How does quasi-indexer ownership affect corporate tax planning?

Shuping Chen*
University of Texas at Austin
shuping.chen@mccombs.utexas.edu

Ying Huang
University of Texas at Dallas
ying.huang1@utdallas.edu

Ningzhong Li
University of Texas at Dallas
Ningzhong.li@utdallas.edu

Terry Shevlin
University of California, Irvine
tshevlin@uci.edu

This Draft: May 15, 2018

*Corresponding author. This paper was previously circulated under the title "Quasi-indexer ownership and corporate tax-planning." We acknowledge comments from Ashiq Ali, Agnes Cheng, John Core (Editor), Amy Hutton, Ross Jennings, Bin Ke, Ken Klassen, Clive Lennox, Alvis Lo, Lillian Mills, Grace Pownall, Sugata Roychowdury, Gil Sadka, Katherine Schipper, Nancy Su, Daniel Taylor, Yen H. Tong, Kenneth Train, Brady Williams, Christopher Yust, Harold Zhang, Jieying Zhang, an anonymous referee, and workshop participants at Boston College, the 6th Conference on Current Research in Taxation EIASM and University of Muenster, Hong Kong Polytech University, Nanyang Technological University, Ohio State University, Texas A&M University, and the University of Texas at Dallas. We thank Russell Investments for providing index membership data, Audra Boone, Matias Cattaneo, Jason Chao and Joshua White for help with the implementation of the regression discontinuity analysis, Brian Bushee for sharing the data on the classification of institutional investor types online, and Scott Dyreng for sharing the foreign tax haven data online. We are grateful to the following board member, tax professional, and asset management professionals for sharing their insights: Ed Blechschmidt, Terri Holbrook, Mark Leitter at TIAA-CREF, Mark Bavoso at Morgan Stanley, and Andrew McCure and Timothy Murray at BlackRock Inc. All errors are our own. Shevlin acknowledges financial support from his Paul Merage Chair in Business Growth at the University of California-Irvine.

How does quasi-indexer ownership affect corporate tax planning?

Abstract

We study whether, and more importantly, through what mechanisms, quasi-indexers affect portfolio firms' tax planning by employing the discontinuity in quasi-indexer ownership around the Russell 1000/2000 index cutoff. Using a regression discontinuity design, we find that higher quasi-indexer ownership leads to greater tax saving. With respect to the mechanisms, we find that the greater tax saving is a result of a focus on improved overall firm performance, not a specific focus on improved tax planning. We further find that the documented tax saving effect is partially due to quasi-indexers' influences on executive equity incentives, corporate governance, and information environment.

JEL Classification: G32, H26

Keywords: tax planning; quasi-indexer; Russell Index assignment; regression discontinuity

How does quasi-indexer ownership affect corporate tax planning?

1. Introduction

Quasi-indexers, investors who hold highly diversified portfolios that are likely to closely mimic an index, are fast becoming the largest institutional investors in the U.S. economy (Craig, 2013). Bird and Karolyi (2017) (hereafter BK) and Khan, Srinivasan, and Tan (2017) (hereafter KST) show that greater quasi-indexer ownership increases portfolio firms' tax avoidance. In this paper, we also study quasi-indexers' impact on investee firms' tax planning. More importantly, we examine through what mechanisms, quasi-indexers affect portfolio firms' tax avoidance. We confirm BK's and KST's findings that higher quasi-indexer ownership leads to greater tax avoidance. However, contrary to BK, our evidence indicates that the tax-saving effect is a result of quasi-indexers' focus on improved overall firm performance, not a result of a specific focus on improved tax planning – quasi-indexers leave it to portfolio firm management to decide on how to improve performance.

A growing number of studies show that quasi-indexers have causal impacts on firms' corporate governance and transparency using the setting of Russell 1000/2000 index assignment (e.g., Appel et al., 2016a; Crane et al., 2016; Boone and White, 2015).[1] This research finds quasi-indexers to be active investors despite their indexing strategy – the fact that they cannot "vote with their feet" gives them incentives to influence managerial actions through multiple channels, such as proxy voting and direct engagement with the management. Quasi-indexers position themselves as long-term investors and their ownership has been shown to be associated with improvements in

[1] While this line of studies may focus on different institutional investors in their main analyses (e.g., passive mutual fund ownership in Appel et al. (2016a) and total institutional ownership in Crane et al. (2016)), the source of identification in the Russell 1000/2000 index setting is that passive investors mechanically follow the indexes for their investment portfolios. Thus, studies that focus on other or total institutional ownership (e.g., Appel et al., 2016a; Crane et al., 2016; BK), all show that their findings are robust to using quasi-indexer ownership.

firms' long-term performance metrics such as Tobin's Q (Appel et al., 2016a). Better performance by portfolio firms enhances the value of quasi-indexers' holdings and helps them attract new fund inflows. As saving taxes can improve after-tax financial performance, we expect higher quasi-indexer ownership could lead to greater tax avoidance.

Regarding the mechanisms through which quasi-indexers affect investee firms' tax savings, we conjecture that the tax saving effect is likely due to quasi-indexers' pushing for better overall firm performance, not specifically tax savings. To the extent quasi-indexers express a desire to the board and managers to improve the firm's after-tax performance and the board and management respond, we could observe a positive relationship between quasi-indexer ownership and tax avoidance. Optimal tax saving strategies are a function of the characteristics of individual firms and developing such strategies requires firm-specific knowledge and tax expertise. Given that quasi-indexers invest in a large number of firms due to their indexing strategy, it is unlikely that they communicate with portfolio firms on specific tax saving actions. In fact, BlackRock and Vanguard leaderships both emphasize that they do not dictate to managers; rather they believe in letting the board and firm management choose the courses of action in addressing their concerns.[2] Consistent with our conjecture, Appel et al. (2016a) and Schmidt and Fahlenbrach (2017) both argue that quasi-indexers may lack the resources necessary for monitoring detailed firm-specific policy choices of every firm in their large, diversified portfolios.

We utilize the cross-sectional discontinuity in quasi-indexer ownership near the Russell 1000/2000 index cutoff created by the annual Russell index assignment to test the causal impact of quasi-indexers on tax planning (e.g., Boone and White, 2015). Each June, Russell Investments

[2] See, for example, BlackRock CEO Larry Fink's 2016 Corporate Governance Letter to CEOs, and "Vanguard and Black Rock Plan to Get More Assertive with Their Investments", The *Wall Street Journal*, March 4, 2015.

ranks all exchange-traded U.S. common stocks based on their May 31 float-adjusted market capitalization, and assigns the 1,000 largest firms to the Russell 1000 index and the next 2,000 largest firms to the Russell 2000 index. Firms close to the cutoff on either side are very similar in size and the inclusion in the index is quasi-exogenous to corporate policies. The value-weighted nature of the indices results in quasi-indexers, who likely benchmark these indices, mechanically holding larger positions in stocks at the top of the Russell 2000 and smaller positions at the bottom of the Russell 1000. The resulting discontinuity in quasi-indexer ownership at the Russell 1000/2000 index cutoff is caused by Russell's index assignment, not firm policies. This quasi-exogenous variation in quasi-indexer ownership allows us to draw causal inference of quasi-indexers' impact on firms' tax planning.

We capture tax planning using two established measures of tax avoidance: GAAP effective tax rate (hereafter, ETR) and cash ETR (e.g., Chen et al., 2010; Cheng et al., 2012).[3] Lower GAAP and cash ETRs are associated with greater tax avoidance. Using a sharp regression discontinuity design (RDD), we find that the higher quasi-indexer ownership in firms at the top of Russell 2000 relative to those at the bottom of Russell 1000 leads to significantly lower ETRs. The documented effects are also economically significant. The GAAP (cash) ETR is significantly lower by 2.7-3.7% (around 4.8%) for firms at the top of the Russell 2000 than those at the bottom of the Russell 1000.

To provide evidence on our conjecture that the documented tax saving effects are likely due to quasi-indexers pushing for better overall firm performance, not tax savings specifically, we conduct the following mechanism analyses. First, if quasi-indexers push for better overall firm performance, we expect that they also have a positive impact on portfolio firms' pretax

[3] In untabulated analyses, we employ two alternative measures of tax avoidance: the book-tax difference measure proposed by Manzon and Plesko (2002) and the book-tax difference measure developed by Desai and Dharmapala (2006). We find qualitatively similar results using these two alternative measures.

performance. Using the same RDD design, we find that the higher quasi-indexer ownership in firms at the top of Russell 2000 relative to those at the bottom of Russell 1000 leads to significantly better pretax performance. Relative to the bottom Russell 1000 firms, the top Russell 2000 firms have higher pretax ROA by 2.6-4.2%, higher pretax margin by 22.6-27.0%, and greater asset turnover by 12.8-19.7%. We further find weak evidence that the effect of quasi-indexer ownership on tax savings is stronger for firms with better performance, measured as both pre-tax and after-tax performance.

Next, we examine the proxy voting records and guidelines of BlackRock and Vanguard, currently the two largest quasi-indexers, to provide evidence on whether quasi-indexers propose and/or vote on tax-related proposals directly. The voting records indicate that BlackRock and Vanguard rarely directly vote on tax-related proposals. Further, the proxy voting guidelines of BlackRock (Vanguard) rarely (never) discuss tax-related voting policies. These findings, combined with our findings on pretax firm performance, suggest that quasi-indexers' impacts on tax savings are likely due to their pushing for better overall firm performance, not their pushing for tax savings specifically.

Given that prior studies using the same setting have shown that quasi-indexers have causal influences on firms' corporate governance and information environment (e.g., Appel et al., 2016a; Boone and White, 2015), we further examine whether and to what extent quasi-indexers influence tax savings through such indirect channels, including CEO equity incentives, corporate governance, and information environment, which are possible economic determinants of tax avoidance (e.g., Wilson and Wilde, 2017). To the extent that quasi-indexers may influence these firm characteristics for reasons other than achieving better firm performance (e.g., Boone and White,

2015), this analysis sheds additional light on the mechanisms underlying the tax saving effects we document.

We start by examining whether CEO equity incentives, corporate governance, and information environment exhibit discontinuity around the index cutoff. For a measure that exhibits a discontinuity, we add it to our RDD regressions of ETRs as an additional explanatory variable. If the effect of quasi-indexer ownership on tax avoidance is mitigated – the estimated coefficient on the indicator of Russell 2000 membership is reduced in magnitude – after adding the new variable, we conclude that the variable is likely an indirect mechanism. After identifying all possible indirect mechanism variables, we include all of them in the RDD regression of ETRs to examine whether quasi-indexer ownership has a residual effect on tax savings after accounting for the indirect effects. We find that quasi-indexers' influences on CEO equity incentives (CEO option vega), certain aspects of corporate governance (e.g., board independence and CEO power), and information environment (e.g., PIN – the probability of informed trading) are indirect mechanisms through which quasi-indexer ownership affects tax avoidance. These indirect mechanisms, however, only partially explain the tax saving effects we document. Quasi-indexer ownership has a residual effect on tax avoidance after accounting for all identified indirect mechanisms related to CEO equity incentives, corporate governance, and information environment. This residual effect is likely due to quasi-indexers directly pushing for improved overall firm performance.

Our paper contributes to the literature on the role of ownership structure on tax planning (e.g., Chen et al., 2010; Badertscher et al., 2013; McGuire et al., 2014) by demonstrating the causal impact of quasi-indexers on corporate tax policies, and more importantly, by providing evidence on the mechanisms through which these investors can affect tax avoidance. Our study differs from BK and KST, which use the same Russell index setting to examine the effect of quasi-indexer

ownership on tax avoidance, in at least two important dimensions.[4] First and most importantly, we provide a detailed mechanism analysis showing that the tax saving effect is due to quasi-indexers pushing for better overall firm performance, not tax savings specifically. This finding is in contrast to BK's argument that "many of these institutional investors make specific claims about their role in shaping corporate tax strategy" (see their page 29). While KST make a similar conjecture as ours, they provide no empirical analyses on this issue.

Second, we examine whether and to what extent quasi-indexers can affect portfolio firms' tax avoidance through three indirect mechanisms: CEO equity incentives, corporate governance, and information environment. While BK also provide mechanisms analyses (in their section IV), they do not intend to examine, as we do, whether a variable of interest is an indirect mechanism in the sense that quasi-indexer ownership affects the variable, which further affects tax avoidance. Rather, they study whether the tax avoidance effect varies cross-sectionally with the *ex ante* level of agency issues in portfolio firms (their Table 9) or changes in director turnover and CEO option awards (their Table 10). To conclude that a variable is an indirect mechanism (in our sense) for the documented tax effect, one has to first demonstrate a discontinuity of the variable around the index cutoff. BK do not examine whether there is an index-inclusion effect on the variables of interest and whether the effect indirectly contributes to the documented tax avoidance effect. In addition, the equity incentive measure BK use, equity compensation as a percentage of total compensation, is a likely poor measure of CEO equity incentives (Core and Guay, 1999). KST, in their robustness section, briefly describe their analyses on CEO delta and vega, two measures of equity incentives (Core and Guay, 1999), but do not tabulate their results. They acknowledge that

[4] We summarize our major differences with BK and KST in Appendix A.

their mechanism analysis is brief and this represents an opportunity for future research (see their page 117).

Our analysis of whether quasi-indexers push for overall performance or tax savings specifically raises an important question for the tax avoidance literature in general. Prior studies on tax avoidance generally ignore whether tax avoidance arises because of a primary focus on improved firm performance or is a result of a specific focus on tax savings (e.g., Chen et al., 2010; Cheng et al., 2012; BK; KST). We call for future research to explicitly explore this issue.

There is currently a debate about how to use an RDD for causal inference with Russell index reconstitutions (e.g., Appel et al., 2016a,b; Schmidt and Fahlenbrach, 2017; Crane et al., 2016; Young, 2018; Wei and Young, 2017). Issues include: (a) the bandwidth choice, (b) which investors are assumed to be affected (e.g., mutual funds that specifically index Russell vs. quasi-indexers), and (c) the fact that the proprietary May end market cap that Russell uses to assign index membership (the forcing variable) is not observable to researchers. In our analysis, we use Calonico, Cattaneo, and Titiunik's (2015) approach to implement a sharp RDD with fixed bandwidths and assume that all quasi-indexers are affected by the index reconstitution.

There are in essence four approaches to dealing with the measurement error that results from non-observability of the forcing variable: (1) proxy for the forcing variable using Russell's June end float-adjusted market capitalization (or equivalently the Russell June index weights), (2) proxy for the forcing variable using CRSP May end market capitalization, (3) use an IV approach to address the measurement error following Appel et al. (2016a), and (4) use an IV approach to address the measurement error following Crane et al. (2016). We provide in Internet Appendix A detailed discussion and comparisons of these approaches using our assumptions about the bandwidth and that all quasi-indexers are affected. We conclude that, due to the measurement error

arising from the non-observability of Russell's May end market capitalization, none of the approaches is perfect and each approach could lead to a potential bias.

In our tests we use approach (1), the June index weights, which is consistent with Boone and White (2015) and Khan et al. (2017). Like Boone and White (2015), we want to "evaluate the effect of institutional holdings after the reconstitution date, which is driven by June portfolio weights" (see their page 515). This approach achieves the best pre-treatment covariate balance of controls relevant to tax avoidance. We detail in Internet Appendix A that if we instead proxy for the forcing variable using CRSP May end market capitalization, results are much weaker. Using Crane et al.'s (2016) IV approach, we find qualitatively similar but weaker results. We do not find any results using Appel et al.'s approach. Thus, we caveat that inference in this area in general and in our paper specifically appears sensitive to method choices and this is a worthwhile area for future research.

The rest of the paper is organized as follows. Section 2 discusses prior studies and develops empirical predictions. Section 3 discusses our main research design. Section 4 presents the baseline RDD results. Section 5 presents our analyses on whether the documented tax saving effects are due to quasi-indexer pushing for better overall performance, or for tax saving specifically. Section 6 examines CEO equity incentives, corporate governance, and information environment as indirect mechanisms of the documented tax saving effects. Section 7 concludes.

2. Background and prediction

Researchers have hypothesized that ownership structure is an important determinant of tax policy because different owners have different mechanisms in resolving the agency conflicts with managers (Shackelford and Shevlin, 2001; Desai and Dharmapala, 2008; Hanlon and Heitzman,

2010). Managers trade off the benefits of tax savings (e.g., higher after-tax cash flows) against a variety of costs, including agency costs (risk- and effort-averse managers' incentives are not always aligned with shareholders), financial reporting costs (to the extent lower reported taxes to the IRS mean lower reported earnings to shareholders), lower stock prices (to the extent stock price is a function of earnings), reputation costs (Hanlon and Slemrod, 2009, Graham et al., 2014) and political costs (e.g., witness the "tax-shaming" in popular press), and increased IRS scrutiny (Mills, 1998). Given the effort involved and potential risks managers face in pursuing tax saving strategies, managers' private assessment of the cost-benefit tradeoff of tax avoidance may diverge from shareholders' assessment of this tradeoff. Large investor presence can alter managers' assessment of this cost-benefit tradeoff, either leading to greater tax avoidance (e.g., Cheng et al., 2012) or less tax avoidance (e.g., Chen et al., 2010; Badertscher et al., 2013).

We study the causal effect of ownership by a particular group of institutional investors, quasi-indexers, on portfolio firms' tax avoidance. Quasi-indexers, investors who hold highly diversified portfolios that are likely to closely mimic an index, are fast becoming the largest institutional investors in the U.S. economy (Craig, 2013). Recent research shows that quasi-indexers affect corporate governance and policies: quasi-indexer ownership leads to fewer votes supporting management proposals, more support for shareholder proposals, more independent directors and fewer dual class shares (Appel et al., 2016a). Quasi-indexer ownership also leads to richer information production by managers, resulting in lower information asymmetry and higher liquidity (Boone and White, 2015).

Quasi-indexers position themselves as long-term investors. In his 2016 Corporate Governance Letter to CEOs, BlackRock CEO Larry Fink urges "that every CEO lay out for shareholders each year a strategic framework for long-term value creation." Quasi-indexers have

incentives to encourage firm managers toward maximizing (after-tax) financial performance which enhances long-term share value through multiple channels, such as proxy voting and direct engagement with the management. Managers, at the same time, have incentives to respond to quasi-indexer demands, for a variety of reasons: out of a sense of fiduciary duty or gratitude, to gain support in future elections, to enhance future job prospects, or because their shareholders will help them fend off take-over threats (Elhauge, 2016). In the tax avoidance setting, the presence of large institutional investors not only mitigates managerial effort-aversion, but also shields managers from the downside of seeking tax avoidance by mitigating the perception of managerial rent extraction from opaque tax avoidance activities (Desai and Dharmapala, 2006; Chen et al., 2010). This in turn gives managers more incentives for tax savings.

Quasi-indexers' influences on portfolio firms' tax savings could be due to quasi-indexers pushing for better overall firm performance, not their pushing for tax savings specifically. Optimal tax saving strategies are a function of the characteristics of individual firms and developing such strategies requires firm-specific knowledge and tax expertise. Given that quasi-indexers invest in a large number of firms due to their indexing strategy, it is unlikely that they provide tax-specific guidance to portfolio firms. Instead, it is much less costly to directly push for better overall performance and let the board and management themselves choose appropriate corporate strategies, including tax saving strategies, to achieve the performance goal.

The above explanation is consistent with arguments in recent studies using the Russell index setting (e.g., Appel et al., 2016a; Schmidt and Fahlenbrach, 2017). Appel et al. (2016a) caution that while quasi-indexers may be effective at engaging in widespread, low-cost, monitoring of firms to ensure they comply with what they consider to be best governance practices, these investors "might lack the resources necessary to research and influence corporate policies that are

inherently more firm-specific" (page 130) due to their large, diversified portfolios. Consistent with this argument, they find little evidence that quasi-indexer ownership is associated with corporate policies related to investment, capital structure, or cash holding. Similarly, Schmidt and Fahlenbrach (2017) argue that more passive ownership may affects corporate governance negatively "when it comes to *high-cost* governance activities such as monitoring of mergers and acquisitions, the choice of board members, or the accumulation of titles that often happen outside of annual general meetings and require continuous monitoring" (page 287, italics original).

This explanation is also consistent with anecdotal evidence. BlackRock reports 1,522 direct engagements with portfolio firms for fiscal year 2015. Similarly, Glenn Booraem, principal and head of corporate governance at Vanguard, characterized Vanguard's program as "one of engagement",[5] and is reported to speak with hundreds of executives every year to press Vanguard's views behind the scenes.[6] These quasi-indexers usually push managers toward better performance without dictating the specific course of action managers should take. Both BlackRock CEO Larry Fink and Vanguard CEO F. William McNabb III publicly issue letters to U.S. CEOs in support of this engagement approach that raises issues with firms while at the same time allowing boards and firm managers to choose the appropriate course of action. Mr. Fink in his 2016 Corporate Governance Letter to CEOs states: "…we believe companies are usually better served when ideas for value creation are part of an overall framework developed and driven by the company, rather than forced upon them in a proxy fight" (emphasis original). In a 2015 Wall Street

[5] Source: Burr 2012. "Money managers increasing activism on corporate governance – but quietly." *Pension and Investment*.

[6] Source: "Passive fund manager Vanguard turns activist in some board votes" by Ross Kerber, http://www.reuters.com/article/2013/09/13/vanguard-proxyvotes-idUSL2N0H00YV20130913.

Journal interview, Mr. McNabb III states: "We're indifferent as to how a board chooses to engage. What's important is that it engages."[7]

Quasi-indexers could also affect firms' tax planning indirectly through their impact on executive equity incentives, corporate governance, or information environment. An indirect mechanism needs to satisfy two conditions: it can affect tax avoidance and itself is affected by quasi-indexer ownership. For the first condition, prior research has shown that certain aspects of managerial incentives, corporate governance, and information environment are associated with tax avoidance (e.g., Rego and Wilson, 2012; McGuire, Wang, and Wilson, 2014; Chen, Chiu, and Shevlin, 2017). For the second condition, prior studies have shown that quasi-indexer ownership has a causal influence on certain aspects of corporate governance (Appel et al., 2016a; Schmidt and Fahlenbrach, 2017) and firm disclosure and information environment (e.g., Boone and White, 2015), and it is conceivable that they may also affect firms' executive equity incentives.

These indirect mechanisms, however, are ambiguous *ex ante* for two reasons. First, while there is evidence that executive equity incentives can affect tax avoidance (e.g., Rego and Wilson, 2012), Desai and Dharmapala's (2006) theoretical model shows that the relationship between equity-based compensation and tax aggressiveness is ambiguous. Furthermore, it is unclear whether and how quasi-indexers affect executive equity incentives. Second, although it has been shown that quasi-indexers can influence corporate governance and firm transparency (e.g., Boone and White, 2015; Appel et al., 2016a), the impacts of corporate governance and transparency on tax avoidance are ambiguous (e.g., Minnick and Noga, 2010; Robinson, Xue, and Zhang, 2012; Chen and Lin, 2017; Chen, Chiu, and Shevlin, 2017). Given the ambiguous nature of these indirect

[7] "Vanguard and BlackRock Plan to Get More Assertive With Their Investments", March 4, 2015, the *Wall Street Journal*.

mechanisms, we treat their effects as open empirical questions and provide detailed analyses in Section 6.

3. Research design

Following prior literature (e.g., Chen et al., 2010; Chyz et al., 2013), we capture tax avoidance using effective tax rates (ETRs). Our first measure is the GAAP ETR (*GAAP ETR*), which reflects tax avoidance activities through permanent book-tax differences, such as investments in foreign countries with lower foreign tax rates (provided foreign source earnings are classified as permanently reinvested), investment in tax exempt and tax favored assets, and participating in transactions that generate tax losses but not book losses (Chen et al., 2010). Our second measure is cash ETR (*Cash ETR*), which reflects both permanent and temporary – deferral of revenue and acceleration of expenses for tax but not book – book-tax differences. Lower values of *GAAP ETR* and *Cash ETR* indicate greater tax avoidance.

Our primary empirical approach for identifying the effect of quasi-indexer ownership on tax avoidance is a sharp regression discontinuity design (RDD) (e.g., Boone and White, 2015; KST), which takes advantage of a known cutoff determining treatment assignment. Our base empirical model is

$$Y_{it} = \alpha + \beta R2000_{it} + \sum_{n=1}^{N} \Theta_n {Rank_{it}}^n + \sum_{n=1}^{N} \lambda_n R2000_{it} \times {Rank_{it}}^n + \varepsilon_{it}, \quad (1)$$

where Y_{it} is one of our tax avoidance measures. In our mechanism analyses in Sections 5 and 6, Y_{it} also refers to each of the pre-tax performance measures and mechanism variables. $R2000_{it}$ is an indicator variable for firms in the Russell 2000, and $Rank_{it}$ is the assignment variable, defined as the integer distance from the index cutoff each year centered at zero around the Russell

1000/2000 cutoff, with positive (negative) values associated with Russell 2000 (1000) firms.[8,9] Following Boone and White (2015), we define $Rank_{it}$ using the actual index assignment based on June weights provided by Russell Investments. The coefficient on $R2000_{it}$ (β) represents the estimated treatment effect of Russell 2000 index assignment, i.e., the effect of exogenously higher quasi-indexer ownership on tax avoidance. A negative estimate of β is consistent with higher quasi-indexer ownership inducing more tax avoidance.

As is standard in the RDD literature (Roberts and Whited, 2013), we control for the higher order polynomial of the assignment variable (*Rank*) and allow the functional form to be different for observations above and below the Russell 1000/2000 cutoff. This choice is reflected in the inclusion of $\sum_{n=1}^{N} \Theta_n Rank_{it}{}^n$, and the interaction terms, $\sum_{n=1}^{N} \lambda_n R2000_{it} \times Rank_{it}{}^n$. Equation (1) shows the RDD is a *cross-sectional* design employing discontinuity in the data.[10] We follow Boone and White (2015) to fit a local third (N=3) order polynomial estimate using a triangular kernel to the left and right of the index cutoff and using the robust bias-corrected *z*-statistic developed in Calonico et al. (2015). Our results are qualitatively similar when we set N equal to 2 (untabulated).

Researchers face a common tradeoff when implementing an RDD: bias versus efficiency, with a small bandwidth associated with smaller bias but lower power (Roberts and Whited, 2013). Widening the area of analysis around the cutoff mitigates power concerns, but increases the bias

[8] The variable *Rank* is coded as 1 for the first (largest) firm in Russell 2000 and -1 for the last (smallest) firm in Russell 1000.

[9] The ETRs are calculated for the fiscal year with at least six months in the twelve-month period after each year's index reconstiution (i.e., July of year *t* to June of year *t+1*). For instance, for index assignment in year 2000 (*t*=2000), if a firm has Feburary as its fiscal year end, Y_{it} is measured for the fiscal year 2000 (March 2000 to Feburary 2001). If a firm has December as its fiscal year end, Y_{it} is still measured for the fiscal year 2000.

[10]The treatment effect in RDD is not identified by a difference between average outcomes of the treatment and control groups, but rather by a discontinuity in the functional relationship between the outcome and the assignment variable at the cutoff point (Trochim and Donnelly, 2008; Roberts and Whited, 2013), which is identified in the very careful control for the effect of the assignment variable $Rank_{it}$. If the terms related to $Rank_{it}$ are dropped from equation (1), β will capture the difference in the average outcome between the treatment and control groups.

in the estimated treatment effect. We employ two bandwidths: 500 and 300; that is, we include in the analysis only 500 (300) firms to the left and right of the Russell 1000/2000 threshold.[11] Our choices of bandwidth are in line with the outcome of the optimal bandwidth selection procedure of Calonico et al. (2014). For instance, when estimating equation (1) with N equal to 3, Calonico et al.'s (2014) procedure selects the optimal bandwidth of 446 and 585 for *GAAP ETR* and *Cash ETR*, respectively.

Note that selection bias, one of the biggest threats to cross-sectional studies, will only be a problem for an RDD if it induces a discontinuity in the bivariate relationship between the outcome variable and its other determinants at the cutoff point (Lee and Lemieux, 2010). Prior studies using the Russell index assignment setting have shown that there is no discontinuity for major pre-assignment firm characteristics around the index cutoff (e.g., Boone and White, 2015).[12] In untabulated analyses, we further demonstrate the absence of discontinuity for major pre-assignment firm characteristics that are potentially related to corporate tax policy (e.g., Chen et al., 2010; Chyz et al., 2013), including firm size (*Market Cap*), profitability (*ROA*), leverage ratio (*Leverage*), market-to-book ratio (*Market-to-book*), net operating loss carryforward (*NOL Dummy* and *NOL Change*), foreign operations (*Foreign Income*), income from equity method (*Equity income*), capital intensity (*PPE*), and the amount of intangible assets (*Intangible*). Appendix B provides detailed definitions of these variables.

4. Baseline empirical results

4.1. Data and descriptive evidence

[11] Our results are qualitatively similar when we use a bandwidth of 400 (untabulated).

[12] The term "pre-assignment" means that these firm characteristics are measured prior to the index reconstitution date (June 30). It does not mean that it is a comparison of these characteristics for the pre- and post-reconstitution periods.

We obtain the members of the Russell 1000 and 2000 indexes and their float-adjusted market value of equity and portfolio weights from 1996 to 2006 from Russell Investments. Similar to recent research using this setting (e.g., Boone and White, 2015), our sample period ends in 2006 due to the banding policy implemented by Russell after 2006 that potentially reduces the local continuity of firm assignment around the threshold.[13] Data on accounting information and market information are sourced from Compustat and CRSP, respectively. Institutional ownership data are from Thomson Reuters 13f File. We merge the Russell data with CRSP and Compustat through CUSIP, the firm identifier provided by Russell Investments. The full sample consists of 33,00 firm-year observations for 6,361 unique firms. We then exclude REITs (Real Estate Investment Trusts) because they are flow-through entities and by design have zero or close to zero ETRs. When restricted to the bandwidth of 500, the sample size reduces to 10,194 firm-year observations for 2,679 unique firms. Table 1 summarizes the sample selection process.

Table 2 reports the summary statistics for ownership percentages of different institutional investors, measured at the end of September following Boone and White (2015), and our tax planning measures for the 100 firms on the top of the Russell 2000 and the 100 firms at the bottom of the Russell 1000.[14] Firms on the top of the Russell 2000 have significantly larger quasi-indexer ownership (by 13.2%) and transient institutional ownership (by 7.1%) than firms at the bottom of the Russell 1000, whereas the difference in dedicated institutional ownership is statistically insignificant. Figure 1 presents graphical evidence on the discontinuity in institutional ownership.

[13] The banding policy after 2006 was designed to maintain some continuity in the indices, which may violate the local continuity assumption of an RDD because the selection of firms into the indices is related to characteristics other than market capitalization (Crane et al., 2016).

[14] Although we use a larger bandwidth in our RDD analyses, we use a narrow bandwidth of 100 to present these summary statistics because in this simple comparison of means, we do not control for the polynomials of the assignment variable (*Rank*). As Roberts and Whited (2013) note, controlling for the polynomials of the assignment variable allows the use of more observations around the cutoff.

The RDD plots represent local sample means using 10 non-overlapping evenly-spaced bins on each side of the threshold following the methodology described in Calonico et al. (2015). The fitted lines represent a third-order polynomial regression curve (that is, equation (1) with N =3). Consistent with the summary statistics in Table 1, we find that total institutional ownership jumps up by around 31% from the Russell 1000 to Russell 2000 at the cutoff. This is primarily due to quasi-indexer ownership, which jumps up by around 23%. Transient institutional ownership also jumps up by around 9%, whereas dedicated institutional ownership drops slightly.[15]

Table 2 also shows that firms on the top of the Russell 2000 have significantly smaller GAAP and cash ETRs. Figure 2 presents the RDD plots for each ETR measure against index rankings around the Russell 1000/2000 cutoff. Figures 2-a and 2-b show, respectively, that firms at the top of the Russell 2000 have lower *GAAP ETR* and *Cash ETR* than those at the bottom of the Russell 1000 after controlling for the effect of the assignment variable (i.e., the variable *Rank*). Taken together, Table 2 and Figures 1 and 2 provide evidence that the Russell 2000 membership is associated with higher quasi-indexer ownership and greater tax avoidance.

4.2. Regression analysis

Table 3 presents the results of estimating equation (1) for the two ETR measures. We report the main results in Panel A and sensitivity tests in Panel B. For each measure, we include the third (N=3) order polynomial of the assignment variable (*Rank*) using the bandwidths of 300 and 500.

[15] Boone and White (2015) point out that, in contrast to the case of quasi-indexers, the index assignment does not force transient investors to mechanically alter their portfolio allocations, though the rebalancing activities of indexing institutions around this event may create price pressure and trading opportunities that could lead transient investors to initially hold firms at the top of Russell 2000 versus those at the bottom of Russell 1000. More importantly, unlike other investors, transient investors are less likely to directly influence managers' actions – their short-term horizon reduces their opportunities to exert influence. In addition, managers associate transient investors with undesirable effects on stock price volatility, and are therefore unlikely to alter policies to cater to this clientele (Beyer, Larcker, and Tayan, 2014). Our results are robust after controlling for transient investor ownership (see Section 4.2.).

In columns 1 and 2 of Panel A, the estimated treatment effect is negative and significant across all regressions. The treatment effect ranges between -2.7% and -3.7% and the average effect is -3.2%, suggesting that, on average, the GAAP ETR is significantly lower by 3.2% for firms at the top of the Russell 2000 than those at the bottom of the Russell 1000. This effect is economically significant, representing around 9% of the average GAAP ETR in the sample (i.e., averaged over 1000 firms, with 500 firms on each side of the cutoff). The results for *Cash ETR* in columns 3 and 4 of Panel A are qualitatively very similar to those for *GAAP ETR*. The treatment effect is around -4.8% and is significant in each regression. This effect represents around 18% of the average cash ETR in the sample.

While in a standard RDD analysis there is no need to control for other covariates, for sensitivity tests, in Panel B, we further include in equation (1) the following control variables: the third order polynomial of transient institutional investor ownership (*Transient*, *Transient*^{2}, and *Transient*^{3}),[16] industry and year fixed effects, as well as the firm characteristics that are potentially associated with *ETRs* (e.g., Chen et al., 2010; Chyz et al., 2013), namely, *Market Cap*, *Market-to-book*, *ROA*, *Leverage*, *NOL Dummy*, *NOL Change*, *Foreign income*, *PPE*, *Intangible*, and *Equity income*. We estimate the augmented equation (1) using OLS as the Calonico et al. (2015) approach does not allow for control variables. For parsimony we report only the results on *R2000*. The sample size drops substantially after including these control variables. Despite the smaller sample sizes, the effect of quasi-indexer ownership on *GAAP ETR* and *Cash ETR* continues to be significantly negative after including these control variables.

[16] As Table 2 and Figure 2 indicates that firms at the top of the Russell 2000 also have larger transient institutional investor ownership than those at the bottom of the Russell 1000, to attribute the documented treatment effects in Panel A to the larger quasi-indexer ownership, we control for the third order polynomial of transient institutional investor ownership.

Our results are consistent with KST and BK. Using a similar RDD analysis, KST find that firms at the top of Russell 2000 have lower GAAP ETR by 5.1 percentage points and lower cash ETR by 7.0 percentage points than firms at the bottom of Russell 1000. Using an instrumental variable approach, BK document that for a 1 percentage point increase in institutional ownership due to inclusion in the Russell 2000 around the index cutoff, a firm's cash ETR and GAAP ETR fall by 0.2-0.3 and 0.1-0.2 percentage points, respectively.

5. Do quasi-indexers push for overall firm performance or tax savings specifically?

As we discuss in Sections 1 and 2, it is not necessary for quasi-indexers to possess tax management expertise and give firm-specific tax strategy guidance to portfolio firms for them to save taxes. Quasi-indexers may directly communicate with the board and managers a desire for improved overall financial performance, and a focus on improving overall performance can motivate managers to achieve better after-tax performance through lower cash taxes and reported tax expense as well. KST makes a similar conjecture, arguing that "as long as tax avoidance is one strategy employed by managers to improve after-tax performance, we expect a positive relation between tax avoidance and institutional ownership even in the absence of institutional owners specifically and explicitly promoting tax avoidance" (see their page 102). However, they provide no direct empirical evidence. In contrast, BK argue that "many of these institutional investors make specific claims about their role in shaping corporate tax strategy" (see their page 29).

KST show that quasi-indexer ownership leads to higher net income margin and the likelihood of meeting or beating analysts' earnings expectations, and Appel et al. (2016a) show that quasi-indexer ownership increases ROA. However, by focusing on after-tax performance measures, they cannot really assess whether the improved measures reflect quasi-indexers' push

for overall improved firm performance, and tax avoidance is just a byproduct of this overall push, or reflect quasi-indexers' specific focus on improved tax planning that drives greater tax avoidance and higher after-tax net income margin and ROA.[17]

To provide evidence on this issue, we examine the effect of quasi-indexer ownership on pretax ROA (*Pretax ROA*, pretax income divided by average assets) and its two components: pretax margin (*Pretax Margin*, pretax income divided by sales) and asset turnover (*Asset Turnover*, sales divided by average assets). If we observe a positive effect of quasi-indexer ownership on pretax ROA and its components, it implies that the lower taxes could be the result of a general focus on firm performance. If quasi-indexer ownership has no significant effect on pretax ROA and its components, then the lower taxes are likely a result of a specific focus on tax saving.

We use *Pretax ROA*, *Pretax Margin*, and *Asset Turnover* as the dependent variables and estimate equation (1). Table 4 reports results for this analysis. In columns 1 and 2, when the dependent variable is *Pretax ROA*, we find the estimated coefficients on *R2000* are consistently positive and significant in all three regressions. These results suggest that firms at the top of the Russell 2000 have higher pretax ROA than firms at the bottom of the Russell 1000. The difference ranges between 2.6% and 4.2% and the average difference is 3.4%. This difference is economically significant.

Columns 3 and 4 present the results for pretax margin and columns 5 and 6 for asset turnover. The sample sizes are smaller than the corresponding regressions for pretax ROA due to missing values for sales.[18] Consistent with our finding for pretax ROA, the estimated coefficient

[17] KST define net income margin as after-tax income divided by sales. Appel et al. (2016) define ROA as net income divided by total assets.

[18] The fraction of observations with missing sales in our sample is higher than all Compustat firms in the same sample period. When restricting the sample to firms with non-missing sales, we find consistent results for the analysis of pretax ROA.

on *R2000* is positive and significant for all six regressions, suggesting that firms at the top of Russell 2000 have both higher pretax margin and asset turnover than firms at the bottom of Russell 1000. The effect for pretax margin ranges between 22.8% and 27.0%; the effect for asset turnover is between 13.7% and 20.5%. Both effects are economically significant. Overall, these results suggest that the top Russell 2000 firms exhibit better non-tax related performance than the bottom Russell 1000 firms, consistent with our conjecture that quasi-indexers may communicate to the board and managers a desire for better overall performance.

Given that ROA is equal to *Pretax ROA*×(1-*GAAP ETR*), our findings on the effect of quasi-indexers on pretax ROA in Table 4 and the effect on GAAP ETR in Table 3 have the following implication on after-tax profitability. For an average Russell 1000 firm within the bandwidth 500, pretax ROA is 5.9% and GAAP ETR is 34.7%, the average pretax ROA effect (3.4%) and the average GAAP ETR effect (-3.2%) suggest for an average Russell 2000 firm within the bandwidth 500 ROA will be 66.1% higher.[19] This estimate is consistent with the finding from estimating equation (1) using ROA as the dependent variable: the average coefficient on *R2000* is 2.7%, compared to the average ROA of 3.8% for Russell 1000 firms within the bandwidth 500.

We further explore whether the tax avoidance effect is stronger for firms with better pretax or after-tax performance.[20] It is conceivable that the tax avoidance effect is stronger when firms are more profitable because these firms' managers are likely more incentivized to improve performance. We partition the full sample into subsamples of firms with high vs. low performance (based on the sample median), measured with both pretax ROA and ROA, and separately estimate equation (1) for each subsample. To mitigate the concern that the documented difference in the

[19] These effects of *R2000* suggest that ROA for an average Russell 2000 firm within the bandwidth 500 will be (5.9%+3.3%)[1-(34.7%-3.2%)]=6.4%. The relative difference in ROA will be 6.4%/[5.9%×(1-34.7%)]-1=66.1%.
[20] We thank the anonymous reviewer for proposing this analysis.

effect of quasi-indexers may be due to the different scale of the dependent variable (i.e., *GAAP ETR* or *Cash ETR*) in the two subsamples, we normalize the dependent variable to having standard deviation 1.

Table 5 present the results of this analysis. Columns 1 to 4 report the results for GAAP ETR. We find that the difference in the estimated coefficients on *R2000* for high vs. low performance groups is insignificant when firm performance is measured with pretax ROA or ROA.[21] Columns 5 to 8 report the results for cash ETR. For high performance firms, measured with either pretax ROA (column 5) or ROA (column 7), the coefficient on *R2000* is more negative than that for the low performance firms (columns 6 and 8, respectively), and the difference is significant.[22] Overall, the evidence from Table 5, albeit somewhat weak, is consistent with our prediction that the tax avoidance effect is stronger for more profitable firms.

While the analyses in Tables 4 and 5 are consistent with our conjecture that quasi-indexers push for overall firm performance, we cannot rule out the possibility that quasi-indexers also push for tax savings specifically. BK argue that one way quasi-indexers play a specific role in shaping corporate tax strategy is through proxy voting (see their footnote 1 on page 29). To explore this possibility, we examine quasi-indexers' proxy voting records and guidelines.

First, we investigate whether quasi-indexers *explicitly* vote on tax-related issues. We extract the voting records of BlackRock and Vanguard, currently the two largest quasi-indexers,

[21] As the dependent variable in each regression in Table 5 is normalized to having standard deviation 1, the coefficient on *R2000* should be interpreted relative to the standard deviation of the un-normalized dependent variable in each subsample. Thus, the magnitudes of the coefficients on *R2000* in Table 5 are not directly comparable to those in Table 3. When un-normalized ETR measures are used as the dependent variables, the coefficients on *R2000* in Table 5 are comparable to those in Table 3 in magnitude and the inferences based on the un-normalized measures are qualitatively similar to those based on the normalized measures (untabulated).

[22] One might expect differences in GAAP ETR because this is the ETR that affects reported after-tax performance. However, the improvement in pretax performance could arise from actions that give rise to temporary differences and not permanent differences - the former (latter) does not (does) affect GAAP ETRs.

from SEC filings (Form N-PX) for the reporting period July 1, 2005 to June 30, 2006, and search for tax-related proposals. We search for the word "tax" in each Form N-PX and then manually check whether each incidence of the word is related to a tax-related proposal. While it is possible that this approach misses implicit tax considerations when BlackRock and Vanguard vote for non-tax-related proposals, it captures all proposals in which a tax issue is explicitly considered. For instance, at the shareholder meeting of Colonial Bancgroup held on April 16, 2006, Vanguard votes for a proposal to "re-approve the material terms of the performance goals under Colonial Bancgroup, Inc. 2001 long-term incentive plan in order that certain awards under such plan be afforded beneficial tax treatment under Section 162(M) of the Internal Revenue Code of 1986."[23] This proposal is likely to be specifically related to a tax issue.

Table 6, Panel A summarizes the findings from this analysis. For the reporting period July 1, 2005 to June 30, 2006, BlackRock files only one Form N-PX for the reporting period under the name "BlackRock Funds." Vanguard files ten Form N-PX for the reporting period, one for each fund. Both BlackRock and Vanguard vote actively. For instance, BlackRock attends the shareholder meetings of 2,796 firms and submits votes for 11,235 proposals; Vanguard Total Stock Market Index Fund submits votes for 10,908 proposals of 3,957 firms. However, the proposals are rarely related to tax issues. We are not able to identify any tax-related proposals (out of 11,235) for BlackRock and are able to identify only 42 tax-related proposals (out of 39,312) for Vanguard. This evidence suggests that quasi-indexers' rarely directly vote on tax-related issues.

Next, to address the concern that the voting records in Form N-PX may miss proposals related to tax issues if the voting summary does not contain the word "tax", we examine the proxy voting guidelines of BlackRock and Vanguard. As we are not able to find their voting guidelines

[23] Source: Form N-PX of Vanguard Small-Cap Value Index Fund filed on August 31, 2006.

related to our sample period, we rely on their current guidelines.[24] Blackrock's guidelines discuss major factors the company considers when voting for proposals related to the following topics: i) boards and directors; ii) auditors and audit-related issues; iii) capital structure, mergers, asset sales, and other special transactions; iv) remuneration and benefits; v) social, ethical, and environmental issues; and iv) general corporate governance matters. Vanguard's guidelines cover similar topics.

For each company, we search for the keyword "tax" in its voting guidelines. In Vanguard's 10-page guidelines, the keyword "tax" never appears, suggesting that Vanguard does not have any specific voting policies related to tax issues. In BlockRock's 19-page guidelines, the keyword "tax" appears four times. All sentences containing "tax" are presented in Panel B of Table 6. The guidelines briefly mention tax issues when discussing voting considerations for poison pill plans, employee stock purchase plans, golden parachutes, and pay-for-performance plans. Given that the guidelines cover 6 topics and 46 subtopics and tax issues are not treated as a separate topic or subtopic but are only briefly mentioned 4 times in the above contexts, we conclude that BlackRock, like Vanguard, does not provide specific guidance for tax savings.

Collectively, our analyses of BlackRock's and Vanguard's proxy voting records and guidelines suggest that quasi-indexers do not specifically push for tax savings through proxy voting. This evidence is inconsistent with BK's argument that quasi-indexers play a specific role in shaping corporate tax strategies. We acknowledge that quasi-indexers may vote on certain proposals that are not directly related to tax issues but have tax saving implications. However, these cases are exactly consistent with our argument that quasi-indexers want to push for better overall firm performance but not tax savings specifically.

[24] BlackRock's proxy voting guidelines are available at https://www.blackrock.com/corporate/en-kr/literature/fact-sheet/blk-responsible-investment-guidelines-us.pdf; Vanguard's at https://about.vanguard.com/investment-stewardship/voting-guidelines.

6. Do quasi-indexers influence tax savings indirectly?

Section 5 provides evidence consistent with quasi-indexers pushing for better overall firm performance which then leads firms to lower their taxes to improve after-tax performance, but not pushing for tax savings specifically. However, prior studies using the same setting have shown that quasi-indexers have causal influences on firms' corporate governance and information environment (e.g., Appel et al., 2016a; Boone and White, 2015). To the extent that these factors may also affect tax avoidance, we examine whether quasi-indexers indirectly influence tax savings through such indirect channels, including CEO equity incentives, corporate governance, and information environment (e.g., Wilson and Wilde, 2017). Because quasi-indexers may influence these firm characteristics for reasons other than achieving better firm performance, this analysis sheds additional light on the mechanisms underlying the tax saving effects we document. For instance, Boone and White (2015) argue that quasi-indexers demand greater firm transparency because their diverse stock holdings make gathering private information on portfolio firms more costly. Their impacts on information environment may indirectly affect the cost-benefit tradeoff related to tax savings faced by managers.

We measure CEO equity incentives with option holding vega (*CEO Vega*) and option and stock holding delta (*CEO Delta*) (e.g., Core and Guay 1999; Rego and Wilson 2012). While Rego and Wilson (2012) empirically show that larger equity risk incentives motivate managers to undertake riskier tax strategies, Desai and Dharmapala's (2006) model linking equity-based compensation and aggressive tax planning shows that the relationship between the two is theoretically ambiguous. Thus, *ex ante,* it is unclear whether greater equity is positively or negatively associated with tax avoidance.

We capture corporate governance using the variables in Appel et al. (2016a): the percentage of independent directors (*Board Independence*), the use of dual class shares (*Dual Class Share*), restriction on shareholders' ability to call for special meetings (*Limit Meetings*), and the use of poison pills (*Poison Pills*). To capture broader aspects of corporate governance, we additionally include CEO being the board chair (*CEO Duality*) and the other governance features (in addition to the use of poison pills) included in Bebchuk et al.'s (2009) calculation of E-Index. These features include staggered board (*Staggered Board*), limits to shareholder bylaw amendment (*Limit Bylaw Amend*), golden parachutes (*Golden Parachutes*), supermajority requirement for mergers (*Supermajority Merger*), and supermajority requirement for charter amendments (*Supermajority Charter*). These governance variables can be grouped into two broad categories: power of the board (*Board Independence*) and CEO power (all other variables). Armstrong et al. (2015) show that the effect of board independence on tax avoidance depends on the existing tax avoidance level: a positive relation for low levels of tax avoidance and a negative relation for high levels of tax avoidance. The prediction on CEO power is also ambiguous. On the one hand, powerful CEOs can be more entrenched and less likely to yield to investor demand. On the other hand, powerful CEOs are better able to implement changes desired by investors when their incentives are aligned. Thus, the relationship between CEO power and tax avoidance is an empirical issue.

The firm disclosure and information environment variables we examine are similar to those in Boone and White (2015): the frequency of management earnings forecasts (*Management Forecasts*), the number of analysts following the firm (*Analyst Following*), the bid-ask spread (*Bid-Ask Spread*), and the probability of an informed trade (*PIN*). The effect of firm information environment on tax avoidance is theoretically ambiguous and related empirical evidence is limited

(e.g., Chen et al., 2017; Chen and Lin, 2017). On one hand, firm transparency may increase detection risk of tax avoidance activities by the public and IRS, leading to a negative relation between transparency and tax avoidance (Chen and Lin, 2017; Chen et al., 2017). On the other hand, Chen et al. (2017) argue that because tax avoidance decreases transparency (Balakrishnan et al., 2012), opaque firms may have a higher cost of tax avoidance if they are concerned about the capital market cost associated with the deterioration of the already poor information environment, resulting in a positive relation between transparency and tax avoidance.

We employ the following approach to examine the indirect mechanisms. First, for each measure of equity incentives, corporate governance, or information environment, we test whether there is a discontinuity around the index cutoff by using it as the dependent variable and estimating equation (1). For any factor to be an indirect mechanism, one necessary condition is that this factor exhibits a discontinuity around the index cutoff; that is, quasi-indexer ownership has a causal effect on the mechanism variable.[25] Second, for each variable that shows discontinuity, we add it to equation (1) and examine whether the effect of *R2000* on tax avoidance is mitigated. For this comparison, we estimate the two regressions using the same sample to ensure that the difference is due to the inclusion of the mechanism variable, not the difference in samples. If the *R2000* effect is mitigated, it suggests that quasi-indexer ownership affects effective tax rates indirectly through the variable and thus this variable captures a likely indirect mechanism.

Our approach is conceptually similar to a path analysis (Zhao et al., 2010; Baron and Kenny, 1986). A path analysis is used to answer how a variable (*X*, Russell 2000 membership in our case) affects another variable (*Y*, ETRs in our case). The impact could work through a third mediating variable (*M*, say CEO vega in our case). In a formal path analysis, a researcher estimates the

[25] We note that because we select some of the variables based on prior studies using this setting (e.g., Boone and White 2015; Appel et al. 2016), we are reproducing the prior results of discontinuity for these variables.

following two equations: $M = \alpha_1 + \beta_1 X + \varepsilon$ and $Y = \alpha_2 + \beta_2 X + \beta_3 M + \varepsilon$. $\beta_1 \times \beta_3$ being significant suggests *M* has an indirect mediation effect, while β_2 being significant suggests that *X* has a direct effect on *Y*; whether *M* attenuates or accentuate the total effect of *X* depends the signs of $\beta_1 \times \beta_3$ and β_2 (see Zhao et al. 2010). Our approach essentially requires that for *M* to be an indirect mechanism for the documented tax effect, β_1 should be significant and the joint effect of β_1 and β_3 makes β_2 smaller in magnitude compared to the case in which *M* is not included in the regression. For *X* to have a direct effect, we also require β_2 to be significant.[26] We note that KST use a similar approach to examine CEO vega and delta as potential mechanisms, though their description is very brief, as we discuss in the Introduction.

Table 7 reports the results for the analysis of discontinuity for the proposed mechanism variables. We estimate two models (N=3; bandwidth = 500, 300) for each possible mechanism variable and conclude that a variable exhibits a discontinuity around the index cutoff if the estimated coefficient on *R2000* is significant in at least one regression. Among the two equity incentive variables, *CEO Vega* is significantly larger for firms at the top of Russell 2000 than firms at the bottom of Russell 1000. All governance variables except *Limit Meetings*, *Staggered Board*, and *Supermajority Charter* exhibit a discontinuity around the index cutoff. Compared to firms at the bottom of Russell 1000, firms at the top of Russell 2000 have greater board independence, lower likelihood of using dual class shares, and their CEOs appear to be more powerful – these firms show higher likelihoods of combining the positions of CEO and board chairman, using poison pills and golden parachutes, limiting shareholder bylaw amendment, and requiring

[26] We do not perform a formal path analysis because it requires a model of structural relations among all variables, whereas when multiple mechanism variables are included in the regression (e.g., Table 9), the theoretical relations among them are unclear ex ante.

supermajority approval for mergers.[27] This latter finding of increased CEO power is also documented by Schmidt and Fahlenbrach (2017). We find a discontinuity for all of our four information environment variables: firms at the top of the Russell 2000 provide more earnings forecasts, are followed by more analysts, and have lower bid-ask spread and lower probability of informed trading. These results are consistent with Boone and White (2015).

For all variables that we identify in Table 7 to exhibit a discontinuity around the index cutoff, we examine whether each of them individually contributes to the documented tax avoidance effects in Table 3. If adding a mechanism variable to equation (1) mitigates the effect of *R2000* in at least one out of the two regressions (N=3; bandwidth = 500, 300) for an ETR measure, we conclude that the variable contributes to the effect for that ETR measure and thus is an indirect mechanism for that ETR measure. For brevity, we summarize the findings in Table 8 and present more detailed results in the Internet Appendix B. Table 8 show that the discontinuity of *Board Independence*, *Dual Class Shares*, *Poison Pills*, and *Golden Parachutes* mitigates the estimated effect of *R2000* on GAAP ETR, but not that on cash ETR; the discontinuity of *CEO Vega* and *Bid Ask Spread* mitigates the estimated effects of *R2000* on both ETR measures; and the discontinuity of *PIN* mitigates the estimated effects of *R2000* on cash ETR, but not that on GAAP ETR. These results suggest that it is more likely that the mechanism variable mitigates the effect of *R2000* on GAAP ETR than that on cash ETR. One explanation for this could be that GAAP ETR is more related to reported earnings, which are more likely to be used as a performance measure for firm

[27] Our findings for board independence and the use of dual class shares are consistent with Appel et al. (2016). They also find that higher quasi-indexer ownership leads to the removal of poison pills (i.e., change of our variable *Poison Pills*) and the reduction of restrictions on shareholders' ability to call for special meetings (i.e., change of our variable *Limit Meetings*). When using the change of *Poison Pills* as the dependent variable, our (untabulated) result is consistent with theirs. When using the change of *Limit Meetings* as the dependent variable, we find no discontinuity around the index cutoff. This difference may be due to the different research designs: we use a regular RDD analysis while they use an instrumental variable approach.

valuation and performance evaluation than cash flows (e.g., Dechow, 1994; Ball and Shivakumar, 2006; Graham et al., 2014).

Finally, we examine whether the Russell 2000 membership still has an effect on tax avoidance after accounting for the combined effects of the indirect mechanisms identified above. Specifically, we add to equation (1) all variables that we have identified as contributing to the estimated tax effects, namely, *CEO Vega*, *Board Independence*, *Poison Pills*, *Golden Parachutes*, *Bid Ask Spread*, and *PIN*, and examine whether the coefficient on *R2000* is still significant. If the coefficient is still significant, it suggests that quasi-indexer ownership has a residual effect on tax avoidance, presumably through quasi-indexers' pushing for overall firm performance.

We report results of this analysis in Table 9. Columns 1-4 report the results for GAAP ETR, and columns 5-8 for cash ETR. In column 1, to establish the benchmark, we estimate equation (1) for a bandwidth of 500 using GAAP ETR as the dependent variable and requiring non-missing indirect mechanism variables but not including those variables. For this reduced sample, the estimated coefficient on *R2000* is still negative and significant. In column 2, we add all indirect mechanism variables and find that the estimated coefficient on *R2000* is still negative but becomes insignificant. The test reported at the bottom of the table suggests that adding the indirect mechanism variables significantly weakens the effect of *R2000*. In columns 3 and 4, however, when the bandwidth is set to 300, we find that although including the indirect mechanism variables still significantly mitigates the effect of *R2000*, the *R2000* effect is still negative and significant. This result suggests that when the bandwidth is 300, quasi-indexer ownership has a residual negative effect on GAAP ETR. To assess to what extent the indirect mechanism variables contribute to the estimated effect of *R2000* on GAAP ETR, for each bandwidth we calculate the percentage decrease in the magnitude of the coefficient on *R2000* after the mechanism variables

are included and report it at the bottom of Table 9. When the bandwidth is set to 500 (300), including the indirect mechanism variables reduces the magnitude of the coefficient on *R2000* by 35.6% (17.0%). This evidence suggests that the effects of the indirect mechanisms are relatively not large.

In columns 5-8, we find that adding the indirect mechanism variables does mitigate the effect of *R2000* on cash ETR for both bandwidths. Although the coefficient of *R2000* becomes insignificant in column 6 when the bandwidth is set to 500 and the indirect mechanism variables are included, the insignificant result is likely due to the already insignificant result in column 5 (*t*-statistic = -1.49) when the sample is restricted to firm-years with non-missing mechanism variables. In column 8, when the bandwidth is set to 300, we find that the effect of *R2000* on cash ETR is still negative and significant when the indirect mechanism variables are included. When the bandwidth is set to 500 (300), including the indirect mechanism variables reduces the magnitude of the coefficient of *R2000* by 24.5% (2.8%). As in the case of GAAP ETR, the effects of these indirect mechanisms are relatively small.[28]

Overall, the results in Table 9 suggest that quasi-indexer ownership has a residual effect on tax avoidance after accounting for all identified indirect mechanisms related to CEO equity incentives, corporate governance, and information environment. When combined with the findings in Section 5, this residual effect is likely due to quasi-indexers directly pushing for improved overall firm performance. With respect to the effects of the indirect mechanism variables, we find that the effects of *CEO Vega* and *Poison Pills* are generally negative and the effect of *Bid Ask Spread* is strongly positive. The effect of *Board Independence* is mixed: it has a negative effect on

[28] When we conduct a formal path analysis by assuming no relations among the indirect mechanism variables, the percentage of the indirect effect of *R2000* relative to the total effect is 33.2%, 14.5%, 19.1%, and 0.1%, respectively. These estimates are comparable to what we document in Table 9 by comparing the estimated coefficients on *R2000*.

GAAP ETR but a positive effect on cash ETR. We caution though these results document associations not causality, and since our goal is to examine whether adding these variables mitigates the effect of *R2000*, we refrain from making strong statements regarding how these individual variables affect tax avoidance.

The indirect mechanisms identified above could work in two possible ways. First, quasi-indexers may influence the mechanism variable (e.g., CEO vega) solely to improve firm performance, and managers' stronger incentives to improve performance lead to greater tax savings (see our analyses in Section 5). Second, quasi-indexers may influence the mechanism variable for a purpose other than improving firm performance, and the change in the mechanism variable affects the cost-benefit tradeoff related to tax avoidance faced by managers, which indirectly leads to greater tax savings. To shed light on this issue, we repeat the analyses in Table 8 using pretax ROA as the dependent variable. For each identified mechanism variable, we find no evidence that including the variable mitigates the effect of *R2000* on pretax ROA (untabulated). This evidence suggests that the second case is more likely: quasi-indexers likely influence these mechanism variables for reasons beyond improving firm performance (e.g., Boone and White 2015).

7. Conclusion

We provide causal evidence on whether, and more importantly, through what mechanisms, quasi-indexers affect firms' tax planning. We employ the plausibly exogenous variation in quasi-indexer ownership as a result of the annual Russell 1000/2000 assignment to draw causal inference. A large discontinuity in Russell index weights drives a substantial cross-sectional difference around the Russell 1000/2000 cutoff in quasi-indexer ownership. Firms at the top of the Russell

2000 have significantly higher quasi-indexer ownership than those at the bottom of the Russell 1000, though they have similar market capitalizations. Using a regression discontinuity design, we find that firms at the top of the Russell 2000 exhibit significantly lower GAAP and cash ETRs. Our mechanism analysis suggests that the tax saving effects are likely due to quasi-indexers pushing for better overall firm performance, not tax savings specifically. We also show that the tax saving effects are partially due to quasi-indexers' influences on CEO equity incentives, certain aspects of corporate governance, and information environment.

Our study contributes to the literature on the impact of ownership structure on tax avoidance (e.g., Desai and Dharmapala, 2006; Chen et al., 2010). While prior studies are subject to the endogeneity concern, our empirical design improves upon these prior studies as we rely on a plausibly exogenous variation in quasi-indexer ownership to draw causal inferences. Our evidence also extends the growing literature documenting the impact of quasi-indexers on firms' corporate governance and transparency (e.g., Appel et al., 2016a; Crane et al., 2016; Boone and White, 2015) by providing evidence on these investors' impact on corporate tax planning. Our study contributes beyond the two concurrent studies on the effect of quasi-indexer ownership on tax avoidance using the same setting, Bird and Karolyi (2017) and Khan, Srinivasan, and Tan (2017), by showing the mechanisms through which quasi-indexers influence tax savings.

One limitation of our study, and of all studies employing the regression discontinuity design, is that the inference of causality is "local causality" only; that is, the causality inference can only be applied to the data points close to the left and right of the threshold – in our case, the Russell 1000/2000 threshold. Thus, we caution against making general inferences of quasi-indexers' impact on all firms' tax savings.

References

Appel, I., Gormley, T., Keim, D., 2016a. Passive investors, not passive owners. Journal of Financial Economics 121, 111-141.

Appel, I., Gormley, T., Keim, D., 2016b. Identification using Russell 1000/2000 index assignments: A discussion of methodologies. Working paper.

Armstrong, C., Blouin, J., Jagolinzer, A., Larcker, D., 2015. Corporate governance, incentives, and tax avoidance. Journal of Accounting and Economics 60, 1-17.

Badertscher, B., Katz, S., Rego, S., 2013. The separation of ownership and control and corporate tax savings. Journal of Accounting & Economics 56, 228-250.

Ball, R., Shivakumar, L., 2006. The role of accruals in asymmetrically timely gain and loss recognition. Journal of Accounting Research 44, 207-41.

Balakrishnan, K., Blouin, J., Guay, W., 2012. Does tax aggressiveness reduce corporate transparency? Working Paper. University of Pennsylvania.

Baron, R., Kenny, D., 1986. Moderator-mediator variables distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology 51, 1173-1182.

Bebchuk, L., Cohen, A., Ferrell, A., 2009 What matters in corporate governance? Review of Financial Studies 22, 783-827.

Beyer, A., Larcker, D., Tayan, B., 2014. Does the composition of a company's shareholder base really matter? Working Paper. Stanford University.

Bird, A., Karolyi, S., 2017. Governance and taxes: Evidence from regression discontinuity. The Accounting Review 92 (1), 29-50.

Boone, A., White, J., 2015. The effect of institutional ownership on firm transparency and information production. Journal of Financial Economics 117, 508-533.

Bushee, B., Noe, C., 2000. Corporate disclosure practices, institutional investors, and stock return volatility. Journal of Accounting Research 38, 171-202.

Calonico, S., Cattaneo, M., Titiunik, R., 2014. Robust nonparametric confidence intervals for regression-discontinuity designs. Econometrica 82, 2295-2326.

Calonico, S., Cattaneo, M., Titiunik, R., 2015. Optimal data-driven regression discontinuity plots. Journal of the American Statistical Association 110, 1753-1769.

Chen, S., Chen, X., Cheng, Q., Shevlin, T. 2010. Are family firms more tax aggressive than non-family firms? Journal of Financial Economics 95, 41-61.

Chen, N., Chiu P., Shevlin, T., 2017. Analyst monitoring, information asymmetry, and tax planning activities: Evidence from a natural experiment. Contemporary Accounting Research, forthcoming.

Chen, T., Lin, C., 2017. Does information asymmetry affect corporate tax aggressiveness? Journal of Financial and Quantitative Analysis, 52, 2053-2081.

Cheng, A., Huang, H., Li, Y., Stanfield, J., 2012. The effect of hedge fund activism on corporate tax avoidance? The Accounting Review 87, 1493-1526.

Chyz, J., Leung, W., Li, O., Rui, O., 2013. Labor unions and tax aggressiveness. Journal of Financial Economics 108, 675-698.

Core, J., Guay, W., 1999. The use of equity grants to manage optimal equity incentive levels. Journal of Accounting and Economics 28, 151-184.

Craig, S., May 18, 2013. The giant of shareholders, quietly stirring. The New York Time.

Crane, A., Michenaud, S., Weston, J., 2016. The effect of institutional ownership on payout policy: Evidence from index thresholds. Review of Financial Studies 29, 1377-1408.

Dechow, P., 1994. Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. Journal of Accounting and Economics 18, 3-42.

Desai, M., Dharmapala, D., 2006. Corporate tax avoidance and high-powered incentives. Journal of Financial Economics 79, 145-179.

Desai, M., Dharmapala, D., 2008. Tax and corporate governance: an economic approach. In: Schön, W. (Ed.), Tax and Corporate Governance. Springer – Verlag, Berlin, pp. 13-30.

Elhauge, E., 2016. Horizontal shareholding. Harvard Law Review 109, 1267-1317.

Graham, J., Hanlon, M., Shevlin, T., Shroff, N., 2014. Incentives for tax planning and avoidance: Evidence from the field. The Accounting Review 89, 991-1023.

Hanlon, M., Heitzman, S., 2010. A review of tax research. Journal of Accounting & Economics 50, 127-178.

Hanlon, M., Slemrod, J., 2009. What does tax aggressiveness signal? Evidence from stock price reactions to news about tax shelter involvement. Journal of Public Economics 93, 126-141.

Khan, M., Srinivasan, S., Tan, L., 2017. Institutional ownership and corporate tax avoidance: New evidence. The Accounting Review 92, 101-122.

Lee, D., Lemieux, T., 2010. Regression discontinuity design in economics. Journal of Economic Literature 48, 281-355.

Manzon, G., Plesko, G., 2002. The relation between financial and tax reporting measures of income tax. Tax Law Review 55, 175-214.

McGuire, S., Wang, D., Wilson, R., 2014. Dual class ownership and tax avoidance. The Accounting Review 89 (4), 1487-1516.

Mills, L., 1998. Book-tax differences and internal revenue service audit adjustments. Journal of Accounting Research 36 (2), 343-356.

Minnick, K., Noga, T., 2010. Do corporate governance characteristics influence tax management? Journal of Corporate Finance 16, 703-718.

Rego, S., Wilson, R., 2012. Equity risk incentives and corporate tax aggressiveness. Journal of Accounting Research 50, 775-810.

Roberts, M. R., Whited, T. M., 2013. Endogeneity in empirical corporate finance. In: Constantinides, M., Stulz, R. (eds.), Handbook of the Economics of Finance. Elsevier, vol. 2.

Robinson, J. R., Xue, Y., Zhang, M. H., 2012. Tax planning and financial expertise in the audit committee. Working Paper. University of Texas at Austin.

Schmidt, C., Fahlenbrach, R., 2017. Do exogenous changes in passive institutional ownership affect corporate governance and firm value? Journal of Financial Economics 124, 582-306.

Shackelford, D., Shevlin, T., 2001. Empirical tax research in accounting. Journal of Accounting and Economics 31 (1-3), 321-387.

Trochim, W., Donnelly, J., 2008. The research methods knowledge base. Cengage Learning: Third Edition.

Wei, W., Young, A., 2017. Selection bias or treatment effect? A re-examination of Russell 1000/2000 index reconstitution. Working paper.

Wilde, J., Wilson, R., 2017. Perspectives on corporate tax avoidance: Observations from the past decade. Working Paper. University of Iowa.

Young, A., 2018. Will the real specification please stand up? A comment on Andrew Bird and Stephen Karolyi. Econ Journal Watch 15: 35-48.

Zhao, X., Lynch, J., Chen, Q., 2010. Reconsidering Baron and Kenny: Myths and truths about mediation analysis. Journal of Consumer Research 37, 197-206.

Appendix A. Comparison with Bird and Karolyi (2017) and Khan et al. (2017)

	Our paper	Bird and Karolyi (2017)	Khan et al. (2017)
Sample period	1996-2006	1996-2006	1988-2006
Main empirical approach	RDD regression with fixed bandwidths	IV approach without using a bandwidth	RDD regression with optimal bandwidths
Tax avoidance measures	GAAP and cash ETRs	GAAP and cash ETRs; book-tax differences; tax shelter	GAAP and cash ETRs; book-tax differences; tax shelter
Quasi-indexers pushing for overall firm performance or tax savings specifically?	Argue that quasi-indexers may push for overall performance but not tax savings specifically, and provide evidence by analyzing pretax performance and voting records and guidelines	Argue that quasi-indexers pushing for tax savings specifically	Argue that quasi-indexers may push for overall performance but not tax savings specifically, but provide no direct empirical evidence
Equity incentives as an indirect mechanism?	Show that CEO vega is an indirect mechanism but vega does not fully explain the tax saving effect	No such analysis	Very briefly describe an analysis suggesting that CEO vega and delta may be indirect mechanisms without tabulating any results; state that the tax avoidance effects disappear after accounting for these indirect mechanisms
Corporate governance as an indirect mechanism?	Show that certain aspects of corporate governance are indirect mechanisms, but these mechanisms do not fully explain the tax saving effect	No such analysis	No such analysis
Information environment as an indirect mechanism	Show that information environment is an indirect mechanisms, but this mechanism does not fully explain the tax saving effect	No such analysis	No such analysis

	Our paper	Bird and Karolyi (2017)	Khan et al. (2017)
Cross-sectional analyses based on equity incentives	No such analysis	Show that the tax saving effect is smaller for firms with higher level of executive equity compensation before the index inclusion	No such analysis
Cross-sectional analyses based on corporate governance	No such analysis	Show that the tax saving effect is smaller for firms with better corporate governance before the index inclusion	No such analysis

Appendix B. Variable definitions

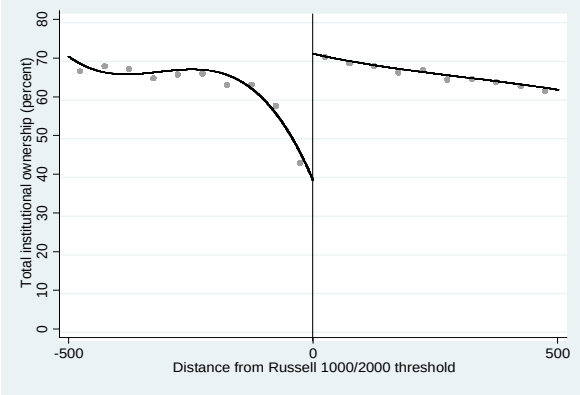
Variable	Description
Tax Planning Measures	
GAAP ETR	GAAP effective tax rate measured as the ratio of total tax expense (TXT) divided by pre-tax income (PI). GAAP ETR is set as missing if pre-tax income is missing or negative. We truncate GAAP ETR to the range [0, 1].
Cash ETR	Cash effective tax rate measured as the ratio of cash tax paid (TXPD) divided by pre-tax income (PI). *Cash ETR* is set as missing if pre-tax income is missing or negative. We truncate *Cash ETR* to the range [0, 1].
Other Variables	
Analyst Following	The number of analyst following a firm in the prior fiscal year.
Asset Turnover	Sales (SALE) divided by average assets (AT).
Bid-Ask Spread	The average of the daily bid-ask spread, computed as (ask – bid) / [(ask + bid) / 2] for the calendar quarter prior to the current fiscal year.
Board Independence	The percentage of board directors being outsiders measured in the prior fiscal year.
CEO Delta	The natural logarithm of one plus CEO's total portfolio delta, which is the dollar increase (in thousands) in the value of stocks and options a firm's CEO holds for a 1% increase in stock price.
CEO Duality	An indicator variable that equals one is a firm's CEO is also the board chairman and zero otherwise.
CEO Vega	The natural logarithm of one plus CEO's total portfolio vega, which is the dollar increase (in thousands) in the value of options a firm's CEO holds for a one percentage point increase in stock return volatility.
Dual Class Share	An indicator variable that equals one if a firm has dual class shares and zero otherwise.
Equity income	Equity income in earnings (ESUB) scaled by lagged total assets (AT).
Foreign income	Foreign income (FIFO) scaled by lagged total assets (AT).
Golden Parachutes	An indicator variable that equals to one if a firm has golden parachutes and zero otherwise.
Intangible	Intangible assets (INTANG) scaled by total assets (AT).
Leverage	Leverage measured as the ratio of long-term debt (DLTT) divided by total assets (AT).
Limit Bylaw Amend	An indicator variable that equals one if a firm has limits to shareholder bylaw amendment and zero otherwise.
Limit Meetings	An indicator variable that equals one if a firm has restrictions on shareholders' ability to call for special meetings and zero otherwise.

Market-to-book	Market-to-book ratio at the beginning of the year measured as market value of equity (PRCC_F × CSHO) scaled by lagged total assets (AT).
Management Forecasts	An indicator variable equal to 1 if a firm issues any management earnings forecasts in the prior fiscal year and 0 otherwise.
Market Cap	The natural logarithm of the end-of-May CRSP market capitalization of a firm's stock.
NOL Dummy	An indicator variable equals to 1 if a firm's net operating loss carry forward (TLCF) is positive as of the beginning of the year and 0 otherwise. This variable is set to 0 if missing.
NOL Change	The change in net operating loss carry forward (TLCF) scaled by lagged total assets (AT). This variable is set to 0 if missing.
PIN	The probability of an informed trade sourced from Stephen Brown's
Pretax ROA	Pretax income (PI) divided by average assets (AT).
Pretax Margin	Pretax income (PI) divided by total sales (SALE).
PPE	Net property, plant, and equipment (PPENT) scaled by total assets (AT).
Quasi	The percentage of shares outstanding owned by quasi-indexers defined by Bushee and Noe (2000).
R2000	An indicator variable that equals to one if a firm is in the Russell 2000 index and zero if it is in the Russell 1000 index.
Rank	The integer distance from the index cutoff each year centered at zero around the Russell 1000/2000 threshold based on the June index weight, with positive (negative) values associated with Russell 2000 (1000) firms. This variable is scaled by 1000 in all related regression analyses.
ROA	Return on assets measured as operating income (PI – XI) divided by average assets (AT).
Staggered Board	An indicator variable that equals one if a firm has a staggered board and zero otherwise.
Supermajority Charter	An indicator variable that equals one if a firm has a supermajority requirement for charter amendments and zero otherwise.
Supermajority Merger	An indicator variable that equals one if a firm has a supermajority requirement for mergers and zero otherwise.

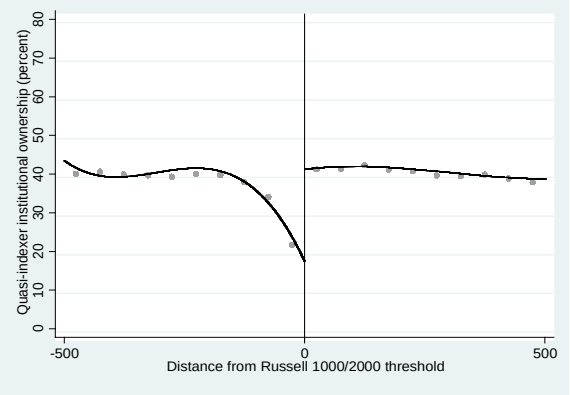
Figure 1
Institutinal ownership around the Russell 1000/2000 cutoff

These figures display fitted regression curves of ownership percentage of all and different types of institutional investors: (a) total institutional ownership, (b) quasi-indexer institutional ownership, (c) dedicated institutional ownership, and (d) transient institutional ownership, for firms around the Russell 1000/2000 cutoff for the years 1996-2006. The decomposition of institutional ownership (into quasi-indexer, dedicated, and transient) follows Bushee and Noe (2000). The ownership is measured at the end of September of each year. The distance is the relative position of a firm to the cutoff point between the Russell 1000 and the Russell 2000 each year based on June weights. Positive values represent the Russell 2000, while negative value represent the Russell 1000. The regression discontinuity plots represent local sample means using 10 non-overlapping evenly-spaced bins on each side of the cutoff following the methodology described in Calonico et al. (2015). The fitted lines represent a third-order polynomial regression curve.

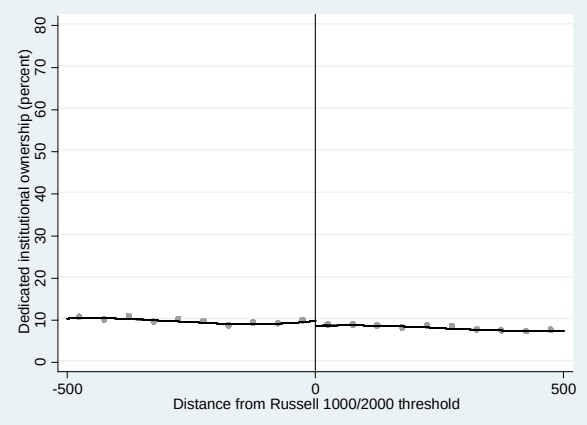
(a) Total institutional ownership



(b) Quasi-indexer institutional ownership



(c) Dedicated institutional ownership



(d) Transient institutional ownership

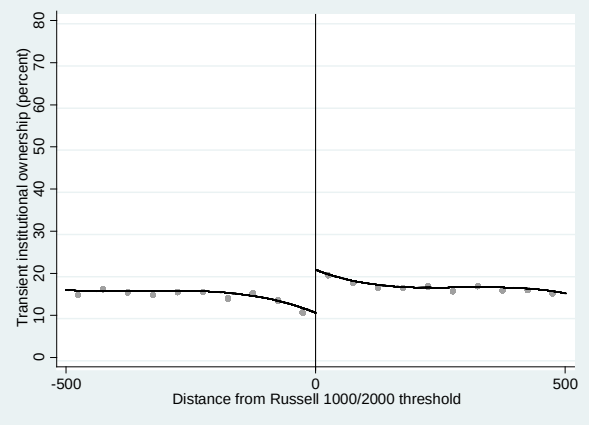
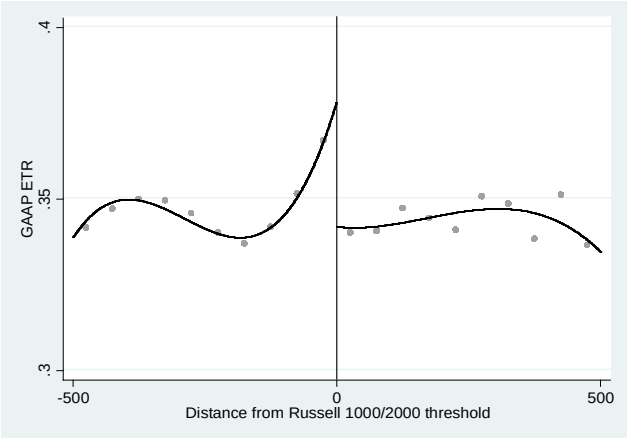


Figure 2
Tax avoidance around the Russell 1000/2000 cutoff

These figures display fitted regression curves of our two measures of tax avoidance: (a) GAAP ETR (*GAAP ETR*) and (b) cash ETR (*Cash ETR*), for firms around the Russell 1000/2000 cutoff for the years 1996-2006. Each measure of tax avoidance is defined in Appendix B. These measures are calcuated for the fiscal year with at least six months in the 12-month period after each year's index reconstiution (June 30). If a firm has two fiscal years each of which with 6 months in the 12-month period (that is, the fiscal year end is December 31), the earlier one is treated as the fiscal year after the index assignment. The distance is the relative position of a firm to the cutoff point between the Russell 1000 and the Russell 2000 each year based on June weights. Positive values represent the Russell 2000, while negative value represents the Russell 1000. The regression discontinuity plots represent local sample means using 10 non-overlapping evenly-spaced bins on each side of the threshold following the methodology described in Calonico et al. (2015). The fitted lines represent a third-order polynomial regression curve.

(a) GAAP ETR



(b) Cash ETR

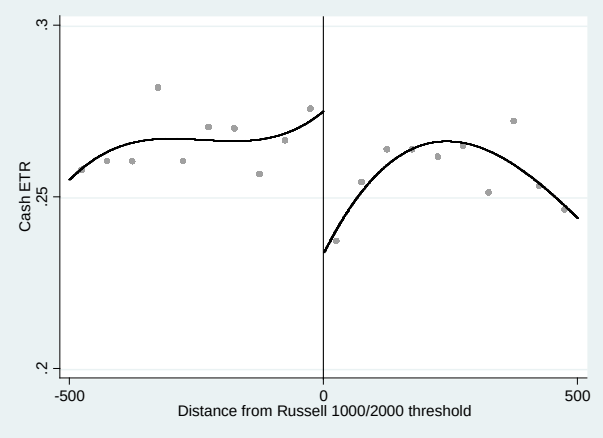


Table 1
Sample selection

		Observations
GAAP ETR sample		
Russell 1000/2000 firms for the sample period 1996-2006		33,000
Drop firms with bandwidth above 500	(22,000)	11,000
Drop REITs	(489)	10,511
Drop firms with missing GAAP ETR	(2,530)	7,967
Cash ETR sample		
Russell 1000/2000 firms for the sample period 1996-2006		33,000
Drop firms with bandwidth above 500	(22,000)	11,000
Drop REITs	(489)	10,511
Drop firms with missing cash ETR	(3,445)	7,051

This table summarizes the sample selection processes for the sample of firms with GAAP ETR and the sample of firms with cash ETR used for the RDD analysis in Table 3. The sample period is 1996 to 2006.

Table 2
Summary statistics

This table reports the summary statistics for ownership percentages of different institutional investors and our tax avoidance measures for 100 firms at the top of Russell 2000 index and 100 firms at the bottom of Russell 100 index. Variable definitions are in Appendix B. Institutional ownership is measured at the end of September of each year. Tax measures are calculated for the fiscal year with at least six months in the 12-month period after each year's index assignment (June 30). If a firm has two fiscal years each of which with 6 months in the 12-month period (that is, the fiscal year end is December 31), the earlier one is treated as the fiscal year after the index assignment. *** denotes statistical significance at the 1% level.

	Russell 2000			Russell 1000			Difference	p-Value for
	N	Mean	Std	N	Mean	Std	in Mean	Diff. in Mean
Institutional ownership (%)								
Total	921	70.5	22.6	1,010	51.0	27.8	19.5***	0.000
Quasi-indexer	921	41.7	15.3	1,010	28.6	18.5	13.2***	0.000
Dedicated	921	8.7	7.6	1,010	9.3	14.7	-0.6	0.269
Transient	921	19.8	12.4	1,010	12.7	11.3	7.1***	0.000
Tax planning measures								
GAAP ETR	772	0.340	0.109	772	0.359	0.095	-0.019***	0.000
Cash ETR	690	0.246	0.163	710	0.271	0.278	-0.026***	0.003

Table 3
Effective tax rates around the Russell 1000/2000 cutoff

This table reports the results of the RDD regressions to identify the effect of quasi-indexer institutional ownership on firms' effective tax rates. We estimate

$$Y_{it} = \alpha + \beta R2000_{it} + \sum_{n=1}^{N} \Theta_n Rank_{it}{}^n + \sum_{n=1}^{N} \lambda_n R2000_{it} \times Rank_{it}{}^n + \varepsilon_{it},$$

where Y_{it} is one of our effective tax rate measures: *GAAP ETR* and *Cash ETR*, $R2000_{it}$ is an indicator variable for firms in the Russell 2000, and $Rank_{it}$ is the integer distance from the Russell 1000/2000 index cutoff each year centered at zero around the cutoff, with positive (negative) values associated with the Russell 2000 (1000) firms. In Panel A, the RDD coefficient (β) is estimated by fitting a local third (N = 3) order polynomial estimate using a triangular kernel to the left and right of the Russell 1000/2000 cutoff using the bias-correction technology developed in Calonico et al. (2015) for two fixed bandwidths: 300 and 500. The reported numbers are estimated RDD coefficients (β) and related *z*-statistics. In Panel B, additional control variables and industry and year fixed effects are further included into the regressions. As Calonico et al.'s (2015) approach does not allow for control variables, the regressions in Panel B are estimated with a regular OLS model with standard errors clustered for each firm. The reported numbers are estimated RDD coefficients (β) and related *t*-statistics. ***, **, * indicate significance at 1%, 5%, and 10% levels, respectively. Variable definitions are in Appendix B.

Panel A: Main results

	GAAP ETR		*Cash ETR*	
	1	2	3	4
R2000	-0.037***	-0.027*	-0.048***	-0.047**
	(-3.24)	(-1.78)	(-2.90)	(-2.10)
Order of polynomial (N)	3	3	3	3
Bandwidth	500	300	500	300
No. of observations	7,967	4,737	7,051	4,226

Panel B: Sensitivity tests

	GAAP ETR		*Cash ETR*	
	1	2	3	4
R2000	-0.028***	-0.030**	-0.033*	-0.043*
	(-2.58)	(-2.04)	(-1.85)	(-1.82)
Control variables	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Order of polynomial (N)	3	3	3	3
Bandwidth	500	300	500	300
No. of observations	6,459	3,832	5,731	3,432

Table 4
Effect of quasi-indexer ownership on non-tax related firm performance

This table reports the results of the RDD regressions to identify the effect of quasi-indexer ownership on non-tax related performance metrics, including pretax ROA, pretax margin, and asset turnover. We estimate equation (1) using each of these measures as the dependent variable. The RDD coefficient (β) is estimated by fitting a local third (N = 3) order polynomial estimate using a triangular kernel to the left and right of the Russell 1000/2000 cutoff using the bias-correction technology developed in Calonico et al. (2015) for two fixed bandwidths: 300 and 500. The reported numbers are estimated RDD coefficients (β) and related *z*-statistics. The sample sizes for pretax margin and asset turnover are smaller than that for pretax ROA due to missing values for sales. ***, **, * indicate significance at 1%, 5%, and 10% levels, respectively. Variable definitions are in Appendix B.

	Pretax ROA		*Pretax Margin*		*Asset Turnover*	
	1	2	3	4	5	6
R2000	0.026*	0.042**	0.228*	0.270**	0.137**	0.205**
	(1.72)	(2.04)	(1.88)	(2.48)	(2.37)	(2.45)
Order of polynomial (N)	3	3	3	3	3	3
Bandwidth	500	300	500	300	500	300
No. of observations	9,911	5,926	8,393	4,998	8,389	4,992

Table 5
Quasi-indexer ownership and tax avoidance: subsamples based on firm performance

This table reports the results for the analysis of whether the effect of quasi-indexer ownership on tax avoidance varies with firm performance. We estimate the following equation separately for firms with high vs. low performance, measured with pretax-ROA (*Pretax ROA*) and ROA (*ROA*):

$$Y_{it} = \alpha + \beta R2000_{it} + \sum_{n=1}^{N} \Theta_n Rank_{it}{}^n + \sum_{n=1}^{N} \lambda_n R2000_{it} \times Rank_{it}{}^n + \varepsilon_{it},$$

where Y_{it} is one of our effective tax rate measures: *GAAP ETR* and *Cash ETR*. $R2000_{it}$ is an indicator variable for firms in the Russell 2000, and $Rank_{it}$ is the integer distance from the Russell 1000/2000 index cutoff each year centered at zero around the cutoff, with positive (negative) values associated with the Russell 2000 (1000) firms. To remove the effect of scale, we normalize the tax measure (to having standard deviation 1) for each subsample. We estimate the two equations for N=3 and bandwidth = 500, and report the estimated coefficient on *R2000* (β) and related *t*-statistics, as well as the *p*-value (in parenthesis) for testing the difference in β between the two subsamples. As Calonico et al.'s (2015) approach does not allow for comparing two coefficients across subsamples, we estimate the two equations using a regular OLS regression and cluster standard errors at the firm level. ***, **, * indicate significance at 1%, 5%, and 10% levels, respectively. Variable definitions are in Appendix B.

	GAAP ETR				*Cash ETR*			
	Pretax-ROA		ROA		Pretax-ROA		ROA	
	High	Low	High	Low	High	Low	High	Low
	1	2	3	4	5	6	7	8
R2000	-0.374**	-0.348**	-0.321**	-0.392**	-0.529***	-0.066	-0.525***	-0.051
	(-2.39)	(-2.09)	(-2.15)	(-2.33)	(-3.57)	(-0.45)	(-3.61)	(-0.34)
Order of polynomial (N)	3	3	3	3	3	3	3	3
Bandwidth	500	500	500	500	500	500	500	500
No. of observations	3,979	3,984	3,984	3,992	3,531	3,535	3,530	3,533
p-value for testing the difference in coefficients of *R2000*	0.9068		0.7467		0.0283		0.0266	

Table 6
Proxy voting records and guidelines for BlackRock and Vanguard

Panel A of this table summarizes the proxy voting history of BlackRock and Vanguard for the reporting period July 1, 2005 to June 30, 2006. The voting information is collected from SEC filings (Form N-PX). BlackRock files only one Form N-PX for the reporting period under the name "BlackRock Funds." Vanguard files ten Form N-PX for the reporting period, one for each fund. Panel B extracts all sentences that contain the word "tax" from BlackRock's and Vanguard's current proxy voting guidelines.

Panel A: Tax-related proposals voted on by BlackRock and Vanguard in 2006

Fund Name	No. of Firms Voted	No. of Proposals Voted	No. of Tax-Related Proposals Voted	Tax-Related Proposals Per Firm
BlackRock Funds	2,796	11,235	0	0.000
Vanguard Value Index Fund	416	1,520	3	0.007
Vanguard Total Stock Market Index Fund	3,957	10,908	14	0.004
Vanguard Small-Cap Value Index Fund	9,75	2,420	2	0.002
Vanguard Small-Cap Index Fund	1,825	4,671	4	0.002
Vanguard Small-Cap Growth Index Fund	1,048	2,760	2	0.002
Vanguard Mid-Cap Index Fund	452	1,357	2	0.004
Vanguard Large-Cap Index Fund	764	2,747	4	0.005
Vanguard 500 Index Fund	510	2,013	2	0.004
Vanguard Growth Index Fund	426	1,523	1	0.002
Vanguard Extended Market Index Fund	3,704	9,393	8	0.002

Panel B: Sentences containing "Tax" in BlackRock's and Vanguard's current proxy voting guidelines

BlackRock	"We may also support a pill where it is the only effective method for protecting tax or other economic benefits that may be associated with limiting the ownership changes of individual shareholders." (page 10, under "Poison pill plans")
	"The most common form of ESPP qualifies for favorable tax treatment under Section 423 of the Internal Revenue Code." (page 11, under "Employee stock purchase plans")
	"In evaluating golden parachute plans, BlackRock may consider several factors, including: … whether excessively large excise tax gross up payments are part of the payout…" (pages 11-12, under "Golden parachutes")
	"In order for executive compensation exceeding $1 million to qualify for federal tax deductions, the Omnibus Budget Reconciliation Act (OBRA) requires companies to link that compensation, for the company's top five executives, to disclosed performance goals and submit the plans for shareholder approval." (page 12, under "Pay-for-performance plans")
Vanguard	None

Table 7
Discontinuity of CEO equity incentives, corporate governance, and information environment around the Russell 1000/2000 cutoff

This table reports the results for the analysis of whether measures of CEO equity incentives, corporate governance, and information environment exhibit a discontinuity around the Russell 1000/2000 index cutoff. We use each measure as the dependent variable and estimate equation (1) using Calonico et al.'s (2015) approach. For each variable, we estimate two regressions (N=3; bandwidths = 500, 300), and for brevity, report only the estimated coefficient on *R2000* and its significance level for each regression. In the last column, we summarize whether there is a discontinuity for a variable. A variable is identified as exhibiting a discontinuity if the estimated coefficient on *R2000* is significant for at least one out of the two regressions. ***, **, * indicate significance at 1%, 5%, and 10% levels, respectively. Variable definitions are in Appendix B.

	Bandwidth=500		Bandwidth=300		Discontinuity?
	β	No. of obs.	β	No. of obs.	
Equity Incentive					
CEO Vega	0.888***	7,147	1.047**	4,129	Yes
CEO Delta	-0.499	7,147	0.097	4,129	No
Corporate Governance					
Board Independence	0.322***	6,180	0.334***	3,580	Yes
Dual Class Share	-0.365***	6,650	-0.305***	3,912	Yes
Limit Meetings	0.076	6,650	0.130	3,912	No
CEO Duality	0.185*	6,180	0.208*	3,580	Yes
Staggered Board	0.173	6,650	0.070	3,912	No
Poison Pills	0.714***	6,650	0.682***	3,912	Yes
Golden Parachutes	0.481***	6,650	0.469***	3,912	Yes
Limit Bylaw Amend	0.070**	6,650	0.105*	3,912	Yes
Supermajority Merger	0.119***	6,650	0.151***	3,912	Yes
Supermajority Charter	0.011	6,650	-0.004	3,912	No
Information Environment					
Management Forecasts	1.589***	10,511	2.101***	6,286	Yes
Analyst Following	3.788***	10,511	4.412***	6,286	Yes
Bid Ask Spread	-0.004***	9,883	-0.005***	5,913	Yes
PIN	-0.047***	9,893	-0.060***	5,920	Yes

Table 8
Effects of individual variables that exhibit a discontinuity around the Russell 1000/2000 cutoff

This table reports results for the analysis of whether a variable that exhibits a discontinuity around the Russell 1000/2000 index cutoff in Table 7 (denoted as *Z*) contributes to the estimated tax avoidance effects in Table 3. For each variable *Z*, we estimate the following two equations *using the same sample*:

$$Y_{it} = \alpha + \beta R2000_{it} + \sum_{n=1}^{N} \Theta_n Rank_{it}{}^n + \sum_{n=1}^{N} \lambda_n R2000_{it} \times Rank_{it}{}^n + \varepsilon_{it}, \quad \text{(a)}$$

$$Y_{it} = \alpha + \beta R2000_{it} + \gamma Z_{it} + \sum_{n=1}^{N} \Theta_n Rank_{it}{}^n + \sum_{n=1}^{N} \lambda_n R2000_{it} \times Rank_{it}{}^n + \varepsilon_{it}. \quad \text{(b)}$$

Y_{it} is one of our effective tax rate measures: *GAAP ETR* and *Cash ETR*. $R2000_{it}$ is an indicator variable for firms in the Russell 2000, and $Rank_{it}$ is the integer distance from the Russell 1000/2000 index cutoff each year centered at zero around the cutoff, with positive (negative) values associated with the Russell 2000 (1000) firms. For each mechanism variable *Z* and each ETR measure, we estimate the above two equations for N=3 and bandwidth=500, 300, and summarize whether the variable of interest is a likely indirect mechanism for the ETR measure. We conclude that a variable is a likely indirect mechanism if it mitigates the effect of *R2000* for at least one of the two bandwidths used. As Calonico et al.'s (2015) approach does not allow for control variables, we estimate the two equations using a regular OLS regression and cluster standard errors at the firm level. Variable definitions are in Appendix B. Detailed regression results are reported in the Internet Appendix B.

	Is the Variable an Indirect Mechanism?	
	GAAP ETR	*Cash ETR*
Equity incentive variable		
CEO Vega	Yes	Yes
Corporate governance variables		
Board Independence	Yes	No
Dual Class Shares	Yes	No
CEO Duality	No	No
Poison Pills	Yes	No
Golden Parachutes	Yes	No
Limit Bylaw Amend	No	No
Supermajority Merger	No	No
Information environment variables		
Management Forecasts	No	No
Analyst Following	No	No
Bid Ask Spread	Yes	Yes
PIN	No	Yes

Table 9
Aggregate effects of indirect mechanisms

This table reports the results for the analysis of whether there is a residual effect of quasi-indexer ownership on tax avoidance after incorporating all indirect mechanisms identified in Table 8. We estimate the following two equations *using the same sample*:

$$Y_{it} = \alpha + \beta R2000_{it} + \sum_{n=1}^{N} \Theta_n Rank_{it}{}^n + \sum_{n=1}^{N} \lambda_n R2000_{it} \times Rank_{it}{}^n + \varepsilon_{it}, \quad \text{(a)}$$

$$Y_{it} = \alpha + \beta R2000_{it} + \gamma Z_{it} + \sum_{n=1}^{N} \Theta_n Rank_{it}{}^n + \sum_{n=1}^{N} \lambda_n R2000_{it} \times Rank_{it}{}^n + \varepsilon_{it}, \text{ (b)}$$

where Z_{it} denotes all variables that are identified in Table 8 as indirectly contributing to the estimated tax avoidance effects in Table 3, namely, *CEO Vega*, *Board Independence*, *Dual Class Shares*, *Poison Pills*, *Golden Parachutes*, *Bid Ask Spread*, and *PIN*. Y_{it} is one of our effective tax rate measures: *GAAP ETR* and *Cash ETR*. $R2000_{it}$ is an indicator variable for firms in the Russell 2000, and $Rank_{it}$ is the integer distance from the Russell 1000/2000 index cutoff each year centered at zero around the cutoff, with positive (negative) values associated with the Russell 2000 (1000) firms. For each ETR measure, we estimate the above two equations for N=3 and bandwidth = 500, 300, and report the estimated coefficients on *R2000* (β) and the mechanism variables (γ) and related *t*-statistics, as well as the *p*-value (in parenthesis) for testing the difference in β between the two equations (a) and (b). As Calonico et al.'s (2015) approach does not allow for control variables, we estimate the two equations using a regular OLS regression and cluster standard errors at the firm level. ***, **, * indicate significance at 1%, 5%, and 10% levels, respectively. Variable definitions are in Appendix B.

	GAAP ETR				*Cash ETR*			
	1	2	3	4	5	6	7	8
R2000	-0.045*	-0.029	-0.106**	-0.088*	-0.049	-0.037	-0.144***	-0.140***
	(-1.90)	(-1.23)	(-2.11)	(-1.77)	(-1.49)	(-1.12)	(-2.71)	(-2.63)
CEO Vega		-0.002*		-0.003*		-0.005***		-0.004
		(-1.89)		(-1.88)		(-2.72)		(-1.60)
Board Independence		-0.027**		-0.030***		0.036*		0.064**
		(-2.39)		(-2.08)		(1.90)		(2.65)
Dual Class Shares		0.004		0.010		0.012		0.017
		(0.66)		(1.39)		(1.25)		(1.47)
Poison Pills		-0.010**		-0.007**		-0.016***		-0.013
		(-2.43)		(-1.30)		(-2.41)		(-1.53)
Golden Parachutes		-0.001		-0.002		-0.001		-0.002
		(-0.23)		(-0.40)		(-0.17)		(-0.21)
Bid Ask Spread		1.221***		0.953***		1.514***		1.472***
		(7.15)		(4.50)		(4.65)		(3.70)
PIN		0.055		0.001		0.144**		0.110
		(1.43)		(0.01)		(2.33)		(1.40)
Mechanism variables (*Z*)	No	Yes	No	Yes	No	Yes	No	Yes
Order of polynomial (N)	3	3	3	3	3	3	3	3
Bandwidth	500	500	300	300	500	500	300	300
No. of observations	4,341	4,341	2,457	2,457	3,915	3,915	2,230	2,230
p-value for testing the difference in coefficients of *R2000*	0.0015		0.0099		0.1070		0.7113	
Percentage reduction in the magnitude of the coefficient of *R2000*	35.6%		17.0%		24.5%		2.8%	