Andrew Yeh
Taylor
Acct 930 Capstone Project: Boone and White (2015)
April 22, 2022

Abstract

Boone and White (2015) state that they investigate the effect of institutional ownership on firm transparency and information production. They do this by using the discontinuity in institutional ownership that results from just missing the cutoff for entering the Russell 1000 instead of the Russell 2000 versus just making the cutoff. They find that firms in a bandwidth around this discontinuity have greater managerial disclosure, analyst following, and a more transparent trading environment. I investigate the third claim specifically by replicating the author's results and then perturbing various dimensions of the research design and variable construction to better understand the sources of statistical significance. I find that perturbations to the construction of the liquidity measures yield substantially similar results, but adjustments to the timing of the ranking (the running variable) and subsetting the proposed bandwidth lead to insignificant or even statistically significant but opposite-signed results.

1. Introduction

Boone and White (2015)'s work with the exogenous shock of index inclusion led to many papers using the same shock afterwards, though he cites working papers from Crane, Michenaud, and Weston (2014) and Lu (2013) that preceded this paper's publication. Crane, Michenaud, and Weston (2014) similarly links index inclusion to institutional ownership and shows an increase in dividend payouts, an effect also found in Table 2 of Boone and White (2015) though the effect is insignificant (z-score of 1.48). Lu (2013) links institutional ownership through index membership to banking practices and loan spreads. Cao, Gustafson, and Velthuis (2014) investigate the discontinuity between joining the Russell 2000 and not being able to join the index at all and claim a higher beta, higher short interest, higher liquidity, and more equity-focused versus debt-focused capital structures.

Since Boone and White's (2015) publication, a flurry of subsequent papers have used the same exogeneous shocks in various applications.

One strand of the literature uses this discontinuity to link institutional ownership to monitoring costs and agency theory. Kang, Juan, and Hyun (2018) use index inclusion as an instrument for institutional ownership's effect on monitoring management though they do not mention the differential monitoring effect of quasi-indexers compared to other institutional investors. Scmidt and Fahlenbrach (2017) claim that exogenous increases in passive investment leads to less monitoring and a more powerful CEO who appoints fewer independent directors and makes worse merger decisions. In this way, their claims are in the opposite direction of Boone and White's who claim that passive investors nevertheless engage in monitoring in order to increase liquidity of the firm and decrease own costs.

Several papers connect institutional ownership to tax planning. Chen, Huang, Li, and Shevlin (2018) show that index membership discontinuity leads to better corporate tax planning and link this outcome to quasi-indexer's influence on executive compensation and corporate governance. Khan, Srinivasan, and Tan (2016) similarly show that index membership discontinuity leads to increased tax avoidance policies, though they disagree with Chen, Huang, Li, and Shevlin (2018) on the mechanism of the reduced taxes, providing evidence that it is the result of tax shelters in order to meet analyst expectations.

Many papers use the discontinuity to demonstrate that institutional ownership is a societal positive. Lin, Mao, and Wang (2018) find that the increase in institutional ownership from the index membership discontinuity are more likely to issue managerial forecasts and explains that this is evidence of "peer pressure" for increased voluntary disclosure. Chen, Dong, and Lin (2019) link institutional ownership to increases in CSR performance. Azar, Duro, Kadach, and Ormazabal (2020) come to the same conclusion that passive institutional investors—in particular, the "Big 3"—cause reductions of corporate carbon emissions in companies they're invested in. Appel, Gormley, and Keim (2016) find that passive institutions influence firm's governance choices to increase the number of independent board directors, remove takeover defenses, and equalize voting rights across different classes of shares, and that these together help explain positive longer-term performance.

2. Replication Process

During the sample period, the authors state that they use 6,472 unique firms across the Russell 1000/2000 indices. I find that there are 7,775 but post-CRSP merge, there are 7,325. The reason they authors find fewer companies may be due to merges with other datasets, e.g. Compustat, First Company Issued Guidance, and I/E/B/S.

2.1 Replicating Inferences

I focus my replication efforts on the link between the exogenous shock of index membership and measures of the "trading environment." Specifically, the authors find that increased institutional ownership leads to:

1. Lower bid-ask spread: bid-ask spread divided by the mean of the bid and ask prices
2. Higher turnover: volume divided by shares outstanding
3. Higher dollar volume: volume times share price

The two main tests are a difference-in-means test across bandwidths of 50, 100, and 200 along the discontinuity as well as a RDD with substantially similar bandwidths. The authors use a "local third-order polynomial estimate using a triangular kernel to the left and right…using the bias-correction methodology in Calonico, Cattaneo, and Titiunik (2014)." I replicate his results using third-order polynomials to the left and right of the discontinuity but without the triangular kernel or bias-correction methodology. Similarly to the author, I find highly significant difference-in-means for all bandwidths as well as highly significant RDD coefficients. See Table 1 for the replication results.

2.2 Perturbations to the Independent Variables of Interest

The authors use three separate measures for liquidity and the trading environment which are fairly comprehensive as their independent variables. To investigate the specific constructs, I perturb their variable measurement formulas as follows:

1. Bid-ask spread: The authors calculate a bid-ask spread that varies over time. I perturb this by calculating the 11-month average.
2. Turnover: The authors calculate an 11 month average of the turnover. I perturb this by allowing turnover to vary over time.
3. Dollar volume: The authors calculate an 11 month average of the turnover. I perturb this by allowing turnover to vary over time.

Put simply, my perturbations to the independent variables were to use the calculations methods they propose on the opposite variables that they used. I find substantially the same results from these perturbations. See Table 2 for a comparison of t-statistics.

2.3 Perturbations to the Running Variable of Interest

The Russell Index reconstitution consists for several dates and stages that runs from the last trading day of May to the end of June.[1] Most importantly, FTSE uses the market capitalizations of firms in the last trading day of May to determine membership into the Russell 1000 and Russell 2000 but uses market capitalization of firms at the end of June to determine the weighting of firms in the index. Both the ranking of firm capitalization at the end of May and end of June, therefore, are possible methods to calculate the running variable up to and beyond the discontinuity. The May-version is closer in principle to the discontinuity of index membership itself. However, the June-version is closer to the effect on institutional ownership by quasi-indexers.

The authors use the June-version of market capitalization ranking to determine the discontinuity and bandwidths. I experiment with using the May-version instead. One confounding factor is that I was unable to find matches in CRSP market capitalization for all 7,775 companies with positive weights in the Russell 1000/2000. I find matches for 7,325 instead. If the ~6% of companies that are missing data are correlated in some way, use of the May ranking will bias results.

See Table 3 for a comparison of t-statistics.

2.4 Perturbations to the RDD Bandwidth

One of the weaknesses of RDD is that they trade off more internal validity with the natural experiment around the discontinuity for less external validity since findings may not apply to firms far away from the cutoff. The authors substantially address this by considering large bandwidths of up to 200 firms: the bandwidths of 200 firms cover 13% of the total firms in either Russell index compared to just 1.7% for the bandwidth of 50 firms, giving the reader more confidence in the external validity of this research design.

However, a visual inspection of Figure 5 or a close inspection of diminishing results in Table 5 suggest that the statistical significance of results rests substantially on the shoulders of only the 1.7% of firms that are very close to the discontinuity of membership. I test this by examining firms that are within the closest 200 firms but outside the closest 50 firms ranked closest to the breakpoint. This is reasonable for two reasons: firstly, this is a strict subset of the bandwidth proposed by the authors themselves of up to 200 firms. Second, the story of the discontinuity increasing institutional ownership holds in the 50-200 distance range as can be verified through a visual inspection of Figure 3 in Boone and White (2015). The difference in institutional ownership is especially prominent in the 50-100 distance range and diminished but still present from 100-200 distance range. Therefore, if differences in institutional ownership cause a more transparent and liquid trading environment, the story should continue to hold in the more distant subset of the proposed bandwidths.

I find that the authors' results often on the trading environment are either not statistically significant or even reverse direction for companies inside the bandwidth but outside the closest 50 firms. This is potentially a cause for concern for inferences linking institutional ownership to the trading environment since firms in the bandwidth outside the closest 50 show clear differences in institutional holding from

[1] See Figure 1 in Boone and White (2015) for an example timeline of the process.

differential index membership without any of the accompanying gains in trading environment liquidity. To control for size as a confounding factor, I also attempt a variant of these results with market capitalization as a control in the RDD and find the same results.

See Table 4 for a comparison of t-statistics.

3. Tables

Green-highlighted entries are statistics that agree with the authors' results. Yellow-highlighted entries are non-statistically significant in contrast to the author's findings and predictions. Red-highlighted entries are statistically significant but in the opposite direction of the author's findings and predictions.

Table 1
t-statistics from Replication Results

Construct	Coefficient [1]	t-statistic
Difference in Means		
Bid-ask spread	-0.018	-8.76
Turnover	3.09	13.5
Dollar volume	11,369,349	22.99
Regression Discontinuity Design		
Bid-ask spread	-0.009	-9.18
Turnover	9.96	10.05
Dollar volume	50,998,650	24.5

[1] The coefficient is calculated as the difference-in-means between the most heavily weighted firms in the Russell 2000 and the most lightly weighted firms in the Russell 1000 for the difference-in-means test, and the coefficient for the "Russell 2000" dummy variable in the regression discontinuity design with triple polynomials on either side of the discontinuity.

Table 2

t-statistics from Perturbations in the Definition of Independent Variables Measuring Liquiditiy

		New	
Construct	**Original**	**Alt 1**	**Alt 2**
Difference in Means			
Bid-ask spread [1]	-8.76	-13.52	-22.58
Turnover [2]	13.5	10.14	--
Dollar volume [2]	22.99	20.02	--
Regression Discontinuity Design			
Bid-ask spread [1]	-9.18	-7.62	-14.86
Turnover [2]	10.05	7.34	--
Dollar volume [2]	24.5	19.5	--

Note: Green-highlighted entries are statistics that agree with the authors' results. Yellow-highlighted entries are non-statistically significant in contrast to the author's findings and predictions. Red-highlighted entries are statistically significant but in the opposite direction of the author's findings and predictions.

[1] Bid-ask spread is calculated as the concurrent bid-ask spread. Alternatively, the 11-month average is calculated. Additionally, the bid-ask spread in cents without price normalization is calculated.

[2] Turnover and dollar volume are calculated originally as the 11-month average of turnover and dollar volume. Alternatively, they are made concurrent.

Table 3

t-statistics from Perturbations in the Definition of the Running Variables Measuring Size

Construct	Original June Definition [1]	New May Definition [2]
Difference in Means [-50,50]		
Bid-ask spread	-8.76	-0.11
Turnover	13.5	0.47
Dollar volume	22.99	2.22
Regression Discontinuity Design		
Bid-ask spread	-9.18	0.7
Turnover	10.05	1.55
Dollar volume	24.5	17.89

Note: Green-highlighted entries are statistics that agree with the authors' results. Yellow-highlighted entries are non-statistically significant in contrast to the author's findings and predictions. Red-highlighted entries are statistically significant but in the opposite direction of the author's findings and predictions.

[1] The June Definition uses the end-of-June market capitalizations to rank securities with respect to the Russell 1000/2000 index inclusion breakpoint. This corresponds to the final index weights in the Russell indices.

[2] The May Definition uses the end-of-May market capitalizations to rank securities with respect to the Russell 1000/2000 index inclusion breakpoint. This corresponds to the inclusion decision into either the Russell 1000/2000 indices.

Table 4

t-statistics from Perturbations in the Bandwidths Around the Discontinuity [1]

	Original [2]		New [3]	
Construct	**0-50**	**0-100**	**50-100**	**100-200**
Difference in Means				
Bid-ask spread	-8.76	-10.39	-9.32	-4.77
Turnover	13.5	10.43	0.02	-0.7
Dollar volume	22.99	12.17	-7.7	-25.33
Regression Discontinuity Design				
Bid-ask spread	-9.18	-14.12	-1.52	0.04
Turnover	10.04	13.14	0.06	-3.48
Dollar volume	24.49	29.1	0.21	-1.36
Regression Discontinuity Design with Market Capitalization Control				
Bid-ask spread	-7.16	-11.89	-1.64	-0.13
Turnover	10.24	13.62	0.02	-3.44
Dollar volume	16.06	20.25	0.48	-1.09

Note: Green-highlighted entries are statistics that agree with the authors' results. Yellow-highlighted entries are non-statistically significant in contrast to the author's findings and predictions. Red-highlighted entries are statistically significant but in the opposite direction of the author's findings and predictions.

[1] Bandwidth lengths are the sub-sample of the data used in difference-of-means and regression discontinuity design tests with 0 marking the point of discontinuity between inclusion in the Russell 1000 and Russell 2000 indices, -1 being the largest firm by index holding percentage in the Russell 2000, and 1 being the smallest irm by index holding percentage in the Russell 1000.

Bibliography

Appel, Ian R., Todd A. Gormley, and Donald B. Keim. "Passive Investors, Not Passive Owners." *Journal of Financial Economics* 121, no. 1 (July 1, 2016): 111–41. https://doi.org/10.1016/j.jfineco.2016.03.003.

Azar, José, Miguel Duro, Igor Kadach, and Gaizka Ormazabal. "The Big Three and Corporate Carbon Emissions around the World." *Journal of Financial Economics* 142, no. 2 (November 1, 2021): 674–96. https://doi.org/10.1016/j.jfineco.2021.05.007.

Boone, Audra L., and Joshua T. White. "The Effect of Institutional Ownership on Firm Transparency and Information Production." *Journal of Financial Economics* 117, no. 3 (September 1, 2015): 508–33. https://doi.org/10.1016/j.jfineco.2015.05.008.

Cao, Charles, Matthew Gustafson, and Raisa Velthuis. "Capital Market Consequences of Index Membership: Evidence from Small Firms," n.d., 59.

Chen, Shuping, Ying Huang, Ningzhong Li, and Terry Shevlin. "How Does Quasi-Indexer Ownership Affect Corporate Tax Planning?" *Journal of Accounting and Economics* 67, no. 2 (April 1, 2019): 278–96. https://doi.org/10.1016/j.jacceco.2018.01.001.

Chen, Tao, Hui Dong, and Chen Lin. "Institutional Shareholders and Corporate Social Responsibility." *Journal of Financial Economics* 135, no. 2 (February 1, 2020): 483–504. https://doi.org/10.1016/j.jfineco.2019.06.007.

Crane, Alan D., Sébastien Michenaud, and James P. Weston. "The Effect of Institutional Ownership on Payout Policy: Evidence from Index Thresholds." *The Review of Financial Studies* 29, no. 6 (June 1, 2016): 1377–1408. https://doi.org/10.1093/rfs/hhw012.

Kang, Jun-Koo, Juan Luo, and Hyun Seung Na. "Are Institutional Investors with Multiple Blockholdings Effective Monitors?" *Journal of Financial Economics* 128, no. 3 (June 1, 2018): 576–602. https://doi.org/10.1016/j.jfineco.2018.03.005.

Khan, Mozaffar, Suraj Srinivasan, and Liang Tan. "Institutional Ownership and Corporate Tax Avoidance: New Evidence." *The Accounting Review* 92, no. 2 (July 1, 2016): 101–22. https://doi.org/10.2308/accr-51529.

Lin, Yupeng, Ying Mao, and Zheng Wang. "Institutional Ownership, Peer Pressure, and Voluntary Disclosures." *The Accounting Review* 93, no. 4 (October 1, 2017): 283–308. https://doi.org/10.2308/accr-51945.

Lu, R., 2013. How does institutional ownership affect bank loan pricing: evidence from a regression discontinuity design. Unpublished Working Paper. National University of Singapore, Singapore.

Schmidt, Cornelius, and Rüdiger Fahlenbrach. "Do Exogenous Changes in Passive Institutional Ownership Affect Corporate Governance and Firm Value?" *Journal of Financial Economics* 124, no. 2 (May 1, 2017): 285–306. https://doi.org/10.1016/j.jfineco.2017.01.005.