Online Appendix for "Do Index Funds Monitor?"

This Online Appendix provides additional empirical evidence to supplement the analyses provided in the published text.

Table of Contents

- Section A: Analysis of fund voting.
- Section B: Analysis of fund exit.
- Section C: Analysis of fund voting using a Heckman (1979) correction model to account for potential selection bias.
- Section D: Balance tests and pretrend comparisons for our research design, to examine the possibility of bias due to preexisting differences between treated and control firms.
- Section E: Discussion of Russell research designs.

A. Voting – Additional Analyses

In this section we examine fund voting using several alternate specifications. All of our conclusions remain unchanged.

A.1. Voting by Passive versus Active Fund Families

Many mutual funds belong to a fund family, and in that case fund voting could be coordinated at the family level. In this section, we investigate fund voting when index versus active status is measured at the level of the fund family rather than at the level of the individual fund (as we do in the main text). The results are in Table A1 Columns 1 and 2. The independent variable for this analysis is $FracPassive_{it}$, which measures the fraction of fund i's family's total AUM that was in passively-managed index funds in year t. We find that the pattern in fund voting is consistent with our main estimates and indeed stronger: Based on the regression estimates, a fund belonging to a family that was 100% passive was 27% more likely to vote with firm management on a contentious agenda item compared to a fund belonging to a family that was 100% active.

In Table A1 Columns 3 and 4, we examine the voting behavior of the Big Three fund families (Vanguard, Blackrock, and State Street) compared to the voting record of other large fund families. The sample consists of all votes on contentious agenda items by funds that belonged to fund families with at least $100 billion in AUM that year (the results are similar if we drop this requirement). We find that the voting behavior of the Big Three is consistent with their high passive fraction of assets under management: Compared to other large fund families, the Big Three are 25 percentage points more likely to vote with firm management on contentious items.*

Finally, consistent with our results in the paper we continue to observe that the comparisons are very similar with firm and year fixed effects (Columns 1 and 3) versus firm-by-year fixed effects (Columns 2 and 4). We draw two conclusions from this observation. First, firm and year fixed effects control for most of the variation in (for example) policy or governance; that is, policy or governance vary considerably among firms and perhaps in the aggregate over time, and much less within a given firm over time. Second, the stability of the coefficients of interest across a variety of specifications suggests that our results are not likely to be affected by an omitted variable (Oster, 2019).

*One concern is that securities lending by index funds could lead to a large difference between the number of shares held and the number of shares voted. To address this point, we located a recent Morningstar report (https://www.morningstar.com/lp/securities-lending-risks-rewards). For the Big Three passive fund families the yearly percentage of their portfolio that was on loan averaged less than 5% in all cases. Thus, securities lending is not a concern for the interpretation of our results.

Table A1
Voting Comparisons at the Fund-Family Level

The table presents comparisons of fund voting on contentious items between fund families. $FractionPassive_{it}$ is the fraction of fund i's family's total AUM that was passively managed in year t. $Big3_{it}$ is an indicator variable that equals 1 if the fund is a member of the Big 3 passive fund families (Vanguard, Blackrock, and State Street). The sample consists of votes on contentious items (*i.e.* those on which ISS and firm management were opposed). In columns 3 and 4 the sample consists of contentious votes by funds within large families (those that had at least $100 billion in AUM that year). Robust standard errors clustered by fund-family and firm are in parentheses. *, **, *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)
	$VotedwithMgmt$	$VotedwithMgmt$	$VotedwithMgmt$	$VotedwithMgmt$
$FractionPassive_{it}$	0.271***	0.273***		
	(0.094)	(0.093)		
$Big3_{it}$			0.254***	0.256***
			(0.051)	(0.048)
Observations	2,530,135	2,529,929	1,149,311	1,149,149
Adjusted R^2	0.085	0.130	0.222	0.320
Firm FE	Yes	No	Yes	No
Year FE	Yes	No	Yes	No
Firm × Year FE	No	Yes	No	Yes

A.2. Voting Comparisons with Agenda-Item Fixed Effects

The comparisons of fund voting presented in the paper use either firm- and year fixed effects, which sweep out differences in governance or policy between firms and changes in overall governance or policy over time, or firm-by-year fixed effects which absorb even time-varying heterogeneity in governance or policy at the individual firm level. That is, the latter set of results compares fund voting between funds within each individual firm-year.

There could be subtler heterogeneity in the types or the contentiousness of the items that come up for votes at a given annual meeting. For example, if (along with the rise of index investing) fewer problematic management proposals are being tabled, or the proposals that are tabled are less problematic, this could bias the comparison toward finding that index funds vote more with firm management.

We examine this possibility by re-examining our voting comparisons with individual fixed effects for each agenda item. That is, these estimates sweep out all variation between individual agenda items, and are identified only by comparing how funds vote *within* each individual agenda item. Results are reported in Table A2. We find that results using these alternative specifications are nearly identical to our main estimates in the paper. We conclude that changes in the contentiousness or the types of agenda items at firm's annual meetings are not a factor in our comparisons of fund voting.

Table A2
Voting Comparisons within Agenda Items

The table presents comparisons of fund voting with a fixed effect for each individual item. The sample consists of votes on contentious items (*i.e.* those on which ISS and firm management were opposed). Robust standard errors clustered by fund and firm are in parentheses. *, **, *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
	$VotedwithMgmt$	$VotedwithMgmt$	$VotedwithMgmt$	$VotedwithMgmt$	$VotedwithMgmt$
Item Type:	All	Board	Compensation	Disclosure	Entrenchment
$IndexFund_i$	0.101***	0.110***	0.113***	0.063**	0.048***
	(0.024)	(0.027)	(0.031)	(0.032)	(0.017)
Observations	2,600,136	1,426,875	35,129	122,322	80,766
Adjusted R^2	0.142	0.126	0.054	0.019	0.127
Agenda Item FE	Yes	Yes	Yes	Yes	Yes

B. Exit – Additional Analyses

According to Edmans (2009), Dasgupta and Piacentino (2015) and others, in addition to voting, shareholders can influence a firm's actions by selling the stock or threatening to sell the stock. While some index funds fully replicate their benchmark index, and therefore have no option to exit, other index funds use statistical sampling to replicate their benchmark index and therefore do have the flexibility to selectively exit firms.

In Table A3, we examine fund exit behavior. The dependent variable *Exit* is equal to one if a fund exits a position that it held the previous year, and zero otherwise. The independent variables of interest are *IndexFund* (our treatment variable), fund size ($logAUM$), and fund diversification ($log(\#Holdings)$). In Columns 1 and 3 we find that overall, index funds were approximately 17 percentage points less likely than active funds to exit a given stock in a given year, and that this comparison holds when sweeping out variation using either firm and year or firm-by-year fixed effects.

However, index funds are on average larger and much more diversified in their holdings than active funds are. Controlling for measures of fund size and diversification (Columns 2 and 4), we find that the results are similar: Index funds are 13 percentage points less likely than active funds to exit a given stock in a given year. Overall, the evidence is clear: index funds are much less likely to exit a given position, and thus have a weaker threat of exit against firm management, than active funds.

Table A3
Fund Exit

The table presents OLS panel regression estimates that compare the propensity to exit a position between index funds and active funds. The dependent variable, $Exit_{ijt}$, equals 1 if a fund exits a position and 0 otherwise. $IndexFund_i$ equals 1 if the fund is an index fund and 0 if the fund is an active fund. Robust standard errors clustered by fund and firm are shown in parentheses. *, **, *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1) $Exit_{ijt}$	(2) $Exit_{ijt}$	(3) $Exit_{ijt}$	(4) $Exit_{ijt}$
$IndexFund_i$	-0.175***	-0.133***	-0.171***	-0.131***
	(0.013)	(0.014)	(0.013)	(0.014)
$logAUM_{it}$		-0.013***		-0.012***
		(0.002)		(0.002)
$log(\#Holdings_{it})$		-0.030***		-0.030***
		(0.005)		(0.005)
Observations	3,289,251	3,289,251	3,284,457	3,284,457
Adjusted R^2	0.081	0.091	0.141	0.151
Firm FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Firm × Year FE	No	No	Yes	Yes

C. Heckman Correction

In Table 5 of the paper we present comparisons of index versus active funds' voting on contentious agenda items. The estimates in the paper use firm- and year fixed effects and firm-by-year fixed effects. The latter approach compares voting by active and index funds at the same annual meeting – that is, the firm-by-year fixed effects specification absorbs even time-varying firm characteristics. As a result, those estimates address any possibility of confounding variables at the firm-year level that could bias our comparisons. However, there is still the potential for selection bias. Selection bias in funds' voting arises from the funds' portfolio choices. If index funds tend to hold better-run firms, or vice versa, then the gap in fund voting behavior might be explained by their holdings and not by their monitoring.

To explicitly correct for selection bias in fund holdings, we use our Russell research design as the first stage in a Heckman (1979) correction model. Specifically, we estimate the following two-stage model:

$$\begin{aligned} Observed_{ijt} = Probit(&\tau IndexFund_i \\ &+ \xi_1 R1000 \rightarrow R2000_{jc} \times PostAssignment_{ct} \times IndexFund_i \\ &+ \xi_2 R2000 \rightarrow R1000_{jc} \times PostAssignment_{ct} \times IndexFund_i \\ &+ \mu_1 R1000 \rightarrow R2000_{jc} \times PostAssignment_{ct} \\ &+ \mu_2 R2000 \rightarrow R1000_{jc} \times PostAssignment_{ct} \\ &+ \phi_{jc} + \chi_t + \gamma_{jt} + \nu_{ijct}) \end{aligned} \tag{1}$$

$$Y_{ijt} = \beta IndexFund_i + \alpha InverseMillsRatio_{ijt} + \lambda_j + \kappa_t + \eta_{jt} + \epsilon_{ijt} \tag{2}$$

Equation (1) uses our cohort difference-in-differences specification to generate exogenous variation in fund ownership. *Observed* equals 1 if fund i holds stock j on date t, and zero otherwise; *IndexFund* equals 1 if the fund is an index fund, and 0 otherwise; *R1000* $\rightarrow$ *R2000* equals 1 if a firm switched from the Russell 1000 to the Russell 2000, whereas *R2000* $\rightarrow$ *R1000* equals 1 if a firm switched from the Russell 2000 to the Russell 1000. $PostAssignment$ equals 1 if the firm-year is post index assignment, and 0 if it is pre index assignment. ϕ_{jc} and χ_t are, respectively, firm-by-cohort and year fixed effects; alternatively, γ_{jt} are firm-by-year fixed effects. The results for the first stage (Equation (1)) are reported in Table A4.

Equation (2) shows the second stage, which examines an outcome (specifically, fund voting) as a function of index fund status after including the *InverseMillsRatio* (i.e., the Heckman correction term from Equation (1)). λ_j are firm fixed effects and κ_t are year fixed effects; alternatively, η_{jt} are firm-by-year fixed effects.

The second-stage Heckman-corrected estimates are shown below in Table A5. In both models with firm and year fixed effects, and firm-by-year fixed effects, after controlling for selection bias concerns, we continue to find that index funds are 8.5 and 7.7 percentage points more likely to vote with management on contentious votes. These results again support the conclusion that index funds cede power to firm management in their voting decisions.

Table A4
Observation Equation

The table presents the estimated observation equation (the Heckman first stage, equation (1)) that a given fund is observed holding a given firm. The sample for this estimate is the panel of all firm-years in the Russell sample, interacted with all mutual funds that held at least one firm in that sample. The dependent variable $Observed_{ijt}$ is a dummy that equals 1 if fund i held a position in firm j in year t. Robust standard errors clustered by fund are in parentheses. *, **, *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

	$Observed_{ijt}$ (1)	$Observed_{ijt}$ (2)
$R1000 \rightarrow R2000_j \times$	0.071***	0.088***
$PostAssignment_t \times IndexFund_i$	(0.015)	(0.017)
$R2000 \rightarrow R1000_j \times$	-0.066***	-0.063***
$PostAssignment_t \times IndexFund_i$	(0.024)	(0.028)
$R1000 \rightarrow R2000_j \times$	-0.231***	-0.033***
$PostAssignment_t$	(0.016)	(0.007)
$R2000 \rightarrow R1000_j \times$	0.063***	0.021**
$PostAssignment_t$	(0.015)	(0.009)
$IndexFund_i$	0.622***	0.706***
	(0.053)	(0.061)
Model	Probit	Probit
Observations	11,907,984	11,907,984
Firm × Cohort FE	Yes	No
Year FE	Yes	No
Firm × Year FE	No	Yes

Table A5
Heckman Corrected Estimates

The table presents the estimated voting equation (the Heckman second stage, equation (2) that a given fund is observed voting with management's recomendation on a contentious vote. The sample for this estimate is the panel of all firm-years in the Russell sample, interacted with all mutual funds that held at least one firm in that sample. Robust standard errors clustered by fund are in parentheses. *, **, *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

	$VotedWithMgmt$ (1)	$VotedWithMgmt$ (2)
$IndexFund_i$	0.085***	0.077**
	(0.028)	(0.031)
$InverseMillsRatio_{ijt}$	-0.089***	-0.097***
	(0.030)	(0.035)
Observations	254,038	254,038
Adjusted R^2	0.72	0.112
Firm FE	Yes	No
Year FE	Yes	No
Firm × Year FE	No	Yes

D. Balance Tests and Pretrends

D.1. Balance Tests

In order to generate a valid comparison, it is important that our treated and control firms in our Russell sample are similar *ex ante* and differ only by their index switching status. For example, if prior to treatment, the firms above the upper band were systematically governed worse than firms below it, our estimates would be biased toward finding a spurious association between index fund investment and good governance. We thus run standard RDD balance tests on pre-treatment firm characteristics to check that treated and control firms on either side of both bands are similar *ex ante*, particularly in terms of fund ownership and firm governance.

In Table A6 we compare pre-treatment means of total index funds ownership (Column 1), stock returns (Column 2), returns volatility (Column 3), E-index (Column 4), Board independence (Column 5), and dual class shares (Column 6) for switchers versus stayers. In each case we measure the dependent variable in the last pre-treatment year. We find no significant differences between treated and control firms in any of the outcome variables examined.

Furthermore, Figure 3 in the paper presents formal regression discontinuity (RD) plots for fund ownership and firm governance, measured in the last pretreatment year for each firm, with local polynomial control functions fitted on either side of each band. Again, we observe no significant difference at the treatment cutoff (the upper or lower band respectively). Furthermore, in each case the treated and control firms also have similar overall *levels* of fund ownership and governance. Hence, we conclude that our treated and control groups are well-balanced *ex ante*, and that our research design does not suffer from the problems highlighted by Wei and Young (2020).

Table A6
Balance Tests: Comparison of Pretreatment Firm Characteristics

The table presents comparisons of pretreatment means between switchers (firms that switched indexes) versus stayers (firms in the same cohort and near the same band that did not switch indexes) on either side of the yearly Russell bands from 2007-2016. Robust standard errors clustered by firm are shown in parentheses. *, **, *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	$IndexOwn_{j,t-1}$	$Return_{j,t-1}$	$RtnVolatility_{j,t-1}$	E-Index$_{j,t-1}$	IndepBoardPct$_{j,t-1}$	DualClass$_{j,t-1}$
$R1000 \rightarrow R2000_j \times$	0.36	-0.03	0.00	-0.02	0.00	-0.03
$PostAssignment_t$	(0.76)	(0.05)	(0.01)	(0.15)	(0.01)	(0.04)
$R2000 \rightarrow R1000_j \times$	0.03	-0.00	-0.00	0.00	0.00	0.03
$PostAssignment_t$	(0.40)	(0.03)	(0.00)	(0.11)	(0.02)	(0.03)
Observations	820	775	793	471	457	471
Adjusted R^2	0.414	0.318	0.445	0.402	0.051	-0.026
Cohorts	2007-2016	2007-2016	2007-2016	2007-2016	2007-2016	2007-2016
Cohort × Band FE	Yes	Yes	Yes	Yes	Yes	Yes

D.2. Comparing Pretrends

The results of the balance tests suggest that our treated and control firms on either side of both bands are similar *ex ante* in terms of fund ownership and firm governance. However, this is not strictly necessary for the validity of our difference-in-differences design, as the firm-by-cohort fixed effects difference out any non-time-varying imbalance between treated and control firms. Rather, the identifying assumption is *parallel trends*: that is, in the absence of treatment, the treated firms would have had the same average trend in outcomes as the control firms did. This is inherently untestable, but a standard check is to compare trends in outcomes between treated and control firms in the years prior to treatment (pretrends).

We compute the pretrend for each firm in each cohort by regressing the outcome variables measured in years -3, -2, -1 on event-time t plus firm and calendar-year fixed effects. In

Table A7 we show the comparison of pretrends between treated and control firms. There is no economically significant difference in pretrends between treated and control firms for any of the firm characteristics. For two characteristics (return volatility and the E-index), across one of the bands, the difference is statistically significant at the 10% level, but the magnitude is small.

Table A7

Pretrend Tests: Comparison of Trends in Pretreatment Firm Characteristics

The table presents comparisons of pretreatment trends between switchers (firms that switched indexes) versus stayers (firms in the same cohort and near the same band that did not switch indexes) on either side of the yearly Russell bands from 2007-2016. Robust standard errors clustered by firm are shown in parentheses. *, **, *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1) $IndexOwn_{j,t-1}$	(2) $Return_{j,t-1}$	(3) $RtnVolatility_{j,t-1}$	(4) E-Index$_{j,t-1}$	(5) IndepBoardPct$_{j,t-1}$	(6) DualClass$_{j,t-1}$
$R1000 \rightarrow R2000_j \times$	-0.02	-0.01	0.01*	-0.09*	0.00	0.01
$PostAssignment_t$	(0.13)	(0.03)	(0.00)	(0.05)	(0.02)	(0.00)
$R2000 \rightarrow R1000_j \times$	0.01	0.04	0.00	0.05	0.00	0.00
$PostAssignment_t$	(0.13)	(0.02)	(0.00)	(0.03)	(0.01)	(0.00)
Observations	808	808	808	808	457	808
Adjusted R^2	0.008	0.347	0.006	0.030	0.051	0.030
Cohorts	2007-2016	2007-2016	2007-2016	2007-2016	2007-2016	2007-2016
Cohort × Band FE	Yes	Yes	Yes	Yes	Yes	Yes

D.3. Varying the Selection Window

We also examine the robustness of our results when we vary the size of the window, around each band in each year, in which our treated and control firms must fall. If measurement error in the forcing variable is affecting our results – in particular, biasing our estimates due to selection or causing our estimates to have low power – then as we narrow the window our results should disappear and vice versa.

Table A8 presents the results when we vary the window of selection around the bands each year. With both the narrower and wider windows, across both bands and across all categories of fund ownership, the results are similar to our main estimates. Thus, our results are robust to alternate window sizes; this finding is inconsistent with selection bias in our setting.

Table A8
Varying Window Size

The table presents estimates of the effects of Russell index switches on investment fund ownership when we vary the window of selection around the bands each year. Panel A shows the results when we narrow the window from 100 ranks to 50 ranks around each band. Panel B shows the results when we widen the window from 100 to 150 ranks around each band. The sample consists of stocks that were "potential switchers" near the yearly Russell upper and lower bands from 2007 to 2016, three years before and after index assignment for each firm in each cohort. Robust standard errors clustered by firm and year are shown in parentheses. *, **, *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Window = ±50 Ranks

	(1) $IndexOwn_{jt}^{R2000}$	(2) $IndexOwn_{jt}^{R1000}$	(3) $IndexOwnjt$	(4) $ActiveOwn_{jt}$
$R1000 \rightarrow R2000_j \times$	1.58***	-0.20***	0.80*	-2.94**
$PostAssignment_t$	(0.18)	(0.03)	(0.38)	(0.99)
$R2000 \rightarrow R1000_j \times$	-1.60***	0.22***	-1.08**	2.22**
$PostAssignment_t$	(0.13)	(0.02)	(0.37)	(0.91)
Observations	2,300	2,300	2,300	2,300
Adjusted R^2	0.469	0.475	0.831	0.729
Window	50	50	50	50
Year FE	Yes	Yes	Yes	Yes
Stock × Cohort FE	Yes	Yes	Yes	Yes

Panel B: Window = ±150 Ranks

	(1) $IndexOwn_{jt}^{R2000}$	(2) $IndexOwn_{jt}^{R1000}$	(3) $IndexOwnjt$	(4) $ActiveOwn_{jt}$
$R1000 \rightarrow R2000_j \times$	1.67***	-0.22***	1.19***	-2.36***
$PostAssignment_t$	(0.11)	(0.02)	(0.29)	(0.73)
$R2000 \rightarrow R1000_j \times$	-1.58***	0.22***	-1.22***	0.98
$PostAssignment_t$	(0.06)	(0.01)	(0.20)	(0.58)
Observations	7,321	7,321	7,321	7,321
Adjusted R^2	0.560	0.562	0.845	0.712
Window	150	150	150	150
Year FE	Yes	Yes	Yes	Yes
Stock × Cohort FE	Yes	Yes	Yes	Yes

E. Russell Research Designs

This section discusses our Russell research design in greater detail, and shows how it differs from previous Russell research designs. As discussed in the main text, one difficulty with this setting is that Russell does not release the true rankings that determined index assignment; instead, researchers must impute them. This introduces noise in the forcing variable which can severely bias estimates in a regression discontinuity design (RDD). Our research design addresses this issue by exploiting the panel nature of our data.

We divide the discussion of Russell methodology into the following subsections:

- In subsection E.1 we discuss the issue of noise in a forcing variable when using an RDD. We simulate data and show that when there is noise in the forcing variable, RDDs recover biased estimates. We then show that our cohort difference-in-differences recovers the true effect.
- In subsection E.2 we describe our research design in greater detail.
- In subsection E.3 we discuss our approach in the context of prior Russell research designs. Consistent with two recent papers by Wei and Young (2020) and Gloßner (2020), we show evidence that the methodologies in prior studies lead to biased estimates. In particular, we replicate selected results in Appel, Gormley, and Keim (2016) and show that Appel et al. (2016) find that index investing leads to a sharp decrease in dual class shares, yet, in the data only a handful of firms *ever* change their share class structure. These results suggest their methodology is subject to selection bias. We then show that our methodology does not suffer from these issues: we find no change in dual class shares correctly estimated.

E.1. Identification when the Forcing Variable is Measured with Noise

The Russell Index rebalancing procedure creates a discontinuity in treatment status: some firms are assigned to the Russell 1000, while other firms are assigned to the Russell 2000. The discontinuity in treatment status across the rank-1000 cutoff (pre-banding) and across the yearly bands (post-banding) suggests a regression discontinuity design (RDD). However, there are features of the Russell setting that make an RDD undesirable. The main issue is that Russell does not publicly release the true *unadjusted* rankings that determine index assignment; instead, researchers must impute them. In our sample, our proxy rankings (based on CRSP and Compustat data) predict the actual index assignments with 99.5% accuracy, but there could still be significant errors in the rankings of individual firms. This is a concern because errors in measuring the forcing variable bias the RDD control function to be too flat, and produce spurious or upward biased estimates of treatment effects (Pei & Shen, 2017). Note that a fuzzy RDD, which adjusts for non-compliance with treatment assignment, does not address this issue.

Pei and Shen (2017) point out that when treatment is assigned by an otherwise arbitrary threshold and treatment status is observed perfectly but the forcing variable is observed with noise, conventional RDD estimates may be biased away from zero – that is, they produce spurious estimates of the treatment effect. To illustrate this, we simulate data and show that (i) a variety of regression discontinuity designs generate spurious estimates under these conditions, and (ii) our panel difference-in-differences research design recovers the true treatment effect.

We simulate a sample of 200 firms and we rank them on a simulated forcing variable from -100 to +100, with a treatment threshold at rank = 0. We construct the forcing variable so that it is a smooth line (i.e., there is no discontinuity at rank = 0 or any other point). We then simulate an outcome variable as a linear function of the ranking (See Figure 4 Panel A in the main paper). We then add normally distributed measurement noise to each firm's ranking and re-sort them on the basis of the ranking measured with noise. Figure 4 Panel B in the main paper shows the result graphically. The control functions on either side of the threshold are attenuated (*i.e.* their slope is too flat), and as a result there is a large, spurious jump in the outcome variable at the threshold. Table A9 columns 1-6 show that various RDD estimators all produce a large and significant spurious treatment effect. Some of these specifications have been used in prior studies using Russell settings.

We proceed from the insight that if the econometrician observes the sample firms repeatedly – at least once before and after treatment – she can instead use a firm-year panel to compare the change in outcomes before and after treatment, for treated versus untreated firms near the cutoff, with firm fixed effects. The firm fixed effects eliminate the need for a control function – because each firm has a single ranking relative to the cutoff at assignment, any control function would be absorbed by the fixed effects – and eliminates the bias that is present in the RDD estimates.

Table A9 column 7 shows the result when we expand the simulated data for each firm to three years pre- and post-treatment, and run the difference-in-differences (DiD) estimate with firm fixed effects. We see that our cohort difference-in-differences design recovers the true treatment effect. Moreover, the results in Table A9 demonstrate another advantage of our research design. The combination of a larger sample size (six years per firm instead of only one) and the firm-by-cohort fixed effects, which sweep out heterogeneity at the level of each individual firm, produce estimates with much higher statistical power than the RDD estimates. In Table A9, our difference-in-differences estimate has a standard error that is less than one-tenth the size of the average standard errors for the RDD estimates.

Table A9

Identification when the Forcing Variable is Measured with Error

The table presents comparisons of regression discontinuity design (RDD) and difference-in-differences (DiD) estimates on simulated data, in which there is no change in the outcome variable across the cutoff (true treatment effect = 0.0) and the forcing variable is measured with error. Columns 1-6 show a variety of RDD estimates. Column 7 shows the cohort DiD estimate. Robust standard errors clustered by firm are shown in parentheses. *, **, *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$Treated_i$	29.058***	23.465***	31.130***	20.920***	28.965***	33.391***	
	(10.658)	(5.770)	(4.927)	(8.984)	(7.403)	(9.848)	
$Treated_i \times PostTreat_t$							-0.004
							(0.707)
Control Function	LPoly	LPoly	LPoly	Linear	Quadratic	Cubic	
Kernel	Triangular	Epachenikov	Rectangular	Rectangular	Rectangular	Rectangular	
Observations	200	200	200	200	200	200	1,200
Adjusted R^2	0.690	0.690	0.690	0.690	0.691	0.690	0.963
Firm FE	No	No	No	No	No	No	Yes

E.2. Panel Difference-in-Differences Design

To deal with the issue of measurement error in the RDD forcing variable, we exploit the panel nature of our data. Specifically, we estimate a cohort difference-in-differences design, with firm-by-cohort fixed effects. To see why this approach addresses measurement error in the forcing variable, consider the RDD estimate from the following model:

$$\begin{aligned} Y_j = \beta_1\, I\{R1000 \rightarrow R2000_j\} + \beta_2\, I\{R2000 \rightarrow R1000_j\} \\ +\gamma\,(truerank_j + measurementerror_j) + \epsilon_j, \end{aligned} \tag{3}$$

where $I\{R1000 \rightarrow R2000_j\}$ equals 1 if a stock switches from the Russell 1000 to the 2000, $I\{R2000 \rightarrow R1000_j\}$ equals 1 if a stock switches from the Russell 2000 to the 1000, and γ is the coefficient on a linear control function (that is measured with error). Standard arguments (Wooldridge, 2008) show that measurement error in *caprank* causes $\hat{\gamma}$ to be biased toward zero. Since *truerank* is correlated with treatment status, the estimated treatment effect $\hat{\beta}$ is biased away from zero. This bias is present in general (*i.e.* it occurs for *any* choice of control function).

Instead, we estimate the following difference-in-differences model:

$$\begin{aligned} Y_{jt} = \beta_1\, I\{R1000 \rightarrow R2000_j\} \times I\{PostAssignment_t\} + \\ \beta_2\, I\{R2000 \rightarrow R1000_j\} \times I\{PostAssignment_t\} + \phi_j + \lambda_t + \epsilon_{jt}, \end{aligned} \tag{4}$$

where ϕ_j and λ_t are firm and date fixed effects and $I\{PostAssignment_t\}$ is an indicator variable that equals 1 after index re-balancing.[†] We compare the outcome variable before

[†] Importantly, this means that β_1 and β_2 – the effects of switching from the R2000 to the R1000 and vice versa – are identified from disjoint sets of treated and control stocks. The stock-by-cohort fixed effects sweep out any non-time-varying differences between treated and control stocks, while the year fixed effects remove aggregate trends in firm behavior or ownership.

treatment versus after treatment, with a fixed effect applied to each firm in each cohort. Because each firm had a fixed ranking within the cohort, the fixed effects ϕ_j absorb any association of the outcome variable with both the true ranking *and* the error in the proxy ranking for each firm. Thus, the specification (4) estimates the treatment effect of switching indexes as would a correctly measured RDD, but in a way that is not sensitive to measurement error in the forcing variable.

This approach is not a panacea. Errors in the proxy rankings could also cause us to select firms that were farther away from the bands than we know, which would introduce selection bias into the sample. We examine the possibility of selection bias in two ways. First, in Section D, we present formal balance tests which show that the firms on either side of each band are indistinguishable, before treatment, on a variety of outcomes both in levels and trends. Second, in Section D.3 we document that our estimates remain stable as we vary the windows around the bands. These results are inconsistent with selection bias.

Our methodology differs from previous papers that use Russell reconstitutions in two important dimensions. First, we develop a research design that explicitly uses Russell index reconstitutions in the post-2006 period. Thus, our results reflect this more recent period, during which index investing is at all-time highs.

Second, unlike previous RDD research designs, our difference-in-differences specification uses firm fixed effects to sweep out unobserved heterogeneity among firms. Among other advantages, this means that our estimates are not biased by noise in the measurement of the forcing variable, which can be an issue in both sharp and fuzzy RDD specifications (Pei & Shen, 2017).

E.3. Comparing Russell Research Designs

In this subsection, we replicate the Appel et al. (2016) Russell research design. For each year from 1998 to 2006, we select firms within a bandwidth of +/- 250 ranks of the yearly rank-1000 cutoff in those prebanding years. We add fund holdings from Thomson Reuters and the presence or absence of a dual-class share structure from ISS (data originally collected by Riskmetrics).

Table A10 shows the results when we run first- and second-stage estimates using the research design and sample period of Appel et al. (2016) (see their Table 3 and Table 7). Our results are not perfectly identical to theirs – likely due to differences in sample construction and the definition of index fund status – but are similar in sign, magnitude and significance in all cases. In particular, we replicate their conclusion that greater index fund ownership, instrumented via assignment to the Russell 2000 index in June, leads to a lower likelihood of the firm having a dual-class share structure by December of the same year.

However, this result is at odds with the fact that firms very rarely change their share class structure. Out of *all* 61,727 firm-years covered by the ISS/Riskmetrics data from 1990 to 2006, we find only 147 cases (0.21%) in which a firm's share class structure changed in either direction. Moreover, the results in Table A10 are estimated using a subsample of all firms – those near the Russell cutoff. Out of the 4,250 firm-years in our replication of Appel et al. (2016)'s estimates, there are only six firms, total, that changed their share structure. For these six firms in the replication sample that did change their share class structure, Table A11 shows that the direction of the change was uncorrelated with their index assignment. Of the four firms that changed from dual- to single-class, three were in the Russell 1000 and one in the Russell 2000. Of the two firms that changed from single- to dual-class, one was in the Russell 1000 and one in the Russell 2000.

Table A10
Replication of Table 3 and 7 of Appel, Gormley, and Keim (2016)

The table presents estimates of the effects of Russell index assignment using the approach of Appel et al. (2016). The sample consists of stocks that were within +/-250 ranks, using the published (float-weighted) Russell rankings, near the yearly rank-1000 cutoff from 1998 to 2006. Panel A shows the results when we regress index fund ownership on Russell 2000 index membership. Panel B shows the results when we regress a dummy variable for dual-class share structure on index fund ownership instrumented using Russell 2000 index membership. Both $IndexFundOwn$ and $DualClass$ are scaled by their sample standard deviation. Robust standard errors clustered by firm are shown in parentheses. *, **, *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

A: First Stage: Effects of Index Assignment on Passive Ownership

	(1) $IndexFundOwn_t$	(2) $IndexFundOwn_t$	(3) $IndexFundOwn_t$
R2000	1.103***	1.086***	1.065***
	(0.084)	(0.083)	(0.089)
Bandwidth	250	250	250
Polynomial Order N	1	2	3
Float Control	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	4,250	4,250	4,250

B: Second Stage: Effects of Index Fund Ownership on Share Class Structure

	(1) $DualClass_t$	(2) $DualClass_t$	(3) $DualClass_t$
$(IndexFundOwn_t = R2000_t)$	-1.226***	-1.189***	-1.274***
	(0.218)	(0.213)	(0.255)
Bandwidth	250	250	250
Polynomial Order N	1	2	3
Float Control	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	1,700	1,700	1,700

Table A11
List of Firms that Changed Share Class Structure in the Replication Sample

The table presents the six firm-years that changed their dual-class share status in our replication of Appel, Gormley, Keim (2016). *DualClass* is an indicator variable that takes the value 1 if a firm has dual class shares in year t, and 0 otherwise.

Ticker	Year t	$DualClass_t$	$DualClass_{t-1}$	$Index_t$
NWAC	2000	0	1	R1000
SPOT	2002	0	1	R1000
IM	2002	0	1	R1000
EXP	2006	1	0	R1000
CBM	2002	1	0	R2000
CW	2006	0	1	R2000

Next, we reexamine these findings using our panel difference-in-differences approach. Each June from 1998 to 2006, we select a cohort of firms within +/-100 ranks of the rank-1000 cutoff that determined index assignment in the pre-banding period. We do not use the published (*float-adjusted*) Russell rankings but instead impute the *unadjusted* Russell rankings following Chang, Hong, and Liskovich (2015) and Gloßner (2020). We include observations for three years pre- and post-assignment for each firm in each cohort. We then run our difference-in-differences estimates on the *prebanding* Russell cohort sample.

Table A12 shows the estimated effects of index switching on dual-class share structure. In both directions we find zero effects, precisely estimated. The precision is important to note, because Appel, Gormley, and Keim (2019) raise the concern that research designs relying only on firms that switch indexes could lack statistical power. This does not appear to be an issue for our design. The standard errors in Table A12 are significantly smaller than those in Table A10. That is, using a sample with fewer firms and a tighter bandwidth ($\pm$100 ranks, compared to $\pm$250 ranks), our cohort difference-in-differences estimator has noticeably higher statistical power.

Table A12
Panel Difference-in-Differences Estimates

The table presents estimates in the prebanding period using our panel difference-in-differences approach. The sample consists of firm-years for three years pre- and post-treatment for firms that were within +/-100 ranks, using the imputed (unadjusted) Russell rankings, near the yearly rank-1000 cutoff from 1998 to 2006. Robust standard errors clustered by firm and year are shown in parentheses. *, **, *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1) $DualClass_t$
$R2000 \rightarrow R1000_j \times$	0.004
$PostAssignment_t$	(0.018)
$R1000 \rightarrow R2000_j \times$	-0.012
$PostAssignment_t$	(0.021)
Window	$\pm$100
Years	1995-2008
Cohorts	1998-2006
Firm $\times$ Cohort FE	Yes
Year FE	Yes
Observations	6,192

References

Appel, I., Gormley, T. A., & Keim, D. B. (2016). Passive investors, not passive owners. *Journal of Financial Economics*, *121*, 111-141.

Appel, I., Gormley, T. A., & Keim, D. B. (2019). Identification using Russell 1000/2000 index assignments: A discussion of methodologies. *Working Paper*.

Chang, Y.-C., Hong, H., & Liskovich, I. (2015). Regression discontinuity and the price effects of stock market indexing. *Review of Financial Studies*, *28*, 212-246.

Dasgupta, A., & Piacentino, G. (2015). The Wall Street Walk when blockholders compete for flows. *The Journal of Finance*, *70*(6), 2853–2896.

Edmans, A. (2009). Blockholder trading, market efficiency, and managerial myopia. *The Journal of Finance*, *64*(6), 2481–2513.

Gloßner, S. (2020). The effects of institutional investors on firm outcomes: Empirical pitfalls of quasi-experiments using Russell 1000/2000 index reconstitutions. *Forthcoming, the Critical Finance Review*.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, *47*, 153-161.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, *37*, 187–204.

Pei, Z., & Shen, Y. (2017). The devil is in the tails: Regression discontinuity design with measurement error in the assignment variable. In *Regression discontinuity designs: Theory and applications* (pp. 455–502). Emerald Publishing Limited.

Wei, W., & Young, A. (2020). Selection bias or treatment effect? A re-examination of Russell 1000/2000 index reconstitution. *Forthcoming, the Critical Finance Review*.

Wooldridge, J. (2008). *Introductory Econometrics: A Modern Approach*. Cengage Learning.