):	1.74**	256.10**	66.84**	11,395
	Excluding firms with high	(0.79)	(116.93)	(29.10)	
	entrenchment index or				
	dual-class stock				

Note: This table reports estimates of the bunching parameter and related measures. Bootstrapped standard errors, computed using 200 replications, are shown in parentheses.
*: significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: Estimates of the Reduced-Form Relationships among Public Float, Market Value, and *α*

	(1)	(2)	(3)	(4)
	Dependent Variable: Market Value		Dependent Variable: *α* (the ratio of public float to market value)	
Public Float ($ Millions)	1.71736***	1.55484***	-0.00009	0.00107***
	(0.436)	(0.294)	(0.0002)	(0.0002)
Total Assets		0.01806		-0.00007***
		(0.016)		(0.00002)
Revenue		0.04024		-0.00003*
		(0.042)		(0.00002)
Firm Fixed Effects?	No	Yes	No	Yes
Year Effects?	No	Yes	No	Yes
Observations	3,769	3,763	5,676	5,670
Number of Firms	2,335	2,332	2,903	2,901
R-squared (within)	0.004	0.096	0.0001	0.092

Note: The regressions in Columns 1 and 2 use only observations for which public float is in the range $50-$66 million. The regressions in Columns 3 and 4 use only observations for which public float is in the range $83-$150 million. Observations for which reported public float exceeds market value are omitted. Robust standard errors are in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Descriptive Statistics for Debt in Current Liabilities and Capital Expenditures

	Debt in Current Liabilities, Scaled by Total Assets			Capital Expenditures, Scaled by Total Assets		
Sample	**Mean**	**St. Dev.**	**No. of Obs.**	**Mean**	**St. Dev.**	**No. of Obs.**
Subsample with public float $66-83 million in the SOX period (2003-2015): "Repeat bunchers"	0.0485	0.0717	207	0. 0396	1,377	119
Subsample with public float $66-75 million in the SOX period (2003-2015): All firm-years other than "repeat bunchers"	0.0380	0.0897	1250	0. 0472	891	1005

Note: This table provides descriptive statistics for the ratio of debt in current liabilities (Compustat variable DCL) to total assets (Compustat variable AT) and the ratio of capital expenditures (Compustat variable CAPX) to total assets (Compustat variable AT). "Repeat bunchers" are firm-year observations in year *t* that are in the bunching region ($66-$75 million), and for which the firm was also in the bunching region in at least one of year (t – 1) or year (t + 1).