Identification using Russell 1000/2000 index assignments: A discussion of methodologies[*]

Ian R. Appel[†], Todd A. Gormley[‡], and Donald B. Keim[§]

October 17, 2018

Abstract

This paper discusses tradeoffs of various empirical methods that rely on Russell 1000/2000 index assignments prior to 2007 for identification and explains why the different approaches reach conflicting conclusions about the effect of index assignment on firm's ownership structure and corporate policies. The paper also discusses changes in Russell's methodology beginning in 2007 and empirical methods that can be used to continue using the setting as a source of identification for more recent years.

(*JEL* D22, G23, G30, G34, G35)

Keywords: instrumental estimation, regression discontinuity, Russell indexes

[*] This discussion originally appeared as a section in our paper "Passive Investors, Not Passive Owners," which was published in the *Journal of Financial Economics*. It was then posted as a separate 8-page working paper on SSRN in March 2016 and revised extended in October 2018. For helpful comments, we thank Alex Edmans and Wei Jiang. We also thank the Rodney L. White Center for Financial Research for financial support.

[†] Carroll School of Management, Boston College, 140 Commonwealth Avenue Chestnut Hill, MA, 02467. Phone: (617) 552-1459. Fax: (617) 552-0431. E-mail: ian.appel@bc.edu

[‡] Olin Business School, Washington University in St. Louis, One Brookings Drive, Campus Box 1133, St. Louis, MO, 63130. Phone: (314) 935-7171. E-mail: gormley@wustl.edu

[§] The Wharton School, University of Pennsylvania, 3620 Locust Walk, Suite 2400, Philadelphia, PA, 19104. Phone: (215) 898-7685. Fax: (215) 898-6200. E-mail: keim@wharton.upenn.edu

Identification using Russell 1000/2000 index assignments: A discussion of methodologies

Abstract

This paper discusses tradeoffs of various empirical methods that rely on Russell 1000/2000 index assignments prior to 2007 for identification and explains why the different approaches reach conflicting conclusions about the effect of index assignment on firm's ownership structure and corporate policies. The paper also discusses changes in Russell's methodology beginning in 2007 and empirical methods that can be used to continue using the setting as a source of identification for more recent years.

(*JEL* D22, G23, G30, G34, G35)
Keywords: instrumental estimation, regression discontinuity, Russell indexes

1. Introduction

There have been numerous papers written that use Russell index assignment as a source of exogenous variation in firms' ownership structures. The underlying idea of this identification strategy is to exploit variation in ownership that occurs around the threshold used to construct two widely-used market benchmarks— the Russell 1000 index and Russell 2000 index. Because portfolio weights assigned to each stock within these indexes are value-weighted and because the Russell 2000 index is a more popular benchmark for index funds, the assignment of an individual stock to one or the other index can have a significant impact on the extent of ownership by index funds. For example, Appel, Gormley, and Keim (2018) [AGK (2018) hereafter] show that ownership by passively managed mutual funds and exchange-traded funds (ETFs) is about 40% higher, on average, for stocks in the Russell 2000 index relative to otherwise similar stocks in the Russell 1000 index.

Despite the large number of papers written using this identification setting, there is little agreement in the literature about exactly how Russell index assignment affects firms' ownership structures and other corporate outcomes. For example, Boone and White (2015), Crane, Michenaud, and Weston (2016), and others argue there is a large 10-40 percentage point difference in total institutional ownership around the Russell 1000/2000 cutoff that is driven by a difference in *both* passive and active institutional ownership (as measured using institution-level (13F) data and Bushee (2001) "quasi-index" and "transient" classifications, respectively). Appel, Gormley and Keim (2016) [AGK (2016) hereafter], however, show there is no statistically significant difference in total institutional ownership around the cutoff, and instead show there is a significant difference in ownership by passively managed mutual funds (as shown using fund-level S12 data which allow for a more precise measure of active versus passive ownership). Moreover, in further contrast to other papers, AGK (2016) fail to detect differences in CEO compensation around the threshold or differences in firm policies related to investments, acquisitions, and capital structure.[1]

This paper discusses the different methodological choices made in the Russell 1000/2000 literature and why they reach different conclusions. The primary reason for these differences resides in the use by many papers of regression discontinuity (RD) methods to compare stock ownership at the Russell 1000/2000 threshold *after* endogenously resorting stocks within indexes using Russell's June *float-adjusted* market cap rankings. This resorting causes a severe endogeneity problem, in that these papers are comparing the least liquid, highest inside ownership stocks at the bottom of the Russell 1000 to the most liquid, lowest inside ownership stocks at the top of the Russell 2000 and incorrectly attributing observed differences to index assignment rather than to the underlying differences in liquidity and inside ownership.

[1] See footnote #6 of AGK (2016). Both Wei and Young (2018) and Glossner (2018) also highlight conflicting findings among various papers that make use of the Russell 1000/2000 threshold for identification.

This inappropriate comparison leads to faulty inferences about the impact of index assignment on firms' ownership structure and various other firm-level outcomes.[2]

Most papers that use the Russell setting for empirical identification limit their analysis to the pre-2007 period because Russell implemented a new policy for determining which stocks are assigned to the two indexes in 2007. This change in methodology eliminated the clear threshold used in earlier years and introduces additional challenges when using the Russell 1000/2000 setting for identification. In addition to evaluating the tradeoffs of the various estimation techniques used in the pre-2007 Russell 1000/2000 setting, we also address in this paper the unique difficulties of making use of the Russell setting after 2006. In particular, we discuss an IV estimation that can be used to restore conditional exogeneity in ownership across the threshold between the two indexes after 2006.

Last, we briefly discuss two new papers that also compare the various methodologies used in the Russell 1000/2000 setting—Wei and Young (2018) and Glossner (2018). A number of the points made in these papers, especially the inappropriate use of June rankings in a regression discontinuity type estimation, mirror those discussed in this paper and earlier papers (e.g., Chang, Hong, and Liskovich 2015; Mullins 2014; Appel, Gormley, and Keim 2016). That said, the papers do shed additional light on the problematic estimations used in the literature and why some findings in the existing literature are misleading. In the last section, we evaluate and compare the relative contributions of these two papers.

The remainder of the paper proceeds as follows. In Section 2, we summarize how the Russell 1000 and 2000 indexes are constructed both before and after 2006. In Section 3, we discuss the five main types of estimation techniques employed by researchers using the Russell 1000/2000 setting on data that ends before 2007 and explain why two of these are problematic. For the remaining three techniques we discuss their various tradeoffs. In Section 4, we discuss how to use the Russell 1000/2000 setting as a source of identification after Russell changed its index assignment rules in 2007. In Section 5, we discuss Wei and Young (2018) and Glossner (2018), and Section 6 concludes.

2. Russell index construction

In this section, we briefly summarize how the Russell 1000 and 2000 indexes are constructed. While the construction of these indexes is discussed in AGK (2016), AGK (2018) and other papers, it is worth including here because the details matter crucially for the empirical specification the researcher should use in the Russell 1000/2000 setting. For example, in addition to the different empirical specifications required before and after 2006, the incorrect use of the end-of-May market capitalization (used to determine index assignment for the next twelve months) and the June *float-adjusted* market

[2] Our discussion of this problem, now contained in this paper, first appeared in the working paper version of AGK (2016). We have since expanded the scope in this paper to include more recent estimation techniques used in this literature (e.g., Schmidt and Fahlenbrach 2017; Wei and Young 2018).

capitalization (used to determine a stock's ranking within the assigned index) can lead to spurious findings.

2.1. Construction of the Russell indexes, Pre-2007

Prior to 2007, the Russell 1000 included the 1,000 U.S. stocks that reflected the largest 1,000 companies in terms of market capitalization and the Russell 2000 included the next largest 2,000 stocks not included in the Russell 1000. To account for changes in stocks' market cap ranking, the Russell indexes were reconstituted each year at the end of June using a proprietary measure of stocks' total market capitalizations calculated by Russell as of the last trading day in May of that year. Specifically, a stock with an end-of-May market cap below (above) the market cap of the $1{,}000^{th}$ ($1{,}001^{st}$) largest market cap was included in the Russell 2000 (Russell 1000) index. Between the yearly end-of-June reconstitutions, index membership remained constant except for the occasional addition or removal due to delistings and IPOs.[3]

After index assignments are determined during the reconstitution, each stock's portfolio weight in the index is then calculated using a <u>second</u> measure of market capitalization. Specifically, Russell uses its June *float-adjusted* market cap to determine portfolio weights. Unlike the end-of-May Russell market cap used to determine index membership, Russell's float adjusted market cap, which is calculated using prices in June, only includes the value of shares that are available to the public. Shares held by someone with more than 10% of shares outstanding, by another member of a Russell index, by an employee stock ownership plan (ESOP), or by a government are excluded from a firm's float-adjusted market cap, as are any shares not listed on an exchange. Therefore, a stock that was the $1{,}000^{th}$ largest stock in total market capitalization need not be the stock with the smallest portfolio weight in the Russell 1000 index.

Russell weights stocks using its float-adjusted market cap to minimize the trading costs of index funds that track each benchmark. By determining portfolio weights using the June float-adjusted market cap, Russell tends to shift the less liquid, high inside ownership stocks toward the bottom of each index and the most liquid, low inside ownership stocks toward the top of each index. Failure to account for this weighting and the resulting imbalances in stock characteristics at the top and bottom of the two indexes is a common flaw of empirical methodologies that make use of the Russell 1000/2000 setting.

2.2. Construction of the indexes, 2007 and later

Beginning with the reconstitution of the Russell 1000 and 2000 indexes in June of 2007, Russell changed its methodology, so the Russell 1000 no longer comprised the 1,000 largest stocks in terms of market cap, and the Russell 2000 no longer comprised the next 2,000 largest stocks. This change in

---

[3] Stocks of companies that are delisted (e.g., because of an acquisition or bankruptcy) are removed from the index immediately. Newly-listed IPO stocks are added on a quarterly basis, and the index in which they were included depends on the market capitalization cutoff used in the most recent end-of-June reconstitution and the newly-listed stock's market capitalization at the time of its IPO. Because of these additions and deletions, the number of stocks in the two indexes can deviate from 1000/2000 during the course of the year. For more details regarding the reconstitution process and eligibility for inclusion in the Russell indexes, see Russell Investments (2013).

methodology was designed to reduce the number of stocks switching indexes each June.

Beginning in June 2007 three factors are used to determine each stock's index assignment: (1) the stock's market capitalization as of the last trading day in May of that year, (2) the stock's index assignment in the previous reconstitution year, and (3) whether the stock's end-of-May market cap falls within a certain range of the cutoff between 1,000th and 1,001st largest end-of-May market caps. Specifically, a stock with an end-of-May market cap below (above) the market cap of the 1,000th (1,001st) largest market cap will be included in the Russell 2000 (Russell 1000) index *unless* that stock was included in the Russell 1000 (Russell 2000) last year *and* its market cap is not below (above) the market cap of the 1,000th (1,001st) largest market cap by more than 2.5% of the cumulative market cap of the Russell 3000E Index (comprising the largest 4000 U.S. stocks; it contains 3561 stocks as of Aug 31, 2018).

This policy, which Russell refers to as "banding," means that stocks previously in the Russell 2000 are only moved to the Russell 1000 index during the annual reconstitution if their end-of-May Russell market cap ranking had increased significantly over the last year. For example, a Russell 2000 stock that moved from a market cap ranking of 1,050 at the end of May last year to a market cap ranking of 950 at the end of May this year will remain in the Russell 2000 even though it is no longer among the 1,000 largest stocks in terms of total market capitalization at the time of reconstitution. Given the bandwidth size (2.5% of the cumulative market cap of the Russell 3000E Index), a Russell 2000 stock typically would need to increase its Russell-calculated market cap to greater than that of the 800th largest market cap before being reassigned to the Russell 1000 at reconstitution. Likewise, a Russell 1000 stock would need to fall below an end-of-May market cap ranking of about 1,200 before being moved to the Russell 2000. This banding policy also causes the number of stocks in the two indexes to deviate from 1,000 and 2,000.

The implementation of the banding policy in 2007 poses an added challenge to researchers attempting to use the Russell 1000/2000 cutoff as a source of identification: the Russell 1000 will be over-weighted in stocks with lower-than-average stock performance during the prior year, and the Russell 2000 will be over-weighted in stocks with higher-than-average stock performance during the prior year. This is because the banding procedure constrains the best performing stocks at the top of the Russell 2000 from naturally moving to the Russell 1000, and conversely for the worst performers at the bottom of the Russell 1000. This constraint will confound comparisons of stocks at the bottom of the Russell 1000 with stocks at the top of the Russell 2000 if not properly accounted for. We discuss this challenge further in Section 3.

3. Pre-2007 specification choices when using the Russell 1000/2000 threshold for identification

In this section, we briefly summarize a selection of the specifications currently used in the literature that make use of the Russell 1000/2000 setting for identification in the pre-banding period (i.e., before the 2007 reconstitution.) These specifications range from using regression discontinuity to using an instrumental variable estimation. While the differences between the specifications are often subtle, the different implementations lead to significant differences in inferences. In discussing each specification, we describe the concerns associated with its use, and its appropriateness in the pre-banding period.

3.1. Sharp regression discontinuity using end-of-May market cap rankings

A seemingly attractive approach to using the Russell 1000/2000 setting is sharp regression discontinuity (RD) estimation. The sharp RD estimation attempts to make use of a discontinuity in index assignment between the 1,000th and 1,001st largest firms at the end of May each year to identify their effect on corporate outcomes. An advantage of this approach is the ability to focus on a subset of firms very close to cutoff, thus reducing concerns that the estimation is not adequately controlling for the one variable that determines index assignment—the end-of-May market cap calculated by Russell.

If the end-of-May market capitalization (calculated by Russell) is observable and perfectly predicts index assignment, then researchers interested in estimating the effect of Russell 2000 assignment on some outcome *Y* prior to 2007 could estimate the following sharp RD:

$$Y_{it} = \alpha + \gamma R2000_{it} + \sum_{n=1}^{N} \phi_n (Rank_{it} - 1000)^n + \varepsilon_{it} \tag{1}$$

where *Y* is the outcome of interest for firm *i* in year *t*, *Rank* is the ranking of firm *i* in year *t* in terms of end-of-May market capitalization (e.g., the 995th largest firm would have a rank of 995), and *R2000* is an indicator that equals one for firms assigned to the Russell 2000. The sample could then be restricted to firms close to the cutoff threshold of $Rank = 1000$, and the polynomial order of controls, *N*, could be varied.[4] Intuitively, the estimate γ identifies the effect of being assigned to Russell 2000 on outcome *Y* by testing for a discontinuity in *Y* between the 1,000th and 1,001st largest firms.

It is not possible to estimate equation (1), however, because the end-of-May market capitalization used by Russell to determine firms' index assignment at reconstitution is not observable to the econometrician. Specifically, Russell uses a proprietary market capitalization value that does not perfectly match up to market capitalizations reported in publicly-available databases like CRSP. Because of this, econometricians can only imperfectly predict firms' index assignments. For example, using 2006 end-of-May market caps, as calculated by CRSP, one finds that only 68% of the firms ranked between 950 and 1,000 were included in the Russell 1000; the remaining 32% were assigned to the Russell 2000 even though they were among the 1,000 largest market caps according to CRSP.

The reason why end-of-May market caps calculated from sources like CRSP do not match Russell-calculated end-of-May market caps is unclear. One possibility is that Russell uses a different source to

[4] One could also add an additional set of controls, $R2000_{it} \times \sum_{n=1}^{N} (Rank_{it} - 1000)^n$, to allow the functional form of the relation between *Rank* and outcome *Y* to vary above and below the cutoff. See Angrist and Pischke (2009), Lee and Lemieux (2010), and Roberts and Whited (2013) for more details regarding regression discontinuity estimations.

obtain the number of shares available at the end-of-May. However, papers that attempt to replicate Russell's end-of-May market caps using a combination of pricing data from CRSP and number of shares that would be publicly available to investors via Compustat quarterly filings (e.g., see Wei and Young 2018) are also unable to perfectly predict Russell's index assignments. A second possibility is that Russell intentionally adds noise to the assignment process to avoid investors being able to perfectly predict which stocks will switch indexes during the reconstitution. Given the magnitude of the stock price effects of such switches (see Chang, Hong, and Liskovich 2015), being able to predict index switches represents a profitable trading opportunity[5] and Russell may want to avoid facilitating such trades. Consistent with this possibility, Russell has been unwilling to share the historical market caps used to determine index assignments (which would be useful in reverse engineering how they calculate end-of-May market caps).[6]

3.2. Sharp regression discontinuity using June float-adjusted Russell rankings

To overcome the aforementioned challenge, some propose using the same sharp RD estimation, but instead construct *Rank* using the within-index rankings that are assigned by Russell *after* index assignments are made. In other words, the Russell 1000 stock with the smallest portfolio weight would be assigned *Rank* = 1,000, and the Russell 1000 stock with the second smallest portfolio weight would be assigned *Rank* = 999, and so on. And, the Russell 2000 stock with the largest portfolio weight in its index would be assigned *Rank* = 1,001, while the Russell 2000 stock with the second largest portfolio weight would be assigned *Rank* = 1,002, and so on. By making this change, the modified sharp RD estimation no longer tests for discontinuities in outcomes using the unobserved threshold between the 1,000th and 1,001st largest stocks based on Russell's proprietary end-of-May market cap. Instead, the estimation tests for discontinuities in outcomes between stocks at the bottom of the Russell 1000 and stocks at the top of the Russell 2000 after sorting the data based on Russell's June float-adjusted portfolio weights.

This modified sharp RD, however, is problematic. Constructing *Rank* in this way ensures that other variables will no longer be continuous at the threshold between *Rank*=1,000 and *Rank*=1,001, which violates the key identification assumption of regression discontinuity (Angrist and Pischke (2009), Lee and Lemieux (2010), Roberts and Whited (2013)). In particular, there will be a discontinuity in firms' float-adjusted market cap since Russell resorts firms within each index based on their June float-adjusted market cap *after* index assignments are made using Russell's end-of-May total market cap; firms at the bottom of the Russell 1000 will have a much smaller June float-adjusted market cap than firms at the top of the Russell 2000. This is seen in the top half of Figure 1, where we plot the average *ln*(float-adjusted market cap) as a

[5] There is a large literature on the Russell reconstitution effect; see, e.g., Madhavan (2003)

[6] In writing AGK (2016), we did obtain from Russell Investments a variable they claimed was their end-of-May market caps for the years 2002 through 2006. However, subsequent analysis revealed that these market caps also did not fully predict index assignment (e.g., see Figure 2). Mullins (2014) obtained and uses similarly noisy proprietary market cap data from Russell and provides more details regarding the likely sources for this noise.

function of this new version of *Rank*. On average, the firm with a Russell-assigned ranking of 1,000 (i.e., the bottom firm in the Russell 1000) has a Russell June float-adjusted market cap that is more than two log points smaller than the firm with a Russell-assigned ranking of 1,001 (i.e., the top firm in the Russell 2000).

This difference in Russell's June float-adjusted market cap between firms at the bottom of the Russell 1000 and the top of the Russell 2000 will cause this modified sharp RD methodology to generate spurious findings. In particular, because firms' float-adjusted market cap is directly related to liquidity, this sharp RD compares the least liquid stocks of the Russell 1000 index against the most liquid stocks of the Russell 2000, and as one might expect, these two sets of stocks will differ in other dimensions for reasons that have nothing to do with index assignment or ownership structure. Moreover, stocks with a smaller float-adjusted market cap relative to their total market cap (i.e., stocks at the bottom of the Russell 1000)), are also stocks where a larger proportion of the firm's equity is held by insiders or non-financial companies. Again, this difference will confound any comparisons using this sharp RD estimation.

The underlying differences in liquidity and inside ownership can explain why papers that use this identification strategy tend to find a much larger 10-40 percentage difference in institutional ownership (both active and passive) around the cutoff. Less liquid stocks with high inside ownership are stocks that institutions endogenously avoid, and the sharp RD is conflated by this endogenous choice.

The problems with this approach have also been restated and shown in recent papers (e.g., see Wei and Young, 2018; Glossner 2018) and were also noted in earlier papers, including Chang, Hong, and Liskovich (2015, Section 1.5), Mullins (2014, footnote #9), AGK (2016, last paragraph of Section 3.2), and AGK (2018, Section 3.2). Despite these warnings, numerous papers continue to use this flawed estimation strategy. A list of such papers can be found in footnote #12 of Wei and Young (2018) and in Table 1, Panel A of Glossner (2018) under papers listed as using "sharp regression discontinuity".

The same underlying problem also applies to papers that use the sharp RD with Russell's June float-adjusted rankings as the first stage in an IV estimation (e.g., for total institutional ownership). The instrument, however, violates the exclusion restriction since it based off of variation in ownership that is driven by endogenous differences in liquidity and inside ownership. A list of papers using this approach can be found in Table 1, Panel A of Glossner (2018) under the subcategory "IV with polynomial rankings."

The modified sharp RD estimation of Crane, Michenaud, and Weston (2016) is also problematic for similar reasons. Crane, Michenaud, and Weston use the above sharp RD as the first stage of an IV estimation, but they instead use end-of-May market caps, as calculated from CRSP, to calculate portfolio weights *after* sorting stocks into the two indexes. While this avoids using the problematic float-adjusted June portfolio weights of Russell, it still leads to a mechanical difference in end-of-May CRSP market caps at the threshold being used for identification by the RD estimation. This occurs because end-of-May CRSP market caps do not perfectly predict index assignment, and the Russell 1000 includes many stocks with end-of-May CRSP market caps that are smaller than the end-of-May market caps of some stocks in the Russell. Therefore, a sorting of stocks within an index based on end-of-May CRSP market caps creates a discontinuity in end-of-May CRSP market caps at the threshold. The jump in end-of-May CRSP market caps is shown in the bottom panel of Figure 1 where we plot the average end-of-May CRSP market caps as a function of ranks constructed using the approach of Crane, Michenaud, and Weston (2016). As can be seen, there a jump in end-of-May CRSP market caps at the threshold of nearly two log points.

Crane, Michenaud, and Weston (2016, p. 1386) acknowledge this jump in market cap at the threshold they use for identification, but argue it isn't problematic for their estimation. We disagree. If end-of-May CRSP market caps are correlated with the outcome of interest (payout policy in their case), either because total market capitalization is a determinant for the outcome of interest or correlated with other determinants, then the jump in end-of-May market caps at the threshold invalidates their RD identification strategy. Specifically, the estimates from this modified sharp RD could be driven by differences in end-of-May CRSP market cap at the threshold rather than index assignment.

3.3. Fuzzy regression discontinuity

Some have proposed switching to a fuzzy regression discontinuity to overcome problems with using the sharp RD and modified sharp RD estimations discussed in Sections 3.1 and 3.2. End-of-May market cap rankings, as calculated using CRSP, might not predict index assignment perfectly, but if they do so with enough power such that there exists a jump in the *probability* of being in the Russell 2000 when you move from an end-of-May CRSP market cap ranking of 1,000 to 1,001, then identification is still possible. In particular, fuzzy RD estimation could be achieved by estimating Equation (1) and using *Treatment* as an instrument for *R2000*, where *Treatment* is an indicator that equals one for firms with a *Rank* greater than 1,000, and *Rank* is determined using end-of-May CRSP market capitalization.

A problem with using the end-of-May CRSP market capitalization as an instrument in a fuzzy RD, however, is that they are a weak predictor of index assignment near the threshold. As noted earlier, nearly a third of the stocks ranked between 950 and 1,000 using end-of-May CRSP market caps in 2006 were assigned to the Russell 2000 index at reconstitution that year. While the predictive power of end-of-May CRSP market caps is better further from this threshold, this is not necessarily helpful in that fuzzy RD estimations that rely on a discontinuity in probability of treatment at the threshold, not at points further away from the threshold (Angrist and Pischke (2009), Lee and Lemieux (2010), Roberts and Whited (2013)). Absent such a discontinuity, the estimation can suffer from a weak instrument problem.

Even the end-of-May market caps provided by Russell (as used in Mullins (2014)) are a weak predictor of index assignment near the cutoff. As can be seen in the top panel of Figure 2, constructed using the Russell-provided end-of-May market caps,[7] having a ranking above or below 1,000 is a poor predictor of being in the Russell 2000 for firms near threshold between the 1,000th and 1,001st largest firms. In fact, firms with a ranking of 995-1,000 are equally likely to be in the Russell 2000 as firms ranked 1,001-1,005.

More recent papers, however, have argued that constructing end-of-May market caps using a combination of pricing data from CRSP and number of shares data from Compustat can yield a sufficiently strong first stage such that the fuzzy RD becomes feasible (e.g., see Wei and Young 2018; Glossner 2018). For example, using stocks with an end-of-May market cap ranking within 200 or 250 rankings of the 1,000th largest stock, Wei and Young (2018) report compelling evidence of a strong first stage that is not sensitive to the polynomial order used to control for the forcing variable, which in their specification is the end-of-May market cap ranking as calculated using a combination of data from CRSP and Compustat. Examples of this are provided in the two left panels of Figure 5 in Wei and Young (2018). These findings suggest the Russell 1000/2000 setting is amenable to using a fuzzy RD in wider bandwidths that make use of observations further from the discontinuity. While the use of wider bandwidths increases the risk of bias that might occur from any failure to control adequately for the importance of the forcing variable, *Rank*, these concerns can be overcome with sufficiently robust controls for *Rank*.

However, at smaller bandwidths, the robustness of the fuzzy RD still remains questionable even using this alternative way to calculate end-of-May market cap rankings. The potential weakness of the first stage when using smaller bandwidths is seen in the right two panels of Figure 5 in Wei and Young (2018). In both cases (especially in the bottom right panel), one can see that the probability of being in the Russell 2000 is converging to 50-50 as you near the threshold. While their inclusion of linear trends on either side give the impression of a large discontinuity, the use of a second- or third-order polynomial (as done in the left two panels of their Figure 5) would likely attenuate this difference. In fact, the paper notes on page 69 that the Kleibergen and Paap (2016) F-stat for the first stage in the 100 bandwidth is only 2.962 when introducing a third-order polynomial. Given this, it is probably not advisable to rely solely on a fuzzy RD estimation in the Russell 1000/2000 setting when using only observations very near the threshold.

Putting aside the potential need to use wider bandwidths, which also is a limitation of the AGK (2016) IV estimation discussed in the next section, the fuzzy RD suffers from two additional limitations. The first limitation of fuzzy RD is that it is not possible to use it for sample years after 2006 when Russell changes their index assignment methodology. With the implementation of banding in 2007, there is no longer a discontinuity in index assignment between the 1,000th and 1,001st end-of-May market cap rankings, regardless of how they are calculated. Instead, index assignment now depends on a combination of factors,

[7] Figure 2 is constructed using proprietary data on end-of-May market capitalizations provided to us for the years 2002 through 2006 by Russell Investments. However, as can be seen, this measure is clearly not the actual end-of-May market cap used by Russell when constructing the indexes; otherwise, it should perfectly predict index assignment.

including last year's index assignment and the distance of the stock from the 1,000/1,001 ranking cutoff. Methodologies that are designed to overcome this post-banding challenge are discussed in Section 4.

A second limitation of the fuzzy RD estimation is that it doesn't provide a direct way to study the importance of firms' ownership structure on other outcomes, which is a key motivating factor for using the Russell 1000/2000 setting for identification purposes. The first stage of the fuzzy RD estimation provides an instrument for index inclusion, not ownership structure. When one uses fuzzy RD with index inclusion as the outcome of the first stage, one is estimating the impact of index inclusion, not ownership structure, on outcomes in the second stage of the IV estimation. However, if one were interested in studying the effect of institutional ownership on other outcomes, one would instead need to use institutional ownership as the outcome variable in the first stage (e.g., AGK (2016) and AGK (2018) show passive institutional ownership is affected by index inclusion in their first stage estimation). Doing this, however, means one is no longer estimating a fuzzy RD estimation but rather estimating an IV regression where predicted index inclusion is used as the IV for passive ownership. This is very similar to the baseline estimation of AGK (2016), except that they use actual index inclusion as the IV for passive ownership, and they robustly control for *ln*(market cap) rather than market cap rankings.

3.4. IV estimation of Appel, Gormley, and Keim (2016)

The IV estimation of AGK (2016) is designed to overcome the shortcomings of the previous methodologies while still allowing the Russell 1000/2000 threshold to be used as a source of identification for passive ownership. Because index assignment prior to 2007 is determined by an arbitrary rule based on Russell's unobserved end-of-May market capitalizations, they argue that the resulting variation in passive institutional ownership can be made conditionally exogenous after robustly conditioning on end-of-May market capitalization, as calculated using other sources, like CRSP. Following that logic, AGK (2016) use an indicator for assignment to the Russell 2000 as an instrument for passive ownership while robustly controlling for the end-of-May market capitalization (as calculated using either CRSP, Compustat, or Russell's noisy measure), similar to what one might do in a RD estimation. In contrast to the empirical strategies above, however, their IV approach does not attempt to directly exploit variation in ownership or index assignment at the threshold between the Russell 1000 and 2000 indexes.

To implement the IV strategy, AGK (2016) estimate the following regression:

$$Y_{it} = \alpha + \beta Passive\%_{it} + \sum_{n=1}^{N} \theta_n \left(Ln(Mktcap_{it}) \right)^n + \gamma Ln(Float_{it}) + \delta_t + \varepsilon_{it}, \quad (2)$$

where Y_{it} is the outcome of interest for firm *i* in reconstitution year *t*; $Passive\%_{it}$ is the percent of a firm's shares held by passively managed mutual funds at the end of the first quarter of the reconstitution year *t*; $Mktcap_{it}$ is the end-of-May CRSP market capitalization of stock *i* in year *t*; and $Float_{it}$ is the June float-adjusted market capitalization calculated by Russell that determines index weights.[8] To address the possibility that passive ownership and corporate outcomes may be jointly determined, AGK (2016) instrument for $Passive\%_{it}$ using $R2000_{it}$, which is an indicator equal to one if stock i is part of the Russell 2000 index in reconstitution year t. The first stage regression takes the following form:

$$Passive\%_{it} = \eta + \lambda R2000_{it} + \sum_{n=1}^{N} \chi_n \left(Ln(Mktcap_{it}) \right)^n + \sigma Ln(Float_{it}) + \delta_t + u_{it} \cdot \tag{3}$$

The identifying assumption for the AGK (2016) framework is that after conditioning on stocks' end-of-May CRSP market capitalization, inclusion in the Russell 2000 index is associated with an increase in *Passive%* (relevance condition) but does not directly affect their outcomes of interest except through its impact on ownership by passive investors (exclusion assumption). AGK (2016) test the relevance condition in Tables 2 and 3 of their paper. They find assignment to the Russell 2000 is associated with a 1.1 percentage point increase in passive mutual fund ownership between 1998 and 2006 (p-value < 1%), an increase of about 0.5 standard deviations (Table 2, Column 2 of their paper)[9]; the effect on total mutual fund ownership is of similar magnitude, but statistically weaker (Table 2, Column 1). They also do not find a statistically significant effect of Russell 2000 assignment on total (13F) institutional ownership (Table 11, Column 1), suggesting that the Russell 1000/2000 setting is not one that can be used to study the effects of total institutional ownership.

As with any IV estimation, the validity of AGK's estimation hinges on the exclusion assumption. AGK (2016) argue the exclusion assumption is plausible in this setting as it is unclear why index inclusion would be directly related to their outcomes of interest after robustly controlling for firms' end-of-May market capitalization (again, as calculated using either CRSP, Compustat, or Russell's noisy measure they were provided) using a first-, second-, or third-order polynomial control for *ln*(*Mktcap*). Moreover, as an additional non-parametric control for size, their baseline estimation restricts their sample to the 250 stocks at the bottom of the Russell 1000 and top 250 stocks of the Russell 2000. In later estimates, they also show robustness of their IV point estimates to varying the number of stocks they include from each index from anywhere between 100 and 500 while still controlling for *ln*(*Mktcap*) using a third-order polynomial (see Appendix Figure 1 of that paper).

One concern with the baseline approach of AGK (2016) is their use of Russell rankings when selecting their sample. Because of how Russell ranks stocks within each index (i.e., putting those with the lowest float-adjusted market cap to the bottom of each index), AGK's sample selection might result in their

---

[8] The baseline specification of AGK (2016) uses end-of-May market caps, as calculated by CRSP. However, in Appendix Tables A.4 and A.5, they show that their findings are not sensitive to instead calculating end-of-May market caps either using Compustat or using the end-of-May market caps provided to them by Russell.

[9] AGK (2016) also find the effect of Russell 2000 assignment on *active* fund ownership is economically and statistically insignificant.

instrument, *R2000*, being correlated with the float-adjusted market capitalization of stocks, which as mentioned above, is related to a stock's liquidity and extent of inside ownership. This correlation, if not accounted for could result in a bias very similar to that of the estimations discussed in Section 3.2.

They mitigate this concern in two ways. First, their baseline IV specification always includes a control for firms' June float-adjusted market capitalization, as provided by Russell. In other words, they directly control for the variable that might differ across the two indexes because of how they select the sample. Second, they show that their findings are similar if they instead choose their sample using only end-of-May CRSP market cap rankings (see Section 7.3 and Appendix Table A.9). In particular, instead of selecting the 250 stocks with the smallest portfolio weights in the Russell 1000 and the 250 stocks with the largest portfolio weights in the Russell 2000, they instead rank stocks based on their end-of-May CRSP market cap and select the sample for each year using firms ranked 750th through 1,250th that year.

The sampling choice boils down to the classic tradeoff between noise and potential bias. An advantage of instead using end-of-May CRSP market caps in selecting the sample is that it eliminates the risk of estimation bias coming from Russell's June float-adjusted reweighting of stocks. A disadvantage of this alternative sampling approach is that the IV estimation is noisier. Because passive investors focus on minimizing tracking error by closely tracking the larger stocks at the top of the index, the expected difference in passive ownership occurring because of Russell 1000/2000 index assignments is largest when comparing stocks at the bottom of the R1000 to stocks at the top of the R2000; using end-of-May CRSP market caps to select the sample is going to yield a smaller and weaker first stage because it is no longer compares stocks with the largest differentials in portfolio weights. AGK (2016) discuss this idea in Section 7.3, and they note that their first stage point estimate for passive ownership falls by 25.2% when using this alternative approach (this can be seen by comparing the coefficient in Table 3, Column 2 to the reported point estimate in Section 7.3); and it falls by about 31.1% in AGK (2018) (this can be seen by comparing the coefficient in Table 3, Column 2 to that first stage point estimate of 0.83 reported in Section 6.1).

Given the inherent tradeoff of the two sampling techniques, there is not a clear preferred specification between the baseline approach in AGK (2016) and the alternative one they propose. However, AGK's finding that the alternative sampling technique has almost no impact on their point estimates suggests that the potential bias from choosing a sample based on Russell-assigned portfolio weights is negligible in their baseline specification.[10]

3.5. Identification using index switchers

---

[10] To see the similarity of the point estimates under the alternative sampling choices, compare the coefficients reported in their Appendix Table A.9 [which only used end-of-May market cap rankings to select the sample] to the corresponding estimates using their baseline estimation [which selects the sample using the actual top and bottom stocks of the Russell 2000 and 1000, respectively], as reported in Table 4 (Column 3), Table 6 (Columns 3 & 6), Table 7 (Column 3), Table 8 (Columns 3 & 6), Table 9 (Column 3), and Table 10 (Column 6).

A final empirical strategy used in the Russell 1000/2000 setting is to make use of stocks that switch indexes as a source of variation in index assignment and ownership structure. In other words, rather than make use of cross-sectional variation in index inclusion after robustly conditioning on end-of-May CRSP market caps or end-of-May CRSP market cap rankings as an exogenous source of variation in ownership structure, this estimation technique instead looks to use time-series variation in index assignment.

At a basic level, there is nothing inherently flawed about this identification strategy. So long as one robustly controls for the factors that drive index switches, one could uncover conditionally exogenous variation in index assignment. In other words, one might robustly control for changes in end-of-May CRSP market cap or changes in end-of-May CRSP market cap rankings. In essence, one is looking to identify the importance of index assignment and ownership structure by comparing (a) the change in outcomes of one stock that changes market cap rankings (market cap) by X spots (amount) and *does* switch indexes, to (b) the change in outcomes for another stock that also changes market cap rankings (market cap) by X spots (amount) but *does not* switch indexes. In fact, a version of this type of estimation can be done by simply adding stock-level fixed effects to the IV estimation of AGK (2016). By adding such fixed effects, the coefficients of their estimation would only be estimated using within-stock variation in index assignment, *R2000*, after conditioning on within-stock variation in their control variables, which includes *ln*(*Mktcap*).

In practice, however, this estimation strategy will typically be less useful. A primary reason is that the estimation will be inherently noisier because it relies on a much smaller set of firms for identification. Most stocks remain in the same Russell index from one year to the next, resulting in relatively few stocks with changes in index assignment. A second reason is that this identification strategy might not capture the relevant variation in index assignment and ownership. Specifically, this estimation relies on more transitory changes in ownership structure, which may not be the type of variation in ownership that is important for driving corporate outcomes. For example, consider a stock that drops into the Russell 2000 in one year and experiences an increase in passive ownership, but then jumps back up to the Russell 1000 the very next year and experiences a reversal in passive ownership. Should we expect that the one-year change in ownership for that stock to have a meaningful impact on persistent firm-level outcomes, like corporate governance? Instead, the relevant variation might be the sustained differences in ownership and index assignment that occurs among stocks that don't switch indexes. This type of problem is similar to the concern about fixed effects raised in McKinnish (2008) and Gormley and Matsa (2014, Section 4.2).

One must also be careful when estimating a model that makes use of switchers. An example of such a switcher estimation can be found in Schmidt and Fahlenbrach (2017). In the first stage of their IV estimation, they are careful to control for changes in end-of-May market cap rankings (see Equation (2) in their paper). Unfortunately, that control is missing from the second stage of their IV estimation (see Equation (1) in their paper). This means that their estimation is actually using this change in ranks over the last year as an instrument, <u>not</u> as a control, which they explicitly acknowledge on page 292 of their paper. This is likely inappropriate because the change in market cap may directly drive many outcomes of interest or be correlated with other factors that might in turn be correlated with changes in ownership and other outcomes of interest. In other words, one would likely be hard pressed to argue that changes in market cap rankings are a valid instrument of ownership, as this estimation strategy assumes.[11]

4. Post-2006 specification choices when using the Russell 1000/2000 threshold for identification

Russell's implementation of a banding policy beginning with the 2007 index reconstitution further complicates the use of their indexes as a source of identification. Following this change in the methodology and as discussed in Section 2.2, index assignment was no longer just a function of Russell's proprietary end-of-May market capitalization. Instead, index assignment became a function of three factors: (1) Russell's end-of-May market capitalization; (2) past index assignment; and (3) whether the firm's end-of-May Russell market capitalization falls within a certain range of the 1,000th largest firm. Because of this change, empirical methodologies used for the Russell setting prior to 2007 are unsuitable for later years. For example, this policy change eliminates any potential discontinuity in index assignment at the cutoff between the 1,000th and 1,001st largest firms based on end-of-May market capitalization (regardless of how it is constructed), precluding the use of a fuzzy RD as proposed for pre-banding time periods (Wei and Young (2018)). And the regression specification used by AGK (2016) is inappropriate because it does not control for the additional factors that contribute to index assignment after 2006.

Failure to control for these additional factors could result in misleading inferences. In particular, the implementation of banding causes other systematic differences in the type of stocks at the bottom of the Russell 1000 versus stocks at the top of the Russell 2000 in addition to differences in end-of-May market capitalization and float-adjusted market capitalization. In particular, stocks in the Russell 1000 with negative changes in end-of-May market cap rankings from year $t-1$ to year t are more likely to remain in the Russell 1000 because banding prevents them from moving down to the Russell 2000; while stocks in the Russell 2000 with positive changes in end-of-May market cap rankings are more likely to remain in the Russell 2000 because banding prevents them from moving up to the Russell 1000.

The differences in past changes in market cap rankings can be seen in Table 1 where we tabulate that average change in end-of-May CRSP market cap rankings before reconstitution (i.e., change in end-of-May CRSP market cap ranking in year t-1 to end-of-May CRSP market cap ranking in year t) for the bottom 500 stocks of the Russell 1000 and top 500 stocks of the Russell 2000 by year from 1998-2014. As

[11] Beyond the problematic exclusion of changes in market cap ranks from the second stage of their estimation, another potential concern with the specification of Schmidt and Fahlenbrach (2017) is that it does not attempt to robustly control for changes in market cap or market cap ranks by testing the robustness of the findings to using a second- or third-order polynomial set of controls like other papers in this literature.

seen in the table, we find no difference in average change in CRSP market cap rankings across the two indexes in the pre-banding years (1998-2006) but begin to see significant differences post-banding (2007-2014). In pre-banding years, the average change in end-of-May CRSP market cap rankings was -90.6 for the bottom 500 stocks of the Russell 1000 and -106.2 for the top 500 stocks of the Russell 2000, and the difference 15.6 is not statistically significant at conventional levels (p-value = 0.145). In post-banding years, however, the difference in end-of-May CRSP ranking changes across the two indexes jumps by an order of magnitude and becomes statistically significant at the 1% level. The average stock at the bottom of the Russell 1000 experienced a change in rankings of 47.3 (i.e., they became smaller relative to other stocks) while the average stock at the top of the Russell 200 experienced a change of rankings of -134.1 (i.e., they moved up in terms of their end-of-May CRSP market cap ranking).

AGK (2018) propose a modification of the methodology in AGK (2016) that allows researchers to continue using a stock's Russell index assignment as a source of exogenous variation in passive ownership in years after 2006. Specifically, they continue to use assignment to the Russell 2000 as an instrument for passive mutual fund ownership but add three additional controls to account for the new factors that determine index assignment for firm i at time t starting in 2007: 1) an indicator for having an end-of-May CRSP market capitalization that ensures firm i will be "banded" by Russell and not switch indexes in reconstitution year t because the distance between its market cap and the Russell 1000/2000 cutoff is less than 2.5% of the Russell 3000E Index cumulative market cap, $band_{it}$; (2) an indicator for being in the Russell 2000 in the last reconstitution year t–1, $R2000_{it-1}$; and (3) the interaction of these two indicators. These three additional controls capture the additional criteria used by Russell beginning in 2007 when determining each firm's index assignment at the annual end-of-June reconstitution for year t.

The authors implement this empirical strategy by estimating the following regression:

$$\begin{aligned} Y_{it} &= \alpha + \beta Passive\%_{it} + \sum_{n=1}^{N} \theta_n \ln\left(Mktcap_{it}\right)^n + \gamma \ln\left(Float_{it}\right) \\ &\quad + \mu_1 band_{it} + \mu_2 R2000_{it-1} + \mu_3\left(band_{it} \times R2000_{it-1}\right) + \delta_t + \varepsilon_{it}, \end{aligned} \tag{4}$$

where $R2000_{it}$ is again used as an instrument for $Passive\%$, and the first stage of the instrumental variable estimation is now given by

$$\begin{aligned} Passive\%_{it} &= \eta + \lambda R2000_{it} + \sum_{n=1}^{N} \chi_n \ln\left(Mktcap_{it}\right)^n + \sigma \ln\left(Float_{it}\right) \\ &\quad + \phi_1 band_{it} + \phi_2 R2000_{it-1} + \phi_3\left(band_{it} \times R2000_{it-1}\right) + \delta_t + u_{it}. \end{aligned} \tag{5}$$

5. Discussion of Wei and Young (2018) and Glossner (2018)

Similar to this paper, both Wei and Young (2018) and Glossner (2018) discuss the various estimation techniques used in the Russell 1000/2000 setting. A number of the points made in these two papers, especially the inappropriate use of June rankings in a RD type estimation, mirror those discussed above and in earlier papers (e.g., Chang, Hong, and Liskovich 2015; Mullins 2014; Appel, Gormley, and Keim 2016). The message of these earlier papers, including earlier drafts of this paper, has not been heeded as many subsequent papers have continued to use the Russell 1000/2000 setting in inappropriate ways. Moreover, readers not familiar with the Russell 1000/2000 setting continue to have difficulty determining which approach is correct, suggesting further clarification of the message will be beneficial.

While we largely agree with many points made in both Wei and Young (2018) and Glossner (2018), we have a few minor disagreements with some of their arguments. In this section, we briefly describe these areas of disagreement.

5.1. Using the Russell 1000/2000 setting for studying ownership

This first point is more a clarification than disagreement. Wei and Young (2018) give the impression that the Russell 1000/2000 setting is not useful for studying the importance of institutional ownership. It should be pointed out, though, that their paper is referring specifically to studying the effect of *total* institutional ownership. With that, we agree; the IV estimations of AGK (2016) also find no effect of index assignment on total institutional ownership. What Wei and Young (2018) don't speak to, however, is the ability to use the Russell 1000/2000 setting to study the importance of *passive* institutional ownership. In our view, the setting does provide a meaningful source of variation in passive institutional ownership, which is also shown in Glossner (2018) using the fuzzy RD estimation preferred by Wei and Young (2018). We suspect Wei and Young (2018) would not disagree with this statement.

5.2. The modified AGK IV estimation of Glossner (2018)

Glossner (2018) argues that when using the IV estimation of AGK (2016), one should not use the actual Russell rankings to select the sample but should instead use the end-of-May market cap rankings, as calculated using CRSP, to create the sample. The concern is that AGK's use of the actual rankings may introduce some imbalance in the sample because of the way Russell resorts stocks within indexes based on their June float-adjusted market caps. Specifically, the chosen sample of Russell 1000 stocks may have larger float adjustments than the chosen sample of Russell 2000 stocks, and this might cause a bias.

This is a reasonable concern, and it is also already discussed in AGK (2016) and AGK (2018). In fact, it is the reason why both AGK papers include the float-adjusted market cap as an additional control in their IV specification; e.g., see the bottom of page 120 in AGK (2016) and Section 3.4 above. Moreover, AGK propose *exactly* the same modification to the sample selection process as an alternative way to deal with this potential concern and do that test in both of their papers. This can be seen in Section 7.3 of AGK (2016) where they discuss "alternative sampling choices" and in Appendix Table A.9 where they repeat their estimations using the modified sampling approach. AGK (2018) discuss using the modified approach as a robustness check in the second paragraph of Section 5.1 and in Appendix Table 4. As noted in Section

3.4 above, however, this sampling choice has little impact on AGK's estimates, suggesting that the potential for bias with their sampling choice was negligible.

5.3. Using covariate balance tests to invalidate some estimation techniques

Wei and Young (2018) make a number of compelling arguments for why using June rankings, as done in a number of existing papers, is problematic. For example, the jump in end-of-May market cap (top-left panel of Figure 2 in their paper), the jump in float adjustments (bottom-right panel of Figure 2 in their paper), and the corresponding discussion on pages 28-29 are convincing and are directly connected to the issues we discuss above, especially the jump in float-adjusted market cap we show in Figure 1.

The paper's use of a covariate imbalance in ownership variables to invalidate these paper's use of June float-adjusted rankings, however, is less convincing. The reason for this is that if index assignment were to affect firms' ownership structures, then we might expect to observe such a covariate imbalance because index assignment is persistent from one year to the next. In other words, if this year's index assignment is correlated with last year's index assignment, and if index assignment does affect ownership, then it seems likely we would observe a covariate imbalance in these outcomes prior to reconstitution each year. This does not imply that papers that use the June float-adjusted rankings in a RD are correct, but it's not clear that covariance imbalance necessarily shows they are wrong.

Covariate imbalance tests for RD are only meaningful when looking at either predetermined outcomes or ex post outcomes that should not be affected by treatment. In our view, Wei and Young's use of covariate balance tests is flawed conceptually because it looks at outcomes argued to be affected by index assignment and the pre-reconstitution outcomes are not truly predetermined when index assignment is persistent (i.e., this year's index assignment is correlated with last year's index assignment). If there was a setting where the Russell indexes did not exist in year *t*-1, then a covariate imbalance before the creation of the indexes in year *t* would be compelling evidence against that methodology.

The authors are aware of this critique and claim on page 15 that the only thing that matters is that index assignment in year *t* is not predictive of index assignment in year *t*-1 for observations sufficiently close to the threshold (e.g., stocks with end-of-May year *t* rankings of 999 and 1,001). But, it's not clear this is correct. As the authors accurately point out on page 25, "RD designs fundamentally rely upon extrapolation, such that data *away* from the threshold must be used to provide an estimate of an effect *at* the threshold". And, the further one moves from the threshold, the more and more predictive this year's ranking will be of last year's index assignment. For example, a stock with a ranking of 1,100 in year *t* is more likely to have been in the Russell 2000 last year than a stock with a ranking of 900.

Additionally, the fact that their fuzzy RD estimation survives the covariate imbalance test when analyzing institutional ownership is not compelling evidence that the June ranking estimation is incorrect. The fuzzy RD estimation might survive simply because it accurately estimates that there is no effect of index assignment on total institutional ownership, as was also shown in Table 11, Column 1 of AGK (2016) using their IV estimation. And, because index assignment in year t-1 has no impact on institutional ownership, there will be no imbalance before reconstitution in year t. However, if index assignment does have an effect on some outcome, then it seems like there might be an imbalance present prior to reconstitution even using the fuzzy RD estimation strategy. But, it's not clear that this would invalidate the fuzzy RD for the same reason it's not clear it invalidates other approaches.

5.4. Using endogenous variables as IVs

We disagree with footnote #15 of Wei and Young (2018) which claims that because actual index assignment is an endogenous outcome of market cap rankings, it cannot be used as an instrument.

To see why we believe this statement is incorrect, consider the example of a RD estimation where the indicator for treatment is always an endogenous outcome of the forcing variable. This is exactly why RD relies on a robust set of controls for the forcing variable. The idea is that the endogenous treatment is *conditionally exogenous* after including the controls that are known to drive treatment. The same logic applies to IVs and fuzzy RD (which is a special type of IV estimation). In the fuzzy RD proposed by Wei and Young (2018) one can view the predicted index assignment as the endogenous outcome of end-of-May market rankings, and yet, fuzzy RD still uses this predicted index assignment as an instrument for actual index assignment in the first stage of the estimation. This is not problematic, however, because the fuzzy RD estimation controls for the variable driving predicted treatment, end-of-May market capitalization rankings. The same is true for AGK's use of actual index assignment as the IV: One can use actual index assignment as an instrument if one is able to include controls that sufficiently capture the underlying factors that determine treatment that might also pose problems for the exclusion assumption. The exclusion assumption does not depend on whether the unconditional IV is endogenous; it instead requires conditional exogeneity.

We also disagree with Wei and Young's claim in footnote #15 that only if the true end-of-May market cap rankings used by Russell were observed would using actual index assignment as an IV be acceptable. The logic behind this statement is unclear (at least to us). Moreover, if Russell's end-of-May market caps just measure the true end-of-May market caps (as measured by CRSP) with noise, then it would be more appropriate (to the extent that the true end-of-May market cap is thought to matter for the outcome of interest) to control for CRSP market caps instead.

6. Concluding remarks

The Russell 1000/2000 index assignments provide a powerful identification tool for helping researchers understand the relation between passive institutional ownership and outcomes related to corporate policy and governance. That said, the setting is widely misunderstood in the literature and authors have used multiple estimation methodologies, often leading to competing findings. For example, some papers argue that the setting provides large and exogenous variation in total institutional ownership, while other papers disagree and conclude that the setting does not provide exogenous variation in total ownership. Rather, the setting provides exogenous variation only in the ownership of passive mutual funds and ETFs. The root cause of these disagreements is the use of inappropriate estimation methods by some authors.

In this paper we attempt to provide clarity to these issues by discussing the various methodologies used in the Russell 1000/2000 setting along with their relative weaknesses and strengths. Our objective is to hopefully provide guidance to future researchers who wish to use the Russell 1000/2000 setting for identification.

References

Angrist, J., and J. Pischke. 2009. *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Appel, I. R., T. A. Gormley, and D. B. Keim. 2016. Passive investors, not passive owners. *Journal of Financial Economics* 121:111–41.

Appel, I. R., T. A. Gormley, and D. B. Keim. 2018. Standing on the shoulders of giants? The effect of passive investors on activism. Forthcoming in *Review of Financial Studies*.

Boone, A., and J. White. 2015. The effect of institutional ownership on firm transparency and information production. *Journal of Financial Economics* 117:508–33.

Bushee, B. 2001. Do institutional investors prefer near-term earnings over long-run value? *Contemporary Accounting Research* 18:207–46.

Chang, Y., H. Hong, and I. Liskovich. 2015. Regression discontinuity and the price effects of stock market indexing. *Review of Financial Studies* 28:212–46.

Crane, A., S. Michenaud, and J. Weston. 2014. The effect of institutional ownership on payout policy: Evidence from index thresholds. *Review of Financial Studies* 29:1377–408.

Glossner, S. 2018. The effects of institutional investors on firm outcomes: Empirical pitfalls of quasi-experiments using Russell 1000/2000 index reconstitutions. Working Paper.

Gormley, T. A., and D. A. Matsa. 2014. Common errors: How to (and not to) control for unobserved heterogeneity. *Review of Financial Studies* 27:617–61.

Madhavan, A. 2003. The Russell reconstitution effect. *Financial Analysts Journal* July/August: 51–64.

McKinnish, T. 2008. Panel data models and transitory fluctuations in the explanatory variable In *Modeling and evaluating treatment effects in econometrics*, eds. Daniel L. Millimet, Jeffrey A. Smith, and Edward J. Vytlacil, 335–58. Amsterdam: Elsevier.

Lee, D. and T. Lemieux. 2010. Regression discontinuity designs in economics. *Journal of Economic Literature* 48: 281–355.

Mullins, W. 2014. The governance impact of index funds: Evidence from a regression discontinuity. Working Paper.

Roberts, Michael R., and Toni M. Whited, 2013, Endogeneity in empirical corporate finance, in George Constantinides, Milton Harris, and Rene Stulz, eds. *Handbook of the Economics of Finance*, Vol. 2., Part A, pp. 493–572.

Russell Investments. 2018. Russell U.S. equity indexes construction and methodology, August 2018. Report. https://www.ftse.com/products/indices/russell-us

Schmidt, C., and R. Fahlenbrach. 2017. Do exogenous changes in passive institutional ownership affect corporate governance and firm value? *Journal of Financial Economics* 124:285–306.

Wei, W., and A. Young. 2018. Selection bias or treatment effect? A re-examination of Russell 1000/2000 Index reconstitution. Working Paper.

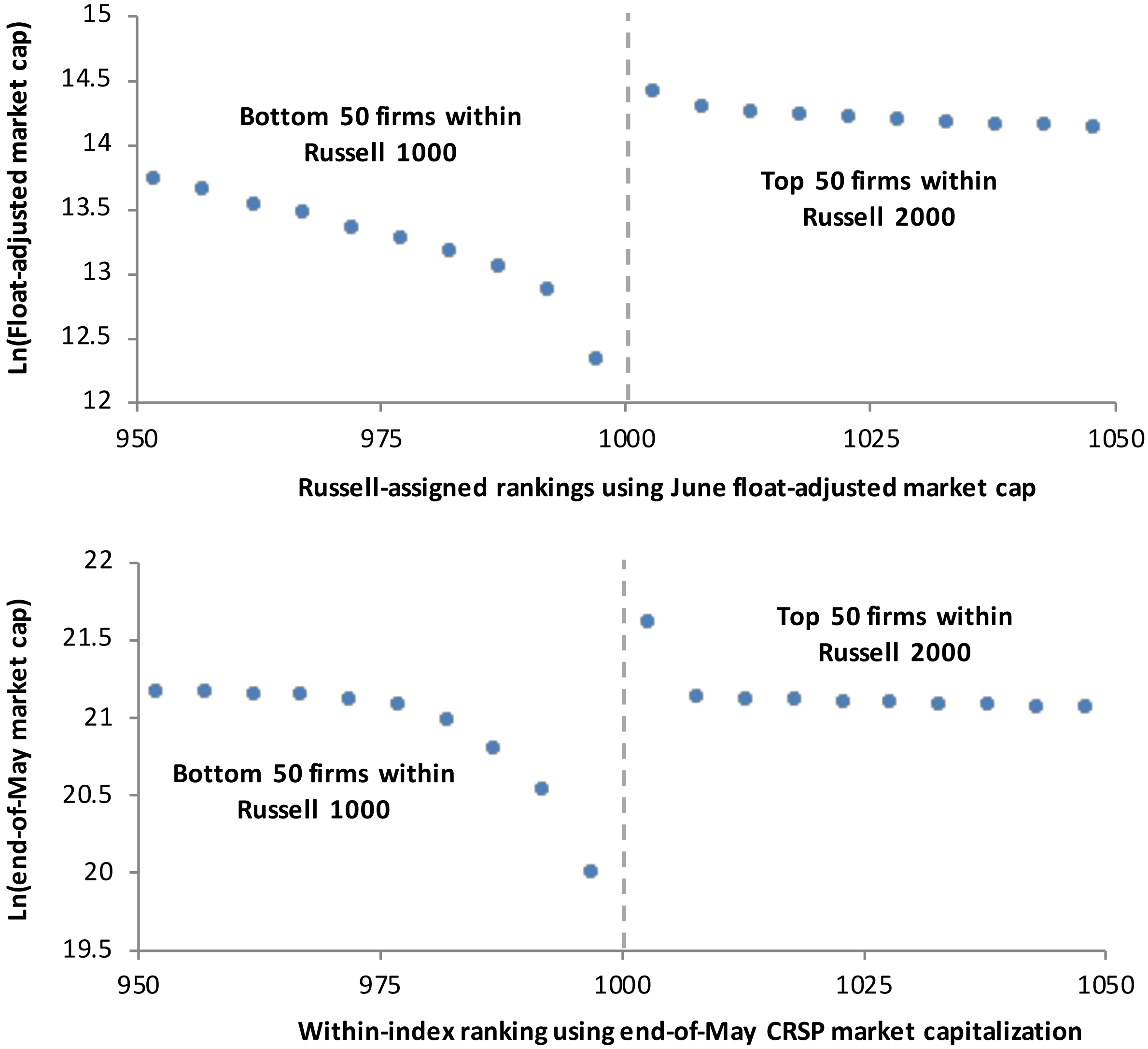

Figure 1

Average Ln(*Float*) and Ln(*Mktcap*) by ranking, where ranking is calculated using either the June float-adjusted portfolio weights assigned by Russell or within-index rankings based on end-of-May CRSP market capitalizations

The top panel of this figure plots the average Ln(float-adjusted market cap) by Russell-determined rankings for the bottom 50 firms in the Russell 1000 index and the top 50 firms in the Russell 2000 index for the years 1998-2006. A ranking of 1000 reflects the firm with the lowest portfolio weight in the Russell 1000 index, while a ranking of 1001 reflects the firm with the highest portfolio weight in the Russell 2000 index. The bottom panel of this figure plots the average Ln(end-of-May CRSP market cap) by size ranking for firms ranked between 950 and 1050, where ranking is determined using within-index end-of-May CRSP market caps. A ranking of 1000 reflects the firm with the lowest end-of-May market cap in the Russell 1000 index, while a ranking of 1001 reflects the firm with the highest end-of-May market cap in the Russell 2000 index. Averages are calculated using bins of five rankings for the years 1998-2006.

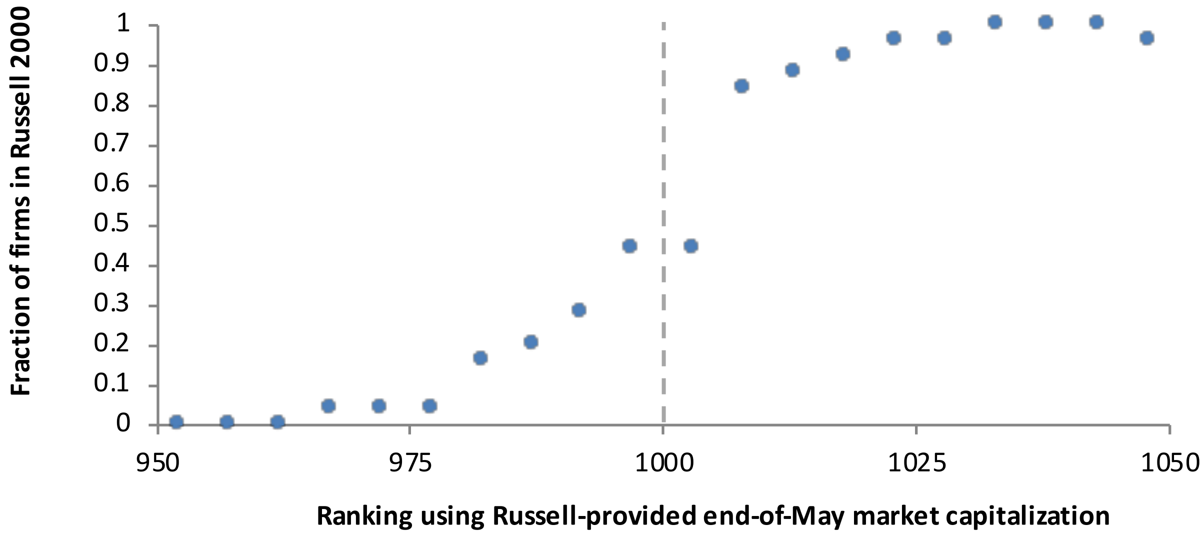

Figure 2

Probability of treatment by ranking near the Russell 1000/2000 threshold using Russell-provided end-of-May market capitalizations

This figure plots the average fraction of firm-year observations in the Russell 2000 by size ranking for the 950th to 1050th largest firms, where ranking is determined using end-of-May market capitalization numbers provided directly by Russell Investments for firms in the Russell 1000/2000 indices between 2002 and 2006. Averages are calculated using bins of five rankings and data from 2002-2006.

Table 1

Average change in end-of-May CRSP market cap ranking by sample period and index

This table reports the average change in end-of-May CRSP market cap rankings prior to reconstitution for both the pre- and post-banding periods (1998-2006 versus 2007-2014) for the bottom 250 stocks of the Russell 1000 index and top 250 stocks of the Russell 2000. Standard deviations are reported below the averages in parantheses. The change in end-of-May CRSP market cap ranking is calculated using the change in ranking from end-of-May in year *t-1* to end-of-May year *t* for each reconstitution year *t*. A ranking of 1 is assigned to the stock with the largest end-of-May CRSP market cap, a ranking of 2 is assigned to the stock with the second largest end-of-May CRSP market cap, and so on. A stock that moves from a ranking of 1,050 in year *t* to a ranking of 950 in year *t* would have a change in end-of-May market cap of -100. In other words, negative changes in rankings reflect stocks that became relatively larger in market cap over the last year. The *p*-values for the difference in means across the two indexes by sample period are also reported.

Time period =	Pre-banding period [1998-2006]	Post-banding period [2007-2014]
Bottom 250 stocks of Russell 1000	-90.6	47.3
	(300.1)	(207.6)
Top 250 stocks of Russell 2000	-106.2	-134.1
	(345.4)	(260.3)
Difference =	15.6	181.4
Difference *p* -value =	0.145	0.000