The Reg SHO Reanalysis Project: Reconsidering Fang, Huang and Karpoff (2016) on Reg SHO and Earnings Management

Bernard Black, Hemang Desai, Kate Litvak, Woongsun Yoo, and Jeff Jiewei Yu

Abstract: During 2005-2007, the SEC conducted a randomized trial in which it removed short-sale restrictions from one-third of the Russell 3000 firms. Early studies found minor effects of the experiment on trading markets. More recently, many studies attribute a wide variety of outcomes to the experiment. We revisit principal findings in Fang, Huang and Karpoff (2016), on the experiment's effect on discretionary accruals (earnings management) of pilot firms. Using multiple measures of accruals, we find no evidence that the experiment affected accruals. Our findings have implications for the robustness of the results from many recent papers that study the short-sale experiment.

Conflict-of-interest disclosure statement

Bernard Black

I have nothing to disclose

Hemang Desai

I have nothing to disclose

Kate Litvak

I have nothing to disclose

Woongsun Yoo

I have nothing to disclose

Jeff Yu

I have nothing to disclose

The Reg SHO Reanalysis Project:

Reconsidering Fang, Huang and Karpoff (2016) on Reg SHO and Earnings Management

Introduction

In July 2004 the SEC announced a randomized trial in which it temporarily suspended short sale restrictions for approximately 1000 ("pilot" or "fully treated") firms in the Russell 3000 Index (below, R3000), leaving some but not all of the prior restrictions in place for the remaining firms (original controls). Specifically, the SEC suspended the New York Stock Exchange (NYSE) uptick rule and the similar Nasdaq bid test for pilot firms during the experiment period (May 2, 2005 to July 6, 2007). Under NYSE Rule 440B, as in effect prior the experiment, a short sale was permitted only following on a plus tick or a zero-plus tick ("uptick rule") and under Nasdaq Rule 3350, short sales in National Market Securities were not allowed at or below the bid when the current bid was at or below the previous bid ("bid test").[1]

Initial studies of the experiment found little to no impact of removing short-sale restrictions on returns, the level of open short interest (below, simply "short interest") or market quality measures. The SEC's Office of Economic Analysis (OEA, 2007) found that relaxing the short-sale constraints did not have a significant impact on returns or short interest of the treated (pilot) firms, but the treated firms did experience an increase in short sale volume. Diether, Lee and Werner (2009) found higher short-sale volume for the pilot firms, but no significant effect on daily returns or volatility. The primary takeaway from these early studies was that removing short-sale restrictions did not significantly affect returns, short interest, or volatility. The SEC relied on these early studies to remove the short-sale restrictions for all firms when the experiment ended in July 2007.

In the years since this experiment however, a large number of finance and accounting papers report evidence that suspension of short sale restrictions had wide ranging direct and

[1] Below, we generally refer to both the NYSE and Nasdaq rules as the "uptick rule." Where we want to distinguish between the two rules, we call them the "NYSE uptick rule" and the "Nasdaq bid test." The American Stock Exchange (AMEX) had an uptick rule similar to the NYSE. We generally refer to both NYSE and AMEX firms a NYSE firm.

indirect effects on pilot firms. For example, these studies report that the pilot program affected share returns, short interest, earnings management, research and development outlays, investments, capital expenditures, acquisition activity, management compensation, audit fees, and more.[2]

The outcomes reported in many of these studies seem implausible for a variety of reasons. First, as explained in Diether et al. (2009) and Alexander and Peterson (2009), suspension of the NYSE uptick rule was expected to have a modest impact on short selling constraints. With a tick size of a penny, heavy trading volume, including heavy arbitrage trading (which will often generate upticks unrelated to fundamentals), the uptick rule was widely believed at the time to provide a real but modest impediment to arbitrage-related trading, but to have little effect on valuation based (fundamentals based) short selling.

Second, regional exchanges which traded shares of NYSE firms did not impose the uptick rule. Similarly, electronic exchanges which accounted for around 40% of the trading volume in Nasdaq firms at that time did not impose the bid test. Thus, firms on these venues had fewer constraints on short selling to begin with. Many firms also had publicly traded options, which provided an alternate avenue to direct short selling of shares. And for Nasdaq firms, the bid test was less restrictive than the uptick rule (Diether et al., 2009; Alexander and Petersen, 2009). Our conversations with practitioners and former regulators suggested that by 2004, the NYSE and Nasdaq restrictions had no meaningful effect on the ability of valuation-based short sellers to take short positions.

Third, in addition to suspending the uptick at all times for the pilot firms, the SEC also suspended this rule for the original control firms in the Russell 1000 (below, R1000) for after-hours trading (from 4:15 p.m. until the next day's opening of trading), suspended the uptick rule for all firms from 8:00 p.m. until the next day's opening, and the Nasdaq's bid test never applied after trading hours. These factors reduced the difference between the rules that applied to pilot

[2] The Internet Appendix to this paper lists more than 30 papers as of August 2019 that we are aware of that document a wide range of indirect effects on the pilot firms including effects on behavior of other capital market participants such as analysts and auditors.

versus other firms.

Finally, the experiment received little publicity nor management opposition when it was announced. We found no articles in the mainstream business press that discuss the rule or suggest manager opposition to the experiment. This is consistent with the notion that the suspension of the experiment was expected to have little economic impact on the firms.[3]

The early studies were consistent with the expectation that removing the short-sale restrictions would facilitate arbitrage-related short selling but have few other meaningful effects (Diether et al., 2009). However, as mentioned above, there are over 30 studies that document wide ranging effects – mostly indirect effects – on the pilot firms. But if, as the early studies show, and Litvak, Black and Yoo (2019) reconfirm, the experiment had limited direct effects on trading markets, with no evidence of an effect on short interest or share prices, it is hard to understand the causal channel that would produce these effects. Thus, it is plausible that that many of the recent findings may not be robust, if re-examined.[4]

In this paper, we re-examine the principal finding in a prominent, recent paper, published in this journal: Fang, Huang, and Karpoff (2016). The authors find that pilot firms reduced earnings management. In particular, FHK report that pilot firms have more negative performance-

[3] The SEC's announcement on June 23, 2004, of its plans to conduct the experiment, and the formal experiment announcement on July 28, 2004, were not covered in any standard business news sources. Between announcement and experiment launch in May 2005, we found only a few technical explanations of how the experiment would work. The experiment launch was noted in a short Dow Jones News Service story a few days earlier, with a Wall Street Journal summary the next day, not long enough to warrant a byline. The SECs decision in March 2007 to eliminate the uptick rule also attracted no opposition that we could find. FHK assertion (at p. 1255) that "The decision to eliminate all short-sale price tests prompted a huge back-lash from managers and politicians," is not supported by evidence from the time when the decision was made. Appendix A of this paper provides details on news stories during the period from announcement through the end of the experiment.

[4] A possible channel would posit that managers of pilot firms were afraid of bear raids, and therefore changed their behavior, even though the bear raids never came. This channel has several problems. First, it is implausible. As we document in the Appendix, prior to and during the experiment period, the experiment received almost no publicity. Second, given the lack of direct effects, one would expect managers to learn over time that the bears were not charging, and thus for any near-term reaction to fade over time. None of the recent studies assesses whether an initial reaction later fades. Third, as a statistical matter, a manager fear channel, engendered by news stories and managers talking to each other, the reactions of different firms would be non-independent, and thus standard errors, which assume independent observations across firms, would be biased downward.

matched discretionary accruals (PMDA) during the experiment period, and lower "F-score (a measure of the likelihood of a future misstatement). We revisit these findings using a variety of specifications, and accrual measures.

We began our reexamination by developing a pre-analysis plan, which specifies how we would approach the FHK research questions, if we wanted to test their conjecture. This pre-analysis plan was completed before measuring any outcomes, with minor exceptions (Black et al., 2019). Using the sample and specifications in the pre-analysis plan, we find no evidence that pilot firms had significantly lower accruals during the experiment period using four different accruals measures, including the PMDA measure used by FHK, or lower F-scores than the control firms. We then methodically move from our sample and specification to theirs but are unable to replicate their findings and find only insignificant results.[5]

I. Sample Selection

The SEC announced the short-sale experiment on July 28, 2004 and launched the experiment on May 2, 2005. It created lists of NYSE, AMEX, and Nasdaq national market firms included in the Russell 3000 (R3000) index and assigned one-third of the firms in each list to be treated, effectively at random. The R3000 list is updated every year on the last Friday in June. The SEC's adopting release states that the SEC used the R3000 list as of June 25, 2004.[6] The SEC's original list includes only 986 pilot firms (which the SEC called "Category A" firms) instead of 1000 because the SEC excluded, prior to randomization, firms that were either not listed on NYSE/AMEX or Nasdaq national market and hence were not subject to the short sale restrictions to begin with, as well as firms which became public after April 30, 2004.

[5] Moreover, FHK obtain statistical significance, given their coefficients, using standard errors clustered on both firm and year, which are far lower, for their outcomes, than standard errors clustered only on firm. But clustered standard errors are potentially downward biased with a small number of clusters (e.g., Cameron, Gelbach and Miller, 2008).

[6] Securities Exchange Act Release 34-50104 (July 28, 2004). The SEC conducted, in effect, a block randomized experiment, in which within each trading market (NYSE, AMEX, and Nasdaq national market), it ranked these firms by trading volume over June 2003 through May 2004, and chose every third firm (the 2nd, 5th, 8th, etc. on the within-market lists) to be treated. See OEA (2007).

We start with the R3000 list as of June 30, 2004 from Bloomberg which maintains monthly historical lists of R3000 index firms. The Bloomberg list includes all the 986 original pilot firms announced by the SEC in July 2004. Next, we merge the R3000 list with the CRSP monthly stock file for June 2004 and can match all 3000 firms. Then we match the R3000/CRSP list with the list of 986 pilot firms announced by the SEC and can match all 986 firms.[7]

As discussed in Litvak, Black, and Yoo (2019), the SEC changed its original experiment. The SEC first randomized the R3000 Index into fully treated (pilot) firms (one-third of the R3000), for which it suspended the uptick rule completely, and "original controls" (the remaining two-thirds of the R3000). Then, for "large" original control firms (those in the Russell 1000 Index, R1000 henceforth), the SEC suspended the uptick rule *after trading hours*. Thus, the original control firms in R1000 can be considered "partly treated" as short sale restrictions for these firms remained in place only during trading hours. The SEC referred to these partly treated large firms as Category B firms. Moreover, all original control firms on Nasdaq, regardless of size, were effectively partly treated, because the Nasdaq bid test never applied after trading hours. Unlike the list of Category A firms, the SEC did not publish an original list of Category B firms as of June 2004, or the full list of original control firms. Thus, we needed to create these lists ourselves. To do so, we follow the SEC's statement of its exclusion rules and exclude firms that were not listed on NYSE/AMEX or Nasdaq national market, and firms that began trading after April 30, 2004. These steps left us with a sample of 2,954 firms comprising 985 pilot firms (the 986 on the SEC list less one firm which was delisted on June 28, 2004) and 1,969 control firms, as of the July 28, 2004 announcement of the short-sale experiment.

We next update the sample to the start of the Reg SHO experiment on May 2, 2005. There were some additional exclusions from the above sample, which was as of July 28, 2004. Though, the SEC did not publish in 2004 a list of either category B or other original control firms, it did publish on April 13, 2005, updated lists of Category A and Category B firms which reflected

[7] We cross-checked the Bloomberg list with a list of R3000 firms, obtained from Russell via academic request. Relative to the Bloomberg list, the Russell list excludes two firms (Westport Resources Corp, listed by the SEC as a pilot firm. and Alaris Medical Systems) that were delisted on June 28, 2004 due to M&A activity.

exclusions principally due to mergers, acquisitions and bankruptcy. We develop an updated 2005 sample as of May 2, 2005 as follows. Following the SEC, we exclude 38 pilot firms and move one Nasdaq firm (ARBA, Ariba Inc.) to Category A. We also exclude 5 firms that ceased trading as of May 2, 2005 (no share prices on CRSP). For Category B firms, we exclude the 15 firms that the SEC excluded; and also exclude 5 other large firms that ceased trading as of May 2, 2005. Finally, we exclude 33 small Nasdaq firms and 24 small NYSE firms that ceased trading as of May 2, 2005.

These steps produce a preliminary 2005 sample of 943 pilot firms and 1,891 control firms, updated to the date of experiment. We then exclude financial firms (SIC 6000-6999), utilities (4900-4999) and three firms that did not file 10-K for fiscal 2004 resulting in a final "2005 Financial Analysis Sample" of 2,115 firms comprising 702 Pilot firms and 1,413 original control firms, which we refer to as our unbalanced panel. This is our preferred sample as it utilizes all R3000 firms with active trading as of May 2, 2005, subject to the exclusions mentioned above. We do not restrict the sample based on availability of covariates as one can obtain unbiased estimates without covariates due to the initial randomization. In contrast, FHK report their results for accruals using a balanced panel of firms with data throughout 2001-2010 (but report that they found similar results with an unbalanced panel, page 1262). Thus, to make our analysis comparable to theirs, we also create balanced panels requiring that data on covariates and the dependent variable (accruals or F-score) be available throughout 2001-2010. This balanced panel is 465 pilot firms and 856 control firms for accruals and 431 pilot firms and 779 control firms for F-score. The FHK balanced panel for both analyses comprises 388 pilot and 709 control firms.

Below, similar to FHK and most of the recent literature on the short-sale experiment, we ignore differences in the rules that applied to large (R1000) versus small (R2000) firms, and between NYSE and Nasdaq firms, and study a "Mixed Experiment," in which we compare all pilot firms to all original control firms (below, simply "control" firms). However, if we had found significant differences between the pilot and original control firms, we would have wanted to see if those results were consistent with the actual, multiple experiments that the SEC conducted, in which some original controls were partly treated (Litvak, Black, and Yoo, 2019), including

differences between partly treated and full control firms, and between the uptick rule and the bid test.

II. Research Design

We began, as detailed in our pre-analysis plan, by asking how *we* would design a study to answer the principal FHK research question - did the pilot firms reduce earnings management in response to the Reg SHO experiment? In particular, FHK test whether pilot firms reduce earnings management relative to control firms during the experiment period, and whether the difference in earnings management disappears after the experiment ends. FHK use performance matched discretionary accruals (PMDA) as their measure of earnings management. They report that pilot firms have more negative discretionary accruals than control firms during the experiment period, and this differential reverses after the experiment ends. We revisit their hypothesis.

To address this research question, we developed different, often simpler specifications than FHK, that rely more heavily on the initial randomization as a basis for balance between pilot and control firms. By relying on the initial randomization, we can use the larger unbalanced panel, and study simpler accruals measures. These design choices enhance precision and make it more likely that an effect, if one exists, will be statistically significant. We also prefer to rely on a simple comparison of means, rather than coefficients from a regression with time-varying, firm-level covariates, which might be outcomes of the short-sale experiment.

Sample Periods: We use 2001-2010 as our main sample period. We define which firm fiscal years are in the *sample* period using the Compustat convention, under which, if the fiscal year-end month is January-May, the fiscal year is the current calendar year minus 1 year; and if the fiscal year-end month falls in June through December, the fiscal year is the current calendar year. Thus, our sample period includes fiscal year ends from June 2001 through May 2011

The experiment started in May 2005 and ended in July 2007, so part of calendar 2005 was before the experiment period and part during the experiment. Similarly, part of calendar 2007 is during the experiment, and part in the post-experiment period. Thus, there is no easy way to define Pre, During and Post periods. FHK use calendar years to define Pre, During and Post. They define Pre as 1 if a firm's fiscal year end falls between January 1 and December 2003, drop calendar 2004

(but report that results are similar if they include 2004 in the Pre period), define During as 1 if the firm's fiscal year end falls between January 1, 2005 to December 2007, and define Post as 1 if a firm's fiscal year end falls between January 1, 2008 and December 31, 2010.

In our pre-specified research plan, we took a slightly different approach. We treat firm fiscal years which include at least 6 experiment months as *during* the experiment period. Thus, our During equals 1 for firm fiscal years ending October 2005 through December 2007. Fiscal years that end earlier than this are Pre, and fiscal years that end later are Post. We prefer to include 2004 in the pre-treatment period, because we would not expect pilot firms to change their accruals in 2004, even if they would do so in response to the actual experiment, which began in May 2005. However, we also report results using FHK's sample periods for closer comparison. Specifically, we (i) exclude fiscal years ending in 2004 from the pre-period; and (ii) use the FHK's definition of Pre, During, and Post periods.

Unbalanced and Balanced Panels: We prefer to use an unbalanced panel of all firms in the 2005 Financial Analysis Sample, for several reasons. First, we expect, and confirm, that sample attrition was not meaningfully affected by the short-sale experiment. Second, the experiment is randomized, so we do not expect there to be substantial treated-vs-control differences during the pre-experiment period. However, it is common, when using an unbalanced panel, to assess robustness with a balanced panel. FHK principally use a balanced panel with data for their entire sample period, but report that results are similar with an unbalanced panel. We will report results using the full 2005 Financial Analysis Sample, and a balanced panel drawn from this sample. [8]

Accruals Measures: We use FHK's PMDA measure and three simpler measures of earnings management. Our first measure is operating accruals. Following Healy (1985) and Sloan (1996), most prior research has focused on operating accruals (OPACC). However, Richardson et al. (2005) show that investing accruals, from capital investments in physical or intangible assets, are associated with lower earnings persistence and are mispriced. Therefore, we also consider total

[8] When we use a balanced panel with FHK specification, we use their definition of years (calendar years and exclude year 2004).

accruals (TOTACC), which includes both operating and investing accruals. These measures are defined as follows:

$$OPACC = (Earnings - CFO) / LagAssets$$

$$TOTACC = (Earnings - CFO - CFI) / LagAssets$$

where Earnings is earnings before extraordinary items from the statement of cash flows (Compustat annual data item IBC), and CFO and CFI are cash flow from operations and cash flow from investing activities, from the statement of cash flows (OANCF minus XIDOC, and IVNCF, respectively). LagAssets is the book value of total assets (AT) at the end of the prior fiscal year. Following Hribar and Nichols (2007), we winsorize accruals at 1% and 99% to reduce the influence of outliers and to address data entry errors in Compustat.

We prefer simpler measures for several reasons. First, given the random assignment of firms to pilot versus control, we expect, and confirm in Table I that there are no systematic differences in the two groups prior to the experiment period. Thus, one does not need to use PMDA to address performance imbalances between pilot and control firms. Second, PMDA has less power than simpler measures of accruals to detect earnings management (Dechow, Ge and Schrand, 2010, Dechow et al., 2011). We confirm below that PMDA estimates are substantially less precise than those for other accruals measures. Third, the papers showing that short selling increases prior to restatements, which provide a principal motivation for FHK, use these simpler measures.

However, we also study PMDA to allow direct comparison of our results to FHK. To facilitate comparison, we closely follow FHK and first estimate operating accruals cross-sectionally within each fiscal year and Fama-French 48 industry:

$$\frac{TA_{i,t}}{ASSET_{i,t-1}} = \beta_0 + \beta_1 \frac{1}{ASSET_{i,t-1}} + \beta_2 \frac{\Delta REV_{i,t}}{ASSET_{i,t-1}} + \beta_3 \frac{PPE_{i,t}}{ASSET_{i,t-1}} + \varepsilon_{i,t} \quad (1)$$

where *i* indexes firms and *t* indexes fiscal years. TA_t is earnings before extraordinary items and discontinued operations (IBC) minus operating cash flows (OANCF minus XIDOC) for year *t*, winsorized at 1% and 99%. $ASSET_{t-1}$ is total assets (AT) at the end of fiscal year *t*-1. So the dependent variable here is simply OPACC. ΔREV_t is the change in sales revenue (SALE) from

year *t*-1 to *t*. PPE_t is gross property, plant and equipment (PPEGT) at the end of year t. We follow FHK and require at least 10 observations in an industry-year to perform each cross-sectional estimation. Next, following FHK, we use the estimated coefficients from this model to calculate normal accruals $NA_{i,t}$:

$$NA_{i,t} = \widehat{\beta_0} + \widehat{\beta_1}\frac{1}{ASSET_{i,t-1}} + \widehat{\beta_2}\frac{(\Delta REV_{i,t} - \Delta AR_{i,t})}{ASSET_{i,t-1}} + \widehat{\beta_3}\frac{PPE_{i,t}}{ASSET_{i,t-1}} + \varepsilon_{i,t} \quad (2)$$

where $\Delta AR_{i,t}$ is the change in accounts receivables (RECT). We then calculate firm-year-specific abnormal accruals as $AA_{i,t} = (TA_{it} / ASSET_{i,t-1}) - NA_{i,t}$. Following FHK, we then match each firm-year observation with another observation from the same year and industry with the closest same-year return on assets (ROA_t defined as income before depreciation and amortization [OIBDP] divided by total assets [TA], for year t).[9] Finally, we calculate PMDA for each firm-year as the firm-year discretionary accruals minus the discretionary accruals for the matched firm-year.[10] Since we have already computed abnormal accruals when calculating PMDA, we decided to also report these abnormal accruals (AA) as a fourth accruals measure.

Difference-in-Differences (DiD) Specification: To test whether pilot firms had lower accruals than control firms during the experiment period, we estimate the following DiD model for each accruals measure over 2001-2010.

$$y_{i,t} = \beta_0 + \gamma_t + f_i + \beta_1 Pilot_i * During_t + \beta_2 Pilot_i * Post_t + \varepsilon_{i,t} \quad (3)$$

Here $y_{i,t}$ is the accruals measure; $Pilot_i$ = 1 for pilot (treated) firms and 0 for control firms; During = 1 during the experiment period (2005-2007), 0 otherwise; Post = 1 during the first three post-experiment years (2008-2010), after the SEC removed short-sale restrictions for all firms, the γ_t are year fixed effects (FE) and the f_i are firm FE. A negative coefficient β_1 on Pilot*During provides evidence that pilot firms reduce earnings management, relative to control firms, during

[9] While we follow FHK (and Kothari et al. 2005) in matching on ROA_t, we note that this is an error in a "causal" project. Since accruals affect ROA, one should match only on pre-treatment values of ROA. In our view, matches should be determined in 2004, and not changed after that.

[10] Following Hribar and Nichols (2007), we winsorize the variables that enter the regressions (equation 1 and 2) used to estimate the discretionary accruals at 1%/99%, compute PMDA, and winsorize again. FHK apparently winsorize the variables that enter the accruals regressions (see their Table I). They do not state whether they winsorize PMDA, but likely do so because they say that they follow Kothari et al. (2005), who do so.

the experiment period. The expected coefficient β_2 on Pilot*Post should be close to zero because short-sale restrictions were removed for all firms following the experiment. We can also directly test for a sign reversal in the Post period, relative to the experiment period, by replacing Pilot*During with (Pilot *(During or Post)) in eqn. (3). With this specification, FHK predict a positive sign for the coefficient on Pilot*Post, similar in magnitude to the predicted negative coefficient on Pilot*During in eqn. (3).

$$y_{i,t} = \beta_0 + \gamma_t + f_i + \beta_1 Pilot_i * (During\ or\ Post)_t + \beta_2 Pilot_i * Post_t + \varepsilon_{i,t} \quad (4)$$

Leads-and-Lags Specification: Both to assess whether pre-treatment trends are parallel, and to allow for the possibility that the treatment effect emerges gradually during the treatment period, we also use a "leads-and-lags" specification, in which we estimate a separate "treatment effect" for each year, before, during, and after the experiment period, and plot the annual coefficients and 95% confidence intervals (CIs) in leads-and-lags graphs. We use the period from 1998 (base year) through 2010; the plots run from 1999-2010. The specification with our sample periods is:

$$y_{it} = \gamma_t + f_i + \sum_{k=1999}^{2010} (\beta^k * D_i^k) + \varepsilon_{izt} \quad (5)$$

$D_i^k = 1$ for pilot firms in year *k*, and 0 otherwise. Thus, in contrast to $Pilot_i$*$During_t$ in eqn. (3), which remains on for pilot firms for the entire experiment period, the D_i^k's turn on for pilot firms only for one specific year, then off again. Therefore, β^{2005} provides the estimated effect for 2005 (first experiment year), β^{2006} provides the effect for 2006, and so on. We adjust the β^k coefficients by subtracting β^{2004} (a year before the experiment starts) from each, so that the *reported* $\beta^{2004} = 0$.

Firm and Year Fixed Effects (FE): FHK do not use firm FE and use year FE only in one specification (in the others, they include During and Post dummies). We prefer to use firm and year FE, as is common in DiD studies, but will also use the FHK specification. Unless specified otherwise, all our regressions use robust standard errors clustered on firm.

Covariates: FHK include the following covariates in their regressions: *ln*(assets); ROA; leverage (total debt/assets); and market-to-book ratio, all winsorized at 1% and 99%.[11] We prefer a

[11] The "note" to FHK's summary statistics table (Table I) states "All variables are winsorized at the 1% and 99% levels. We assume that they also winsorize the covariates that enter regressions, including those used to estimate

specification without these covariates. Given that we start with a randomized trial, covariates are not needed to address omitted variable bias. Including them should have only minor effects on point estimates, and can affect standard errors if the covariates are powerful predictors of the outcome, which we do not expect here.

Moreover, one should not include covariates that could potentially be affected by the treatment. Doing so can introduce bias, if these covariates also predict the outcome. Given the broad range of outcomes that researchers have argued could be affected by the short-sale experiment, it is hard to see what covariates could be both useful and not also potential outcomes of treatment. For example, Grullon, Michenaud and Weston (2015) argue that the short-sale experiment affected capital raising (and thus assets and leverage) and share price (and thus market-to-book ratio), and FHK argue that the experiment affected abnormal accruals, and thus ROA. Nonetheless, for a closer match to FHK, we will report a specification with their covariates, in which we add $\boldsymbol{\lambda}^{*}\mathbf{x}_{i,t}$ to eqns. (3)-(4), where $\mathbf{x}_{i,t}$ is a vector of covariates (for each i, t) and $\boldsymbol{\lambda}$ is a coefficient vector.

Firms Used to Compute AA and PMDA: Following Kothari et al. (2005), we compute AA and PMDA using all firms on Compustat. FHK do not specify which firms they use to compute PMDA, but state that they follow Kothari et al. (2005). Thus, they likely include firms not in the R3000 as well in computing abnormal accruals. This could explain why they report that both pilot and control firms have negative PMDA in both the pre and experiment periods (their Figure 2). Otherwise, average PMDA for all firms must be close to zero by construction of this measure.

F-score: FHK also study the effect of the SEC experiment on the likelihood of future misstatement of earnings (measured using the F-score, developed by Dechow et al. 2011) for pilot firms relative to control firms. Dechow et al. (2011), develop this measure using a variety of firm characteristics to predict the likelihood of a material misstatement. Dechow et al. develop three summary measures of the likelihood of a future restatement (we refer to them as F1, F2 and F3) using a

PMDA. They do not state whether they winsorize PMDA, once computed. Below, we assume that they did, but we also expect (and confirm) that if one winsorizes covariates, also winsorizing PMDA should make little difference.

sample of firms subject to the SEC's Accounting and Auditing Enforcement Releases. The first measure (F1) is based on variables from firm financial statements such as operating performance and accruals. The second measure (F2) adds off-balance sheet items such as operating leases and non-financial measures such as change in the number of employees. The third measure (F3) adds market-related variables such as market-adjusted stock returns.

FHK do not study these F-score measures directly. They instead create a binary variable (HF) that equals 1 if the firm's F-score in the top 1% of F-scores in their sample and study the resulting HF1, HF2, and HF3 measures. They do not explain this choice nor do they assess the robustness to other thresholds for the binary variable. FHK report that the coefficient on Pilot*During is significantly negative (HF of pilot firms declines during the experiment period) for all three measures, both with and without covariates. They do not, however, find evidence of a sign reversal after the experiment ends. We seek below to replicate their results. We will report results using the original F-score, the FHK variant, and variants with less extreme thresholds.[12] Our DiD specification for the F-score will be the same as above with the only change being to the dependent variables and as mentioned earlier, the sample size will be a bit smaller than for the accruals analysis, due to the need for data on all elements of the F-score.

III. Results

A. Summary Statistics, Covariate Balance, and Sample Attrition

Panels A, B and C, of Table I report variable definitions, summary statistics, and covariate balance for fiscal year 2004 for the 2005 Financial Analysis Sample, respectively. Panel B of Table I compares pilot firms to original control firms along a wide range of variables including our four accruals measures, the covariates that FHK use, prior performance, trading volume and other firm characteristics during 2004, for the unbalanced panel. Pilot and control firms are similar, as expected given the initial randomization. In the Internet Appendix, we confirm covariate balance for the balanced panel as well. In their Table I, FHK show balance on the covariates that enter estimation of accruals and their regressions.

[12] Our pre-analysis plan specified only the original F-score measures.

We also confirm that pre-treatment trends are parallel for covariates and outcomes from 1999-2004. This is also expected given the underlying randomization. A threat to validity for any causal design is differential attrition, potentially caused by the treatment. It seems unlikely that the removal of short-sale restrictions was powerful enough to affect attrition. In the pre-analysis plan, we confirm similar attrition for the pilot and control firms for the firms in the 2005 Financial Analysis Sample during the experiment period. [13]

B. Graphical Evidence

Given the initial randomization and no evidence of differential attrition, we start with a simple comparison of means. Our Figure 1 is comparable to FHK Figure 2, and presents univariate means during the Pre, During, and Post periods, separately for the pilot and the control firms. The top panel copies their Figure 2. The other panels present our results for each accruals measure for the unbalanced panel (left graphs) and the balanced panel (right graphs). The means are averaged over the respective periods. For all four accruals measures, pilot and control firms follow similar paths, and those paths are not at all similar to FHK's Figure 2.

In Figure 2, we also provide leads-and-lags graphs for all four accruals measures, which show annual coefficient estimates on Pilot*year, following eqn. (5). Small vertical lines around each point estimate show 95% confidence intervals (CIs). Larger vertical lines indicate the start and end of the experiment period. Across all the accruals measures, pre-treatment trends are reasonably parallel, as expected given the initial randomization. For all measures, there is no evidence of a treatment effect during 2005-2007, nor of a reversal after that.

These initial results provide strong evidence against the FHK story. With a simple specification, across four accruals measures, with both balanced and unbalanced panels, there is no evidence that pilot firms reduced accruals during the experiment period, relative to control firms, nor evidence for a rebound toward parity with control firms when the experiment ended.

[13] Attrition over the two-year experiment period is 13.5% for pilot firms versus 14.7% for control firms (z-statistic for difference in attrition rates = 0.73).

C. Regression Evidence

Table II presents results with our preferred specification, with firm and year FE and without covariates, for both the unbalanced and balanced panels, for all four accruals measures: operating accruals; total accruals; abnormal accruals measured using the modified Jones model (AA); and PMDA (AA, less AA for the matched firm).[14] For the first three measures, across both unbalanced and balanced panels, the coefficients on Pilot*During are economically modest, relative to the absolute values of the 2004 sample means of around 0.06 for operating accruals, 0.07 for total accruals, 0.05 for AA, and 0.04 for PMDA.[15] None are close to being statistically significant. The PMDA coefficient is larger in magnitude, but has the wrong sign (positive) relative to the FHK's prediction and result. Coefficient signs are mixed more generally: they are positive (opposite from the sign predicted by FHK) for the balanced panel for Total Accruals and PMDA.[16]

---

[14] The operating accruals, total accruals, and PMDA measures were pre-specified. We added abnormal accruals (without matching), during the course of our work, as part of an effort to understand why we could not come replicate the FHK results for PMDA.

[15] The mean for AA is not close to zero because we used the Compustat universe as the peer group for the industry-year estimation of normal accruals, rather than using only our sample firms. If we restrict the peer group to firms in our unbalanced panel, the sample mean of AA becomes 0.0015 for 2004 and 0.0008 across all sample years, very close to zero, as expected.

[16] In preparing Table II, we realized that we had underspecified exactly how we would winsorize when computing AA and PMDA. There are two plausible choices, both used in the accruals literature: (i) winsorize the dependent and independent variables used to estimate AA within each year; or (ii) winsorize these variables across the entire sample period. We chose to winsorize when computing eqn. (3) within each year, following Armstrong et al. (2013), and the related code posted by Dan Taylor at http://acct.wharton.upenn.edu/~dtayl/code.htm (SAS code at http://acct.wharton.upenn.edu/~dtayl/DACCs.sas). We chose this approach because it makes it less likely that AA and PMDA will be influenced by outliers, and also because this way, AA_t and $PMDA_t$ do not depend on which other years are included in the sample period. In particular, AA_t and $PMDA_t$ will not depend on whether one drops 2004 from the sample period, as FHK do in their preferred specification.

Hribar and Nichols (2007, App. B) state that they winsorize the variables used to estimate AA and PMDA at 1%/99%, but do not specify whether they do so by year or across their sample period. Jackson (2018) also does not specify whether he winsorized by year versus across the full sample period. He states that "I winsorize all variables before estimating discretionary accruals, and then subsequently winsorize [discretionary accruals] at the 1st and 99th percentiles." Joost Impink of University of Florida has posted SAS code for PMDA at: https://gist.github.com/JoostImpink/8ce0af0a0a0bbb31c8e0. His default winsorization is across the full sample period, but his code allows winsorizing by year as an option.

Note that we still winsorize continuous covariates across the entire sample period when estimating eqn. (4). In the Internet Appendix, we explore with care what difference the choice of winsorization strategy makes.

We also find no evidence of the sign reversal that FHK report (negative coefficient on Pilot*During, then near zero coefficient on Pilot*Post). In the fourth row of Table II, we show the *change* in the coefficient from the During to the Post period. This change is never meaningfully positive and, for three of the four measures (total accruals, AA, and PMDA), the coefficient is negative, opposite from the FHK prediction and finding.

We noted above that FHK's choice of PMDA as their sole accruals measure was surprising, given that (i) the randomization ensures that pilot and control firms were balanced prior to the experiment, on performance and all else; and (ii) as noted above, PMDA typically have lower power to detect earnings management than simpler measures. Table II also reports the standard errors for Pilot*During; PMDA has the highest standard error. Since operating accruals are an unbiased measure of earnings management (due to the randomization), and are more precisely estimated than other measures, if one wanted to use a single accruals measure, operating accruals would be the most preferred, and PMDA the least preferred. For similar reasons, AA should be preferred over PMDA, since AA, without matching, is unbiased and less noisy.

D. Moving Toward the FHK Specification

In Tables III-VII, we methodically change from our preferred specification to the FHK specification, one step at a time, to see whether we can come close to replicating their results. We fail. For operating accruals, the negative coefficient on Pilot*During in Table II increases in magnitude from |0.003| in Table II to |0.007| in Table VI and becomes close to marginally significant with covariates (t = -1.60). The coefficients on Pilot*During for the other three accruals measures remain insignificant, and the coefficient for PMDA remains positive, although close to zero. For the two measures with negative coefficients on Pilot*During (operating accruals and AA), there is no evidence of a sign reversal in the Post period. The FHK combined results – lower accruals for Pilot firms during the experiment period, and a reversal in the post-experiment period, cannot be replicated. Below, we discuss each individual step.

1. Adding Covariates

In Table III we add the four FHK covariates (*ln*(assets), ROA, leverage, and market-to-book ratio) as independent variables. We again present results for all four accruals measures, and

both unbalanced and balanced panels. As noted above, we would prefer not to use these covariates because they are possible outcomes of the short-sale experiment, especially ROA, which is directly affected by accruals. For the other covariates, given our prior that large effects of the experiment are unlikely, there is little harm from adding covariates, apart from the potential for specification search in the choice of covariates.

In practice, adding these covariates makes little difference. Most of the coefficients on the two interaction variables remain very close to those in Table II; and all remain insignificant. The largest change is an increase (opposite from the FHK prediction) in the coefficient on Pilot*During, for total accruals with the balanced panel, from 0.003 to 0.008. There is still no evidence of a reversal in relative accruals after the experiment ends. As expected, there is a strong positive association between ROA and accruals.

2. Using FHK Sample Periods

In Table IV and later tables, we present results only for the balanced panel, which is FHK's preferred specification. In Table IV, we switch to the FHK sample periods, including dropping 2004 from the Pre period and using calendar years. We present results both with and without covariates, for all four accruals measures. Once again, we find no support for lower accruals for the pilot firms during the experiment period, nor for a reversal after the experiment ends. For their preferred PMDA measure, the coefficient on Pilot*During, which was +0.009 with our preferred sample periods, drops substantially to almost zero, but remains positive (opposite from their prediction), at +0.0007. The coefficient on Pilot*During also drops meaningfully for the other accruals measures when we use the FHK sample periods. Results continue to be very similar with and without covariates.

3. Removing Firm FE

In Table V, we move yet closer to the FHK specification, by dropping firm FE and adding Pilot dummy, which would otherwise be absorbed by firm FE. Using firm and year FE is a standard panel data design, which is routinely used in DiD studies. Still, since we begin with a randomized experiment, the pilot and control firms should have similar firm effects, just as for any other covariate, and estimates should be unbiased with or without firm FE.

When we remove firm FE, very little changes. Coefficients are similar and standard errors are slightly smaller. The drop in operating accruals for pilot firms during the experiment period becomes close to marginally significant with covariates (t = -1.60). But the coefficients for the other accruals measures remain far from statistical significance, and total accruals and PMDA have the wrong sign. There is still no evidence for a sign reversal after the experiment ends.

4. Removing Year FE

FHK depart from the norm in panel data studies by using neither firm nor year FE. They use Pre, During, and Post dummies instead of year FE, although they do show that their results are similar with year FE. In Table VI, we drop year FE and add Pre, During, and Post dummies, which would otherwise be absorbed by the year effects. The results are essentially identical to those with Table V with year FE. Inference does not change.

E. Hitting Pause and Re-Checking Our Results

At this point in our replication, we hit "pause." We expected the FHK results to be sensitive to their choice of PMDA to measure accruals, and to their other design choices, including use of balanced versus unbalanced panel, choice of sample periods, using covariates, and not using either firm or year FE. We found some sensitivity when we dropped 2004, in the direction we expected – coefficients on Pilot*During drop substantially. But the coefficient estimates are not sensitive to their other design choices.

However, our coefficient estimates remain far from theirs. For Pilot*During, our estimate for PMDA is positive and near zero (+0.0005), and not close to significant (t = 0.05); theirs was -0.010 with much smaller standard errors than we find (they report s.e. of 0.004, while we find s.e. of 0.104), leading to their t-statistic around 2.5. For sign reversal, our estimate for PMDA is near zero at -0.002, versus their estimate of +0.013. These are large differences. We return below to the difference in standard errors.

We sought to see if we could understand why our coefficient estimates were so different from theirs. Differences in sample size are one possible explanation but seemed insufficient. Two coauthors wrote statistical code independently and confirmed that they obtained similar results.

We also compared our AA and PMDA code to the code publicly posted by Dan Taylor and Joost Impink. It is unlikely that a coding error can explain the differences in coefficients.

FHK do not state clearly what winsorization decisions they made in estimating PMDA, so we explored whether differences between our winsorization decisions and theirs could explain the differing results. In the Internet Appendix, we present results with a variety of winsorization choices: (i) our "full winsorization" approach, in which we winsorize variables used in estimating accruals and then again in estimating the coefficients on Pilot*During and Pilot*post; (ii) winsorize only in the second-stage regression but not the first-stage estimation of AA and PMDA; (iii) winsorizing when estimating AA and PMDA, but not in the second-stage regression and (iv) not winsorizing at either stage. We then use the same four alternatives, but estimate equation (1) using only the firms in the 2005 Financial Analysis Sample, instead of all firms in Compustat. Call these choices (v)-(viii). With choice (ii), the coefficient on Pilot*During for PMDA becomes large and positive (opposite from predicted), at +0.114 with covariates. With choice (iii), this coefficient is positive but smaller at +0.006. With choice (iv) the coefficient increases further to 0.288 and become statistically significant. If we estimate accruals using only the firms in our sample, the Pilot*During and sign reversal coefficients have the same signs as in FHK, at [-0.005, +0.003], [-0.006, +0.004], [-0.003, +0.003], and [-0.003, +0.032] for choices (v)-(viii), respectively, although none are close to being statistically significant. So perhaps one of these choices is closer to what FHK did. However, we remain far from their estimates of [-0.010, +0.013].

This effort makes it apparent that small variations in winsorization choices when computing AA, and in choosing the population of peer group firms used to estimate AA, can have a large effect on coefficient estimates. Yet a result which is not robust to minimal winsorization at 1% and 99%, and variations in exactly how one winsorizes, is not reliable evidence of a causal effect.

As a further robustness check (results in the Internet Appendix), we used a "doubly balanced" panel, in which we further limit the balanced panel for accruals to firms which have F-score available for all sample years. This shrinks the balanced panel from 465 pilot firms and 856 control firms to 419 pilot and 763 control firms. With this "doubly balanced" panel, the coefficient

on PMDA becomes slightly negative at -0.002 (without covariates), but remains far from the FHK coefficient, and statistically insignificant.

F. Two-Way Clustering and Standard Errors

In Table VII, we return to our full winsorization choice (i), and switch from standard errors clustered on firm to the FHK approach of two-way clustering on both firm and year. This makes a huge difference. Standard errors for the coefficient on Pilot*During for PMDA plummet from 0.0104 to 0.0048, close to FHK's reported s.e. of 0.004. Across the other accruals measures and both interaction variables, s.e.'s with two-way clustering are sometimes similar, sometimes somewhat larger, but sometimes much lower. It is not apparent to us why two-way and one-way clustered standard errors are sometimes similar and sometimes far apart, and why differences are found for some accruals measures and not others. Nor is it apparent why one-way and two-way clustered standard errors can be very different for Pilot*During, yet similar for Pilot*Post (as is the case for PMDA), or similar for Pilot*During and very different for Pilot*Post (as is the case for operating accruals). A further sign of the unpredictability of two-way clustered standard errors: for PMDA, our two-way clustered standard errors are close to FHK for Pilot*During, but far apart for Pilot *Post (our s.e.'s are 0.0097, they report 0.004).

Clustered standard errors are correct only asymptotically. As we wrote in our pre-specified design: "With one-way clustering, clustered standard errors with a small number of clusters are downward biased (e.g., Cameron, Gelbach and Miller, 2008). A similar concern applies to the second clustering dimension. We are aware of only limited research, principally Kezdi (2004), on the importance of this potential bias for the second clustering dimension (here, by year) if one has a large number of clusters for the first clustering dimension (here, by firm)." Table VII provides evidence that this bias can be important for panel datasets of the form often used in finance and accounting research, with firm*year data, many firms, but a short time dimension. Moreover, with a small-number-of-clusters, behavior of two-way clustered standard errors is unpredictable – with large differences from one-way clustering found for one dependent variable but not another; and within a single regression, found for one predictor variable but not another.

Two-way clustered standard errors are used with some frequency in the finance and accounting literatures. We are not aware of prior discussion of the touchiness of two-way clustered standard errors with a short time dimension. Similar to the advice from Angrist and Pischke (2008 § 8.1) that one should report the larger of ordinary or robust standard errors when one has a small number of observations, a sensible, conservative choice for standard errors, if two-way clustering is a plausible choice, is to report the larger of the one-way or two-way clustered standard errors, and to do so variable by variable and regression by regression. It is very likely that if FHK had either adopted this approach, or used one-way clustering on firm, none of their results would have been close to being statistically significant, even if their coefficient estimates were reproducible and robust.

Another way to think about the issue of which standard errors are believable: Consider the annual coefficient estimates and 95% CIs for the PMDA measure in Figure 2. These CIs are quite wide; for 2005, the CI is [-0.046, +0.023]; the corresponding s.e. is 0.0176. Averaging over multiple years can increase precision at the usual $n^{0.5}$ rate. Thus the s.e. for the 3-years covered by the Pilot*During variable should be around $0.0176/3^{0.5} = 0.0102$. This is close to what we find with standard errors clustered on firm (0.0104 in our Table VII). However, there is no magic way to get from one-year s.e.'s of around 0.0176 to the s.e. of 0.048 in Table VII (or the 0.04 reported by FHK) for a three year period. In this setting, the much lower, two-way clustered s.e.'s are simply wrong.[17]

IV. Results for F-Score and the FHK "HF" Measure

Having found no evidence of lower accruals for pilot versus control firms, either during or after the Reg SHO experiment (the central FHK results), we seek in this section to replicate another FHK result, also related to earnings management. FHK report that during the experiment period,

[17] A further problem with standard errors for AA and PMDA is violation of the assumption of independent errors. Mean AA for all firms in an industry-year is zero by construction; this creates dependence across these firms. Non-independence is much more acute for PMDA, in which firm 1 is matched with firm 2, and vice versa. This induces dependence between the PMDA's for the two firms, and thus to downward biased standard errors, if both firm 1 and 2 are in the sample. If results had been statistically significant, we would have wanted to address this violation of independence by using applying randomization inference methods to estimate confidence intervals.

a measure of the likelihood of a material misstatement declined for pilot firms relative to control firms (using the binary version of the F-score measure explained earlier).

A. F-score

We begin our reconsideration of the FHK results for F-score by studying the original F-score, using a specification similar to Table II: both unbalanced and balanced panels; no covariates; with firm and year FE.[18] We then systematically move towards the FHK specification, much as we did for accruals. To conserve space, we present in the text only results for the FHK specification: balanced panel, FHK sample periods and no firm or year FE. We tabulate and discuss only two specifications: one with the original F-score and one using the HF binary variable created by FHK; the Internet Appendix contains additional results.[19]

As for our analysis of accruals, we start with graphical evidence. Figure 3 is similar in structure to Figure 2. It provides univariate means for the Pre, During, and Post periods, separately for pilot and control firms, for a balanced sample with data on all three F-scores available for all sample years, using the FHK sample periods. Panel A provides means for the three F-scores, and Panel B provides means for the three HF measures. In the Internet Appendix, we report the leads-and-lags graphs for F-score, again using the balanced panel and the FHK specification. In both Figure 3, Panel A and the leads-and-lags graphs, there is no evidence of a relative drop in F-score during the experiment period. Figure 3, Panel B is similar to Panel A, but shows results for the HF measures. There is a relative drop in the HF-1 and HF-2 measures, but not in HF-3. However, this relative drop is driven entirely by a modest rise for control firms; there is no drop in the likelihood that pilot firms exceed the HF measure threshold.

[18] Following FHK, we require a balanced panel based on F-scores and covariates (sample firms must have data available to calculate F-scores and covariates over the entire FHK sample period). Note that this "F-score balanced panel" is slightly different from the "accruals balanced panel" that we used to study accruals, because the F-score balanced panel requires that firms have complete data on F-score (but not necessarily accruals) and the accruals balanced panel requires that firms have complete data on accruals (but not necessarily F-score).

[19] For the HF measure, firm FE cannot be combined with the probit estimation used by FHK due to the incidental parameters problem.

Table VIII reports the results using F-score as the dependent variable. Across all three F-score measures, both with and without covariates, the coefficients on Pilot*During have varying signs and are not close to being statistically significant. The same is true using our specification (see Internet Appendix). Thus, we fail to observe a reduction in F-score for the pilot firms during the experiment period.

B. Binary HF Score: Balanced Panel, FHK Sample Periods and Specification

FHK do not report results using F-score but rather use a variant of the F-score, a binary variable which they call HF. The HF variable reflects an unexplained choice, to turn a continuous variable into a binary equivalent (by itself not unusual), and to do so using an extreme threshold (F-score within top 1% of all F-scores).

Nonetheless, we use their approach to create the binary HF variable. In Table IX, we report results using their specification, except that we use standard errors clustered on firm, where they use two-way clustered standard errors. We follow FHK in using probit regressions but report average marginal effects (FHK report probit regression coefficients but discuss some marginal effects in the text). As was the case for accruals, the coefficient on Pilot*During is not significant for any of the specifications. We come close to replicating their probit coefficients for HF-1 and HF-2 (FHK report a -0.4% marginal effect for HF-1, versus our -0.5%), but the opposite sign for HF-3.

C. Varying the Binary Variable Threshold

In Figure 4, we use the FHK approach to create a binary HF-style variable which turns on at different thresholds: 99%, 97.5%, 95%, 90% and 80%. We graphically report the results for these thresholds, for HF-1, for which we find an average marginal effect similar to theirs (although they find statistical significance and we do not). Thresholds are shown on the x-axis and average marginal effects for specific thresholds are shown as circles. Dashed lines show the upper and lower bounds of the 95% CIs around the coefficient estimates. The coefficient on Pilot*During is

not significant for any of the thresholds. Moreover, the average marginal effects are positive (opposite from the FHK prediction) for all thresholds except the 99% threshold.[20]

D. One-Way versus Two-Way Clustered Standard Errors

Lastly, we assess the impact of FHK's choice of two-way clustering on firm and year, versus clustering only on firm, on standard errors for their HF measure. In Table X, we show standard errors for the HF score using both clustering choices for probit coefficients (note, in Table IX, where we report average marginal effects). As was the case with PMDA for Pilot*During, but this time in more extreme fashion, across all three HF measures, and both Pilot*During and Pilot*Post, two-way clustering produces much lower standard errors. The HF-1 and HF-2 coefficients become significant, with and without covariates.[21]

As was the case for the PMDA accruals measure, the s.e.'s with two-way clustering are too low. Annual s.e.'s for the probit coefficients are around 0.25. A reduction by a factor of $3^{0.5}$, to reflect three years of data for the experiment period, should produce s.e.'s around 0.14, close to what we find with one-way clustering. Two-way clustering cannot dramatically improve on the usual $n^{1/2}$ reduction in s.e.'s, relative to one-year s.e.'s, where n is the number of years in the treatment period. Yet for HF-1, two-way clustered s.e.'s are over 4 times lower than one-way clustered s.e.'s without covariates, and 8 times lower with covariates.[22] This supports our

[20] FHK apparently defined their HF measure based on the distribution of F-score over their entire sample period, using values for both pilot and control firms. We believe that a preferable approach, would be to define the threshold annually (or looking back but not forward) to avoid look ahead bias, and to define the threshold using only control firms. In the Internet Appendix, we provide figures similar to text Figure 4, but with varying annual thresholds for HF-1 and HF-2, defined using only control firms. Results are similar to Figure 4.

[21] A further problem with standard errors for the HF measures is violation of the independence assumption underlying standard error computation. If, say, pilot firm 1 reduces its F-score and thus has HF = 0 instead of 1, some other firm with a similar F-core will now have HF = 1, by construction of the HF measure; this creates dependence across firms with close-to-threshold scores. If results for HF had been statistically significant, we would have wanted to address this violation of independence by applying randomization inference methods to estimate confidence intervals.

[22] The two-way clustered standard errors for the HF measures are also highly sensitive to small changes in sample specification. In the Internet Appendix, we reproduce Table IX, making only one small change – we use the "accruals balanced panel" instead of the "F-score balanced panel. This minor change has a huge effect on the two-way clustered s.e.'s. For the HF-1measure with covariates, the 2-way clustered s.e. with the F-score balanced panel is .0209 (Table X), versus .0788 with the accruals balanced panel in Table A-VII, panel B in Internet Appendix. The std. error with clustering only on firm is 0.1680 in Table X). By contrast, s.e.'s clustered on firm are stable, across samples, both

recommendation above that researchers, if they consider two-way clustering, should report the larger of one-way or two-way clustered errors.

E. Summary for F-Score and HF Measures

Overall, we fail to replicate FHK's result that the probability of material misstatement is lower for Pilot firms during the experiment period, for either F-score or their HF measure. Even apart from their HF measure perhaps reflecting specification search, and our failure to replicate their results for HF-2 and HF-3, all of their results would be insignificant if they had used standard errors clustered only on firm. These null results are consistent with the null results for the accruals measures.

V. Implications from Our Reanalysis of FHK

In this paper we tried to replicate/reassess two key findings in Fang, Huang and Karpoff (2016). Specifically, we reassessed FHK's findings on the impact of the Reg SHO experiment on earnings management by the pilot firms. Their accruals results cannot be replicated.[23] Across four accruals measures, including theirs, we find no evidence of significantly lower accruals for pilot firms during the experiment period, nor evidence for a post-experiment reversal. There is also no evidence of a drop in F-score for pilot firms during the experiment period. For their binary HF measure, we can replicate their coefficients for HF-1 and HF-2 (although not HF-3). But the HF-1 and HF-2 coefficients are statistically significant only with two-way clustering, and would not be close to statistical significant if one uses standard errors clustered on firm. Moreover, the results for HF measure are highly sensitive to specification (HF=1, if F-score is in the 99th percentile or higher): we find no evidence of relative change for the underlying F-score, nor for binary measures using less extreme thresholds such as 97.5%, 95%, etc.

---

with and without covariates.

[23] See also Black, Litvak and Yoo (2019), who reassess two initial findings in Grullon, Michenaud and Weston (2015) on the causal channel – running from short interest to effect on share price to indirect effects – but cannot replicate either of these findings. They find neither: (i) an increase in open short interest for smaller Pilot firms during the period from experiment announcement to launch; and (ii) a fall in share prices for smaller Pilot firms during the two weeks before the experiment was announced. Neither of these results is replicable.

Our findings have implications for other studies of the short-sale experiment. Similar to FHK, several recent studies rely on a managerial "fear" channel to explain firm actions in response to the Reg SHO experiment. Lack of robustness for one set of outcomes (reduced earnings management), which the authors justify based on a presumed fear channel provides at least some reason to doubt the robustness of other outcomes attributed to the managerial fear channel.

Our results suggest that two-way clustering on firm and year requires great care for panels with a short time dimension, because two-way clustering can substantially reduce standard errors and hence inflate the apparent statistical significance of estimates. We recommend that authors who consider using two-way clustering should report both one-way or two-way clustered standard errors., and rely on the larger of the two for inference

An important takeaway from our reanalysis is that even when researchers begin with a randomized trial, they must make many design decisions. For the FHK research question, those include: defining the sample (including choice of balanced versus unbalanced panel); choosing the sample periods; deciding which outcomes to study; using covariates or not (and which ones); handling outliers; and choosing how to compute standard errors. These decisions offer opportunities for specification search or "lucky" choices to produce false positive results.

References

Alexander, Gordon J. and Mark A. Peterson (2008), The Effect of Price Tests on Trader Behavior and Market Quality: An Analysis of Regulation SHO, *Journal of Financial Markets* 11, 84-111.

Angrist, Joshua D. and Jörn-Steffen Pischke (2008), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton university press.

Armstrong, Chris, David F. Larcker, Gaizka Ormazabal and Daniel J. Taylor (2013), The Relation Between Equity Incentives and Misreporting: The Role of Risk-Taking Incentives, *Journal of Financial Economics* 109: 327-350.

Black, Bernard, Hemang Desai, Kate Litvak, Woongsun Yoo, and Jeff (Jiewei) Yu (2019), Pre-Analysis Plan for the Reg SHO Reanalysis Project, working paper, at http://ssrn.com/abstract=3415529).

Black, Bernard, Alex Hollingsworth, Leticia Nunes and Kosali Simon (2019), The Effect of Health Insurance on Mortality: Statistical Power and What We Can Learn from the Affordable Care Act Coverage Expansions, NBER working paper 25,568, at http://ssrn.com/abstract=3336520.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller (2008), Bootstrap-Based Improvements for Inference with Clustered Errors, *Review of Economics and Statistics* 90, 414-427.

Dechow, Patricia M., Weili Ge, Chad R. Larson, and Richard G. Sloan (2011), Predicting Material Accounting Misstatements, *Contemporary Accounting Research* 28, 17-82.

Diether, Karl, Kuan-Hui Lee, and Ingrid Werner (2009), It's SHO Time! Short-Sale Price-Tests and Market Quality, *Journal of Finance* 64, 37-73.

Fang, Vivien W., Allen Huang, and Jonathan Karpoff (2016), Short Selling and Earnings Management: A Controlled Experiment, *Journal of Finance* 71, 1251-1293.

Grullon, Gustavo, Sebastien Michenaud, and James Weston (2015), The Real Effects of Short-Selling Constraints, *Review of Financial Studies* 28, 1737-1767.

Harvey, Campbell R. (2019), Replication in Financial Economics, working paper, at http://ssrn.com/abstract=3409466 .

Healy, Paul M. (1985), The Effect of Bonus Schemes on Accounting Decisions, *Journal of Accounting and Economics* 7, 85-107.

Hope, Ole-Kristian, Danqi Hu, and Wuyang Zhao (2016), Third-Party Consequences of Short-Selling Threats: The Case of Auditor Behavior, *Journal of Accounting and Economics* 63, 479-498.

Hribar, Paul, and D. Craig Nichols (2007), The Use of Unsigned Earnings Quality Measures in

Tests of Earnings Management, *Journal of Accounting Research* 45, 1017-1053.

Jackson, Andrew B. (2018), Discretionary Accruals: Earnings Management or Not?, *Abacus* 54, 136-153.

Kezdi, Gabor (2004), Robust Standard Error Estimation in Fixed-Effects Panel Models, *Hungarian Statistical Review* 9, 95-116.

Kothari, Sagar P., Andrew J. Leone, and Charles E. Wasley (2005), Performance Matched Discretionary Accrual Measures, *Journal of Accounting and Economics* 39, 163-197.

Litvak, Kate, Bernard Black, and Woongsun Yoo (2019), The SEC's Busted Randomized Experiment: What Can and Cannot Be Learned, working paper, at http://ssrn.com/abstract=2647418.

Richardson, Scott A., Richard G. Sloan, Mark T. Soliman, and Irem Tuna (2005), Accrual Reliability, Earnings Persistence and Stock Prices, *Journal of Accounting and Economics* 39, 437-485.

Securities and Exchange Commission, Office of Economic Analysis (2007), Economic Analysis of the Short Sale Price Restrictions under the Regulation SHO Pilot, at https://www.sec.gov/news/studies/2007/regshopilot020607.pdf.

Sloan, Richard G. (1996), Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings? *The Accounting Review* 71, 289-315.

Table I

Variables, Summary Statistics, and Pretreatment "Covariate Balance" for Accruals

Panel A: Variable Definitions

Balance sheet and income statement values are from Compustat Annual. Except as specified below, we use the most recent fiscal year ending before May 2005, and all variables are winsorized at 1% and 99%. Compustat variable names are indicated below.

Variables	Definitions
Assets ($M)	Total Assets (AT)
Sales ($M)	Net Sales (SALE).
Market Cap ($M)	Market Capitalization (PRC*SHROUT/1000). Based on price and shares outstanding from CRSP as of June 30, 2004 for 2004 Analysis Sample, and as of April 29, 2005 for 2005 Analysis Sample.
Tobin's *q*	Tobin's *q*, defined as total assets (AT) minus book value of common equity (CEQ) and deferred tax (TXDB), plus market capitalization, scaled by total assets. (AT-CEQ-TXDB+(PRCC_C*CSHO)/AT)
Short Interest (% of Shares Outstanding)	Average *% monthly open short interest* during 12 months from July 2003 to June 2004 (for 2004 Analysis Sample) or from May 2004 to April 2005 (for 2005 Analysis Sample), defined as monthly short interest reported on 15th of each month (from Compustat) scaled by shares outstanding at the start of the month (from CRSP)*100 (100*SHORTINT/(SHROUT*1000)).
Capex/Assets	Capital Expenditures scaled by Total Assets (CAPX/AT)
R&D/Sales	R&D scaled by Net Sales (XRD/SALE). Missing R&D is replaced with 0 and negative Net Sales are treated as missing. Winsorized at 1.
ROA	Return on Assets, defined as operating incoome before depreciation scaled by lagged total assets (OIBDP/AT).
Leverage	Total Debt/Total Assets ((DLC+DLTT)/(DLC+DLTT+SEQ))
Book/Market Ratio	Book-to-Market Ratio (CEQ/(CSHO*PRCC_F))
Trading Volume	Average *fractional trading volume* during 12 months from July 2003 to June 2004 (for 2004 Analysis Sample) or from May 2004 to April 2005 (for 2005 Analysis Sample), defined as monthly trading volume (from CRSP) scaled by shares outstanding at the end of the month (from CRSP) (100*VOL/(SHROUT*1000)). Winsorized at 99%.
Beta	Beta from regression of daily return (RET) on market value weighted return from CRSP (VWRETD) over 250 trading days preceding July 28, 2004 (for 2004 Analysis Sample) or May 2, 2005 (for 2005 Analysis Sample)
Share Returns	$\prod_i(1 + RET_i) - 1$, where *i* includes 12 months from July 2003 to June 2004 (for 12-month pre-announcement period) or 10 months from July 2004 to April 2005 (for 10 month period between experiment announcement and experiment launch)
Operating Accruals	Operating Accruals, defined as Earnings Before Extraordinary items and discontinued operations on the cash flow statement (IBC), minus operating cash flows (OANCF) before extraordinary items and discontinued operations (XIDOC), scaled by beginning-of-the-year total assets. ((IBC-(OANCF-XIDOC))/AT_{t-1})
Total Accruals	Total Accruals, defined as Earnings Before Extraordinary items and discontinued operations on the cash flow statement (IBC), minus operating cash flows (OANCF) before extraordinary items and discontinued operations (XIDOC), minus investing cash flow (IVNCF), scaled by beginning-of-the-year total assets. ((IBC-(OANCF-XIDOC)-IVNCF)/AT_{t-1})
Abnormal Accruals (AA)	Measured using the modified Jones model, as described in the text.
PMDA	Performance-matched discretionary accruals, measured as described in the text.

Panel B: Summary Statistics and Covariate Balance

Table provides summary statistics and evidence on pretreatment balance for the 2005 Financial Analysis Sample (unbalanced panel). See Internet Appendix for a similar table for a balanced panel drawn from this sample. Variables are defined in Panel A of Table I above. Except as stated in the variable definitions, (i) we assess balance using the most recent Compustat datadate before May 2005. All variables are winsorized at 1%/99% except as indicated in the Table. Table show mean and median for Pilot and original control firms, together with a two-sample *t*-statistic for difference in means, and a z-statistic from a rank-sum test for difference in medians. Sample for most variables is indicated in the table; sample is slightly smaller for some variables due to missing data on Compustat**.** *, **, *** indicates statistical significance at the 10%, 5%, and 1% levels; significant results at 5% or better are in **boldface**. Financial data is from Compustat; trading volume and share returns are from CRSP.

	Pilot Firms		**Original Controls**		**Norm. Diff.**	***t*-test for Means**	**Rank-sum test**
Number of firms	**702**		**1,413**				
	Mean	**Median**	**Mean**	**Median**			
ln(Assets, $M)	6.796	6.662	6.720	6.546	0.050	1.08	1.28
Assets, $M	3,215	782	3,136	697	0.010	0.22	1.28
ln(Sales, $M)	6.543	6.598	6.484	6.491	0.032	0.68	0.70
Sales, $M	3,049	731	2,944	657	0.015	0.32	0.78
ln(Market Cap, $M)	6.996	6.792	6.888	6.666	0.074	1.59	1.79*
Market Cap, $M	4,152	890	3,705	785	0.043	0.94	1.79*
Tobin's q	2.333	1.849	2.284	1.825	0.035	0.76	0.30
Short Interest (%) (05/2004-04/2005)	4.795	3.146	4.798	3.404	-0.001	-0.01	-0.19
Capex/Assets	0.048	0.033	0.048	0.031	-0.004	-0.09	0.41
R&D/Sales (win at 1.00)	0.099	0.005	0.100	0.009	-0.005	-0.12	-1.51
ROA	0.114	0.128	0.103	0.120	0.076	1.63	1.48
Leverage	0.283	0.240	0.278	0.232	0.017	0.36	0.82
Book/Market Ratio	0.403	0.382	0.400	0.361	0.011	0.23	0.30
Trading Volume (win at 99%)	0.193	0.155	0.198	0.158	-0.033	-0.71	-0.58
Beta	1.441	1.399	1.445	1.358	-0.007	-0.15	-0.01
Share Returns (07/2003-06/2004)	0.492	0.336	0.477	0.338	0.024	0.52	0.09
Share Returns (07/2004-04/2005)	-0.023	-0.041	-0.032	-0.036	0.026	0.57	0.07
Operating Accruals	-0.057	-0.050	-0.058	-0.052	0.010	0.20	0.86
Total Accruals	0.078	0.030	0.071	0.026	0.032	0.69	0.53
Abnormal Accruals (AA)	0.053	0.023	0.051	0.026	0.011	0.25	-0.20
PMDA	-0.032	-0.012	-0.044	-0.020	0.040	0.83	0.89

Panel C: Graphical Evidence on Covariate Balance

Figure provides a graphical overview of covariate balance for the unbalanced panel. It shows t-statistics for differences between pilot and original control firms, for the variables listed in Panel A of Table I above. Vertical lines indicate t-statistics of at -1.96, 0, and +1.96.

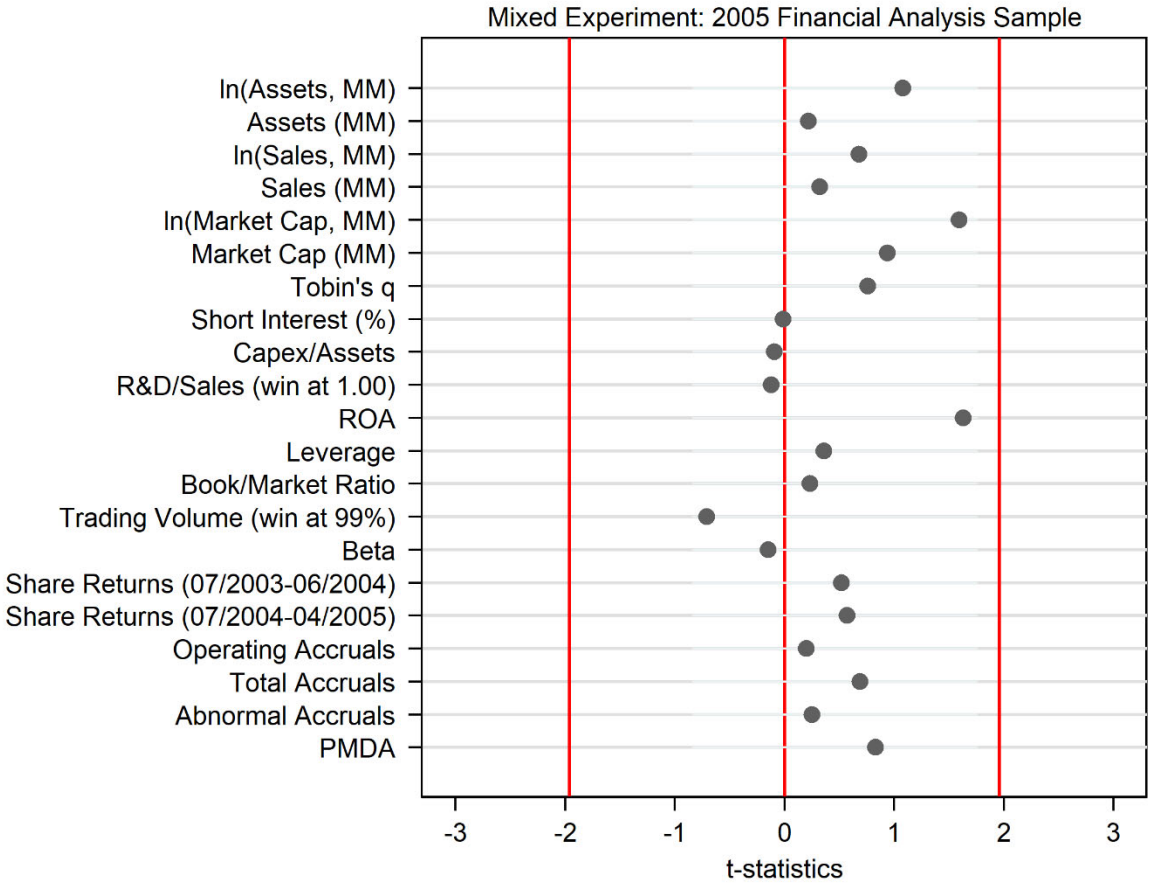


Table II

Our Preferred Specification for Accruals: Firm and Year FE, No Covariates

Unbalanced panel uses all firms in 2005 Financial Analysis Sample for Mixed Experiment with data to calculate accruals. Balanced panel further requires firms to have data to calculate accruals for each year in sample period. Top two rows show coefficients on interaction variables of interest, for indicated accruals measures, from regressions following eqn. (3), with firm and year fixed effects, over fiscal years 2001-2010 (our sample period). In models (1) and (5), the dependent variable is *Operating Accruals*, defined as Earnings Before Extraordinary items and discontinued operations (IBC), minus operating cash flows (OANCF) before extraordinary items and discontinued operations (XIDOC), scaled by beginning-of-the-year total assets. In models (2) and (6), the dependent variable is *Total Accruals,* calculated as IBC-(OANCF-XIDOC)-IVNCF)/$AT_{t-1,}$ where IVNCF denotes investing cash flow. Abnormal accruals (AA) in models (3) and (7) and PMDA in models (4) and (8) are computed using Modified Jones model following FHK. Pilot*During =1 for pilot firms during experiment period, 0 otherwise; Pilot*Post =1 for pilot firms during post-experiment period, 0 otherwise**.** Pre, During, and Post periods are specified in the text. Sign reversal row reports (coefficient on (Pilot*During) – coefficient on (Pilot*Post)), *t*-statistic is computed using eqn. (4). Accruals measures are winsorized at 1%/99%. *t*-statistics, using standard errors clustered on firm, in parentheses. Standard errors (*s.e.*) shown in brackets are for Pilot*During. "Sample mean in 2004" reports sample mean for each accruals measure for fiscal 2004.

		Unbalanced Panel				**Balanced Panel**			
	FHK sign	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Accruals type	**predict**	**Operating**	**Total**	**AA**	**PMDA**	**Operating**	**Total**	**AA**	**PMDA**
Pilot*During	negative	-0.0024	-0.0000	-0.0013	0.0024	-0.0036	0.0034	-0.0025	0.0082
		(-0.64)	(-0.00)	(-0.23)	(0.27)	(-0.98)	(0.43)	(-0.39)	(0.82)
s.e.		[0.0037]	[0.0081]	[0.0059]	[0.0092]	[0.0037]	[0.0080]	[0.0064]	[0.0100]
Pilot*Post	near zero	-0.0014	-0.0052	-0.0042	0.0002	-0.0032	-0.0054	-0.0059	0.0049
		(-0.34)	(-0.67)	(-0.64)	(0.02)	(-0.77)	(-0.72)	(-0.87)	(0.50)
Sign Reversal	positive	0.0010	-0.0049	-0.0027	-0.0024	0.0006	-0.0085	-0.0032	-0.0036
(Post – During)		(0.26)	(-0.59)	(-0.46)	(-0.26)	(0.16)	(-1.06)	(-0.56)	(-0.37)
Firm, Year FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm-Year Obs.		18,785		18,468		14,640		14,080	
Pilot (Control) Firms		702 (1413)		695 (1408)		517 (947)		491 (917)	
Adjusted R^2		19.8%	8.9%	15.9%	4.4%	21.1%	8.8%	16.4%	4.0%

Table III

Move toward FHK: Adding Their Covariates

Samples and specification are the same as Table II, except we include the FHK covariates as additional independent variables and require data on these covariates. Accruals measures are defined in Table II notes. *ln(assets)* is the natural logarithm of total assets (AT). *Market-to-book ratio* is the market value of equity (PRCC_F × CSHO) divided by the book value of equity (CEQ). *ROA* is return on assets, defined as operating income before depreciation (OIBDP) scaled by total assets at the beginning of the fiscal year. *Leverage* is long-term debt (DLTT) plus debt in current liabilities (DLC), scaled by the sum of DLTT, DLC and total shareholders' equity (SEQ). Top two rows show coefficients on interaction variables of interest, for indicated accruals measures, from regressions following eqn. (3), with firm and year fixed effects, over fiscal years 2001-2010. Accruals and covariates are winsorized at 1%/99%. *t*-statistics, using standard errors clustered on firm, are in parentheses. *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Significant results, at 5% level or better, in **boldface**.

Panel	**FHK sign predict**	**Unbalanced Panel**				**Balanced Panel**			
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Accruals type		**Operating**	**Total**	**AA**	**PMDA**	**Operating**	**Total**	**AA**	**PMDA**
*Pilot*During*	negative	-0.0028	-0.0002	-0.0016	0.0047	-0.0037	0.0078	-0.0026	0.0088
		(-0.77)	(-0.03)	(-0.28)	(0.51)	(-1.02)	(1.08)	(-0.40)	(0.87)
*Pilot*Post*	near zero	-0.0010	-0.0011	-0.0041	0.0005	-0.0023	-0.0014	-0.0055	0.0046
		(-0.25)	(-0.15)	(-0.64)	(0.05)	(-0.58)	(-0.19)	(-0.80)	(0.46)
Sign Reversal	positive	0.0018	-0.0009	-0.0025	-0.0042	0.0014	-0.0091	-0.0028	-0.0042
		(0.45)	(-0.12)	(-0.43)	(-0.45)	(0.37)	(-1.22)	(-0.47)	(-0.42)
ln(assets)		-0.0018	**0.0726*** **	**-0.0088** **	**-0.0224*** **	-0.0002	**0.0602*** **	**-0.0093** **	**-0.0271*** **
		(-0.59)	**(13.76)**	**(-2.08)**	**(-3.52)**	(-0.05)	**(11.79)**	**(-2.07)**	**(-3.92)**
market-to-book ratio		-0.0000	**0.0024*** **	0.0006	-0.0006	0.0002	**0.0024*** **	0.0001	-0.0003
		(-0.06)	**(3.93)**	(1.11)	(-0.75)	(0.45)	**(3.84)**	(0.13)	(-0.29)
ROA		**0.2213*** **	**0.5760*** **	**0.2441*** **	**0.0739** **	**0.1963*** **	**0.5790*** **	**0.2343*** **	0.0695*
		(13.18)	**(17.75)**	**(11.56)**	**(2.37)**	**(10.25)**	**(16.95)**	**(9.83)**	(1.82)
Leverage		**-0.0380*** **	0.0084	**-0.0360*** **	**-0.0397*** **	**-0.0356*** **	0.0131	**-0.0393*** **	**-0.0336** **
		(-5.28)	(0.69)	**(-3.82)**	**(-2.86)**	**(-4.35)**	(1.02)	**(-3.57)**	**(-2.08)**
Firm and Year FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm-Year Obs		18,454		18,156		13,720		13,250	
Pilot (Control) Firms		702 (1413)		695 (1408)		488 (884)		465 (860)	
Adjusted R^2		24.8%	20.8%	18.5%	4.8%	24.1%	20.8%	18.1%	3.9%

Table IV

Move toward FHK: Use only Balanced Panel, Switch to FHK Sample Period

In this and later tables, we use only a balanced panel, following FHK. Specification is same as Table II and III, except we use the FHK sample periods (calendar years 2001-2010, dropping 2004). For comparison, FHK's balanced panel includes 9,843 firm-year observations over the same period, for 1,097 firms (388 pilot and 709 control). Top two rows show coefficients on interaction variables of interest, for indicated accruals measures, from regressions following eqn. (3), with firm and year fixed effects. Accruals and covariates are winsorized at 1%/99%. See Table II and III notes for detailed variable definitions. *t*-statistics, using standard errors clustered on firm, are in parentheses. *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Significant results, at 5% level or better, in **boldface**.

Accruals Type	**Operating**		**Total**		**AA**		**PMDA**	
*Pilot * During*	-0.0066	-0.0059	0.0049	0.0090	-0.0076	-0.0066	0.0007	0.0007
	(-1.50)	(-1.38)	(0.57)	(1.16)	(-1.05)	(-0.94)	(0.07)	(0.06)
*Pilot * Post*	-0.0048	-0.0039	-0.0053	-0.0004	-0.0084	-0.0074	-0.0015	-0.0021
	(-1.02)	(-0.87)	(-0.64)	(-0.05)	(-1.09)	(-1.00)	(-0.14)	(-0.20)
Sign reversal	0.0018	0.0019	-0.0102	-0.0094	-0.0008	-0.0008	-0.0022	-0.0027
	(0.44)	(0.47)	(-1.24)	(-1.22)	(-0.12)	(-0.13)	(-0.21)	(-0.26)
ln(assets)		-0.0015		**0.0550*** **		-0.0090*		**-0.0300*** **
		(-0.44)		**(10.02)**		(-1.95)		**(-4.28)**
market-to-book		0.0007		**0.0029*** **		0.0005		0.0001
		(1.41)		**(3.89)**		(0.68)		(0.06)
ROA		**0.1875*** **		**0.5796*** **		**0.2420*** **		0.0558
		(9.52)		**(15.59)**		**(9.98)**		(1.47)
Leverage		**-0.0385*** **		0.0129		**-0.0452*** **		**-0.0333** **
		(-4.24)		(0.95)		**(-3.85)**		**(-2.06)**
Firm, Year FE	YES							
Firm-Year Obs.	11,889							
Pilot (Control) Firms	465 (856)							
Adjusted R^2	18.9%	23.1%	9.0%	20.9%	15.3%	18.1%	3.6%	4.0%

Table V

Move toward FHK: Replacing Firm FE with Pilot Dummy

Specification and sample is same as Table IV, except we drop firm fixed effects and add Pilot firm dummy (which would be absorbed by firm FE). *Pilot dummy* takes value 1 for pilot firms and zero for control firms. Accruals and covariates are winsorized at 1%/99%. See Table II and III notes for detailed variable definitions. *t*-statistics, using standard errors clustered on firm, are in parentheses. *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Significant results, at 5% level or better, in **boldface**.

Accruals Type	**Operating**		**Total**		**AA**		**PMDA**	
*Pilot * During*	-0.0066	-0.0065	0.0049	0.0074	-0.0076	-0.0071	0.0007	0.0005
	(-1.59)	(-1.60)	(0.60)	(0.97)	(-1.12)	(-1.06)	(0.07)	(0.05)
*Pilot * Post*	-0.0048	-0.0044	-0.0053	-0.0032	-0.0084	-0.0080	-0.0015	-0.0016
	(-1.08)	(-1.03)	(-0.67)	(-0.44)	(-1.15)	(-1.12)	(-0.15)	(-0.16)
Sign reversal	0.0018	0.0021	-0.0102	-0.0106	-0.0008	-0.0009	-0.0022	-0.0021
	(0.46)	(0.55)	(-1.32)	(-1.44)	(-0.13)	(-0.15)	(-0.22)	(-0.21)
ln(assets)		**0.0043*****		0.0006		**-0.0033****		-0.0002
		(4.71)		(0.52)		**(-2.50)**		(-0.16)
market-to-book		-0.000		**0.0023*****		**0.0012****		-0.0011
		(-0.00)		**(3.32)**		**(2.40)**		(-1.37)
ROA		**0.0718*****		**0.3411*****		**0.0866*****		-0.0067
		(5.89)		**(15.26)**		**(5.89)**		(-0.28)
Leverage		**-0.0266*****		-0.0023		**-0.0222*****		0.0017
		(-5.04)		(-0.32)		**(-3.20)**		(0.21)
Pilot dummy	0.0065*	0.0056	0.0010	-0.0016	0.0067	0.0056	-0.0006	-0.0005
	(1.68)	(1.47)	(0.16)	(-0.27)	(1.14)	(0.97)	(-0.08)	(-0.07)
Firm FE	NO							
Year FE	YES							
Adjusted R^2	2.7%	5.5%	2.1%	11.9%	1.1%	2.4%	-0.04%	-0.04%

Table VI

Move toward FHK: Pre-During-Post Dummies instead of Year FE

Regression specification and sample is same as Table V, except that we replace year fixed effects with Pre, During, and Post dummy variables. Accruals and covariates are winsorized at 1%/99%. See Table II and III notes for detailed variable definitions. *t*-statistics, using standard errors clustered on firm, are in parentheses. *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Significant results, at 5% level or better, **in boldface.**

Accruals Type	**Operating**		**Total**		**AA**		**PMDA**	
*Pilot * During*	-0.0066	-0.0065	0.0049	0.0074	-0.0076	-0.0070	0.0007	0.0005
	(-1.59)	(-1.60)	(0.60)	(0.98)	(-1.12)	(-1.06)	(0.07)	(0.05)
*Pilot * Post*	-0.0048	-0.0044	-0.0053	-0.0032	-0.0084	-0.0079	-0.0015	-0.0016
	(-1.08)	(-1.03)	(-0.67)	(-0.43)	(-1.16)	(-1.12)	(-0.15)	(-0.16)
Sign reversal	0.0018	0.0021	-0.0102	-0.0106	-0.0008	-0.0009	-0.0022	-0.0021
	(0.46)	(0.55)	(-1.32)	(-1.44)	(-0.13)	(-0.15)	(-0.22)	(-0.21)
ln(assets)		**0.0044*****		0.0006		**-0.0033****		-0.0002
		(4.80)		(0.54)		**(-2.48)**		(-0.12)
market-to-book		0.0001		**0.0024*****		**0.0014*****		-0.0010
		(0.23)		**(3.47)**		**(2.61)**		(-1.31)
ROA		**0.0742*****		**0.3442*****		**0.0885*****		-0.0062
		(6.06)		**(15.31)**		**(6.02)**		(-0.26)
Leverage		**-0.0276*****		-0.0033		**-0.0235*****		0.0014
		(-5.24)		(-0.45)		**(-3.38)**		(0.17)
Pilot dummy	0.0065*	0.0056	0.0010	-0.0017	0.0067	0.0056	-0.0006	-0.0005
	(1.68)	(1.46)	(0.16)	(-0.28)	(1.14)	(0.96)	(-0.08)	(-0.07)
During	**0.0258*****	**0.0222*****	**0.0165*****	0.0086*	**-0.0192*****	**-0.0202*****	-0.0007	-0.0003
	(10.64)	**(9.29)**	**(3.40)**	(1.92)	**(-4.85)**	**(-5.14)**	(-0.11)	(-0.05)
Post	0.0029	0.0013	**-0.0292*****	**-0.0280*****	**-0.0110*****	-0.0080*	0.0065	0.0059
	(1.09)	(0.50)	**(-5.96)**	**(-6.34)**	**(-2.62)**	(-1.91)	(1.07)	(0.95)
Firm FE	NO							
Year FE	NO							
Adjusted R^2	1.4%	4.5%	1.5%	11.5%	0.4%	1.8%	-0.0%	-0.0%

Table VII

Move toward FHK: Two-way Clustering by Firm and Year

Regression specification and sample is same as Table VI, except we use two-way clustering of standard errors, on both firm and year, using the cluster2.ado Stata package developed by Mitchell Peterson. Coefficients on covariates are suppressed. *t*-statistics with two-way clustering are in parentheses. For the principal interaction variables, we report standard errors both with clustering on firm, and with clustering on firm and year, to show differences between these approaches. *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Significant results, at 5% level or better, in **boldface**. PMDA results with covariates are nearly identical if we include year fixed effects, as in FHK Table III model (3).

Accruals Type	**Operating**		**Total**		**AA**		**PMDA**	
*Pilot * During*	-0.0066	-0.0065*	0.0049	0.0074	-0.0076	-0.0070	0.0007	0.0005
	(-1.62)	(-1.70)	(0.52)	(0.87)	(-0.95)	(-0.90)	(0.15)	(0.10)
s.e (cluster on firm)	[0.0042]	[0.0041]	[0.0082]	[0.0076]	[0.0067]	[0.0067]	[0.0104]	[0.0104]
s.e. (2-way cluster)	[0.0041]	[0.0038]	[0.0094]	[0.0085]	[0.0080]	[0.0078]	[0.0047]	[0.0048]
*Pilot * Post*	**-0.0048****	**-0.0044****	-0.0053	-0.0032	-0.0084	-0.0079	-0.0015	-0.0016
	(-1.96)	**(-2.03)**	(-0.79)	(-0.51)	(-1.15)	(-1.12)	(-0.15)	(-0.16)
s.e (cluster on firm)	[0.0044]	[0.0043]	[0.0078]	[0.0073]	[0.0072]	[0.0071]	[0.0099]	[0.0099]
s.e. (2-way cluster)	[0.0025]	[0.0022]	[0.0067]	[0.0062]	[0.0073]	[0.0071]	[0.0097]	[0.0097]
Sign reversal	0.0018	0.0021	-0.0102	-0.0106	-0.0008	-0.0009	-0.0022	-0.0021
	(0.45)	(0.55)	(-0.99)	(-1.14)	(-0.19)	(-0.23)	(-0.24)	(-0.23)
s.e (cluster on firm)	[0.0039]	[0.0038]	[0.0077]	[0.0073]	[0.0061]	[0.0060]	[0.0100]	[0.0099]
s.e. (2-way cluster)	[0.0041]	[0.0038]	[0.0104]	[0.0093]	[0.0042]	[0.0039]	[0.0092]	[0.0092]
Covariates	NO	YES	NO	YES	NO	YES	NO	YES

Table VIII: F-Score, for Balanced Panel, FHK Sample Periods and FHK Specification

Regressions of F-score1 through F-score3 respectively on Pilot*During, Pilot*Post, and for even numbered models, indicated covariates, winsorized at 1%/99%. Regression specification is the same as Table VI, except that the outcome variable is F-score instead of accruals. F-score1, F-score2, and F-score3 are calculated using the set of coefficient estimates provided in Dechow et al. (2011) based on their Model (1), Model (2) and Model (3) respectively, and following their F-score formula on page 60 and example on page 61. See Table II and III notes for detailed variable definitions on independent variables. Following FHK, we require a balanced panel based on F-scores and covariates (sample firms must have data available to calculate F-scores and covariates over the entire FHK sample period). *t*-statistics, using standard errors clustered on firm, are in parentheses. *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Significant results, at 5% level or better, in **boldface**.

	F-score1		***F-score2***		***F-score3***	
	(1)	(20	(3)	(4)	(5)	(6)
*Pilot * During*	0.0069	0.0097	0.0110	0.0140	0.0242	0.0283
	(0.33)	(0.47)	(0.49)	(0.63)	(0.83)	(1.00)
*Pilot * Post*	-0.0118	-0.0092	-0.0037	-0.0008	-0.0046	-0.0015
	(-0.53)	(-0.43)	(-0.16)	(-0.04)	(-0.15)	(-0.05)
Sample mean	0.9787		1.0084		1.1028	
ln(assets)		**0.0288*** **		**0.0324*** **		**0.0258*** **
		(3.77)		**(3.99)**		**(2.62)**
market-to-book		**-0.0086*** **		**-0.0091*** **		**-0.0087*** **
		(-4.37)		**(-4.42)**		**(-3.45)**
ROA		**0.6009*** **		**0.6089*** **		**0.8184*** **
		(9.20)		**(8.78)**		**(9.31)**
Leverage		0.0186		0.0339		0.0387
		(0.55)		(0.93)		(0.88)
Pilot dummy	-0.0048	-0.0048	-0.0122	-0.0120	-0.0210	-0.0211
	(-0.18)	(-0.18)	(-0.43)	(-0.43)	(-0.57)	(-0.59)
During	**0.1006*** **	**0.0778*** **	**0.1120*** **	**0.0878*** **	**0.0491*** **	0.0231
	(7.99)	**(6.15)**	**(8.34)**	**(6.52)**	**(2.92)**	(1.39)
Post	-0.0120	**-0.0330** **	-0.0044	**-0.0280** **	**-0.0514*** **	**-0.0713*** **
	(-0.95)	**(-2.57)**	(-0.33)	**(-2.05)**	**(-2.87)**	**(-3.95)**
Intercept	**0.9514*** **	**0.7041*** **	**0.9760*** **	**0.7001*** **	**1.1087*** **	**0.8490*** **
	(59.04)	**(14.69)**	**(57.34)**	**(13.70)**	**(50.74)**	**(13.48)**
Firm-Year Obs.	10,890	10,890	10,890	10,890	10,890	10,890
Pilot (Control (firms)	431 (779)		431 (779)		431 (779)	
Adjusted R^2	1.0%	5.2%	1.0%	5.1%	0.4%	4.1%

Table IX: HF Measure, for Balanced Panel, FHK Sample Period and FHK Specification

Sample and covariates are the same as Table VIII. Following FHK, the dependent variables HF-1 is a dummy variable that equals to one if the firm's F-score1 is greater than or equal to the 99th percentile of the sample, and zero otherwise. HF-2 and HF-3 are similarly defined using 99th percentile of F-score2 and F-score3 as the cutoff, respectively. Average marginal effects from probit regressions are reported in lieu of probit coefficients to facilitate interpretation. *z*-statistics, using standard errors clustered on firm, are in parentheses. *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Significant results, at 5% level or better, in **boldface**. ,

	HF-1		***HF-2***		***HF-3***	
*Pilot * During*	-0.0054	-0.0054	-0.0051	-0.0054	0.0014	0.0010
	(-1.26)	(-1.25)	(-1.18)	(-1.25)	(0.30)	(0.23)
*Pilot * Post*	-0.0047	-0.0050	-0.0010	-0.0014	0.0022	0.0018
	(-0.87)	(-0.93)	(-0.19)	(-0.27)	(0.39)	(0.32)
ln(assets)		-0.0001		0.0004		-0.0009
		(-0.14)		(0.55)		(-1.07)
market-to-book		**-0.0008****		**-0.0011*****		-0.0006*
		(-2.51)		**(-3.14)**		(-1.89)
ROA		**0.0464*****		**0.0564*****		**0.0453*****
		(3.72)		(4.41)		(3.85)
Leverage		-0.0041		-0.0022		-0.0023
		(-1.01)		(-0.54)		(-0.59)
Pilot dummy	0.0038	0.0037	0.0035	0.0037	-0.0017	-0.0014
	(1.07)	(1.05)	(0.95)	(1.00)	(-0.51)	(-0.42)
During	**0.0054****	0.0050*	**0.0070****	**0.0065****	-0.0014	-0.0013
	(2.03)	(1.90)	**(2.50)**	**(2.36)**	(-0.54)	(-0.55)
Post	-0.0053	-0.0049	-0.0051	-0.0050	**-0.0095****	**-0.0086****
	(-1.54)	(-1.49)	(-1.43)	(-1.45)	**(-2.46)**	**(-2.24)**
Firm-Year Obs.	10,890	10,890	10,890	10,890	10,890	10,890
Pilot (Control) firms	431 (779)	431 (779)	431 (779)	431 (779)	431 (779)	431 (779)
Pseudo R^2	1.6%	4.7%	1.7%	6.1%	1.2%	4.3%

Table X. HF Measure: One-way vs. Two-way Clustered Standard Errors

Probit coefficients for indicated variables from same probit regressions as in Table IX (which reports average marginal effects), except that regression are run with both (i) standard errors clustered on firm, and (ii) standard errors clustered separately on firm and year. Coefficients on covariates are suppressed for brevity. Standard errors (*s.e.*) are in parentheses. *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Significant results, at 5% level or better, in **boldface**.

	HF-1		***HF-2***		***HF-3***	
*Pilot * During*	-0.2067	-0.2084	-0.1945	-0.2114	0.0518	0.0399
s.e (cluster on firm)	(0.1652)	(0.1680)	(0.1660)	(0.1702)	(0.1695)	(0.1722)
s.e. (2-way cluster)	**(0.0399)*** **	**(0.0209)*** **	**(0.0938)*** **	**(0.0807)*** **	(0.1198)	(0.1287)
*Pilot * Post*	-0.1793	-0.1939	-0.0376	-0.0550	0.0825	0.0688
s.e (cluster on firm)	(0.2072)	(0.2092)	(0.1996)	(0.2034)	(0.2111)	(0.2165)
s.e. (2-way cluster)	(0.1721)	(0.1653)	(0.0996)	(0.0826)	(0.1464)	(0.1504)
Covariates	NO	YES	NO	YES	NO	YES

Figure 1. Replication and Extension of FHK Figure 2

The panels below present our results for our four accruals measures, Operating Accruals, Total Accruals, AA, and PMDA, using the same format as FHK Figure 2. The left panels presents results for our preferred specification: unbalanced panel, our sample periods, including all firms with sufficient data to compute accruals. The right panels present results for the FHK specification: their sample periods, balanced panel, and sample limited to firms with data on covariates for all years. Point estimates are sample means, for pilot and control firms, over the Pre, During, and Post periods.

FHK Figure 2 from their Paper

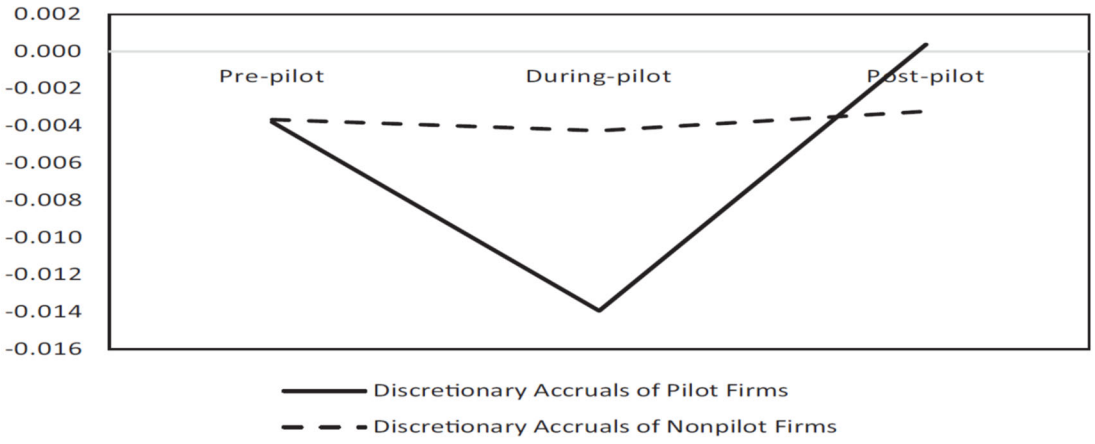
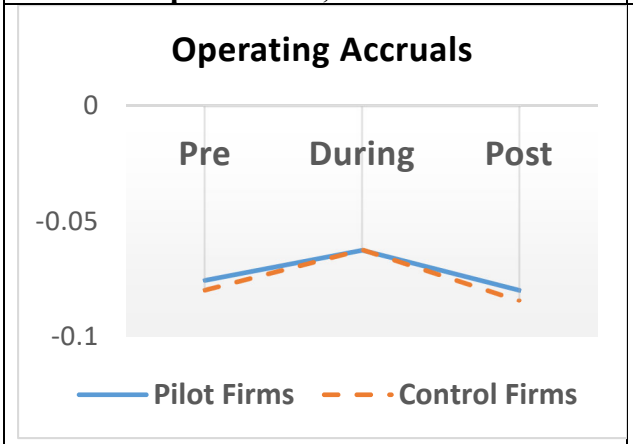
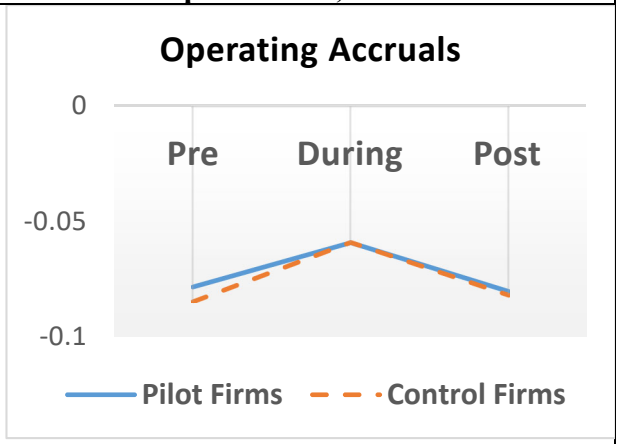
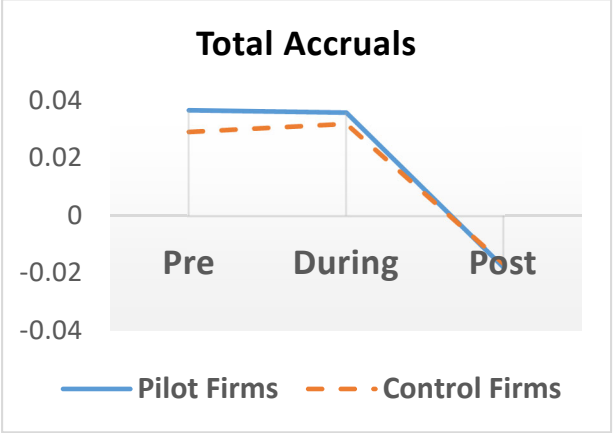


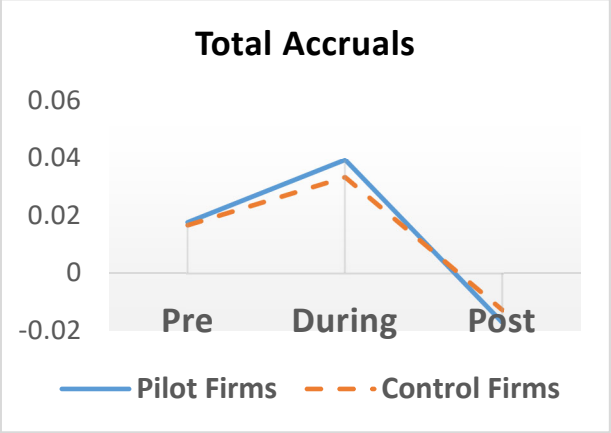
Figure 2. Discretionary accruals for pilot vs. nonpilot firms. This figure displays the results reported in Panel A of Table II. It depicts the mean *Discretionary accruals* for the balanced panel sample of the treatment group and control group for the periods before, during, and after Regulation SHO's pilot program, that is, 2001 to 2003, 2005 to 2007, and 2008 to 2010. The sample comes from the 2004 Russell 3000 index and contains firms that have data available to calculate financial characteristics and discretionary accruals over the entire sample period (i.e., 2001 to 2003 (inclusive) and 2005 to 2010 (inclusive)).

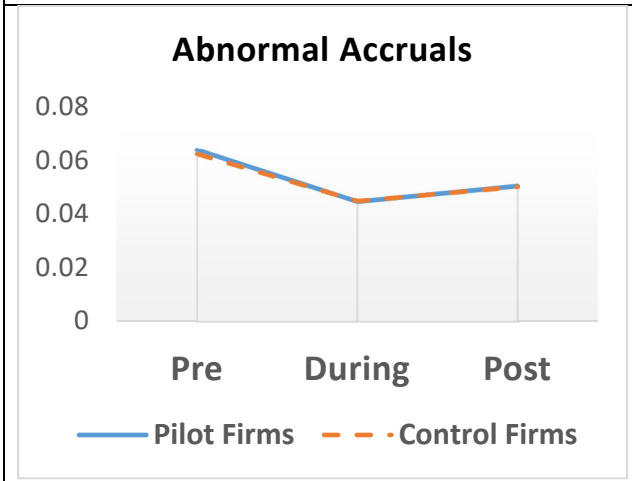
Results for Univariate Means, Across Accruals Measures

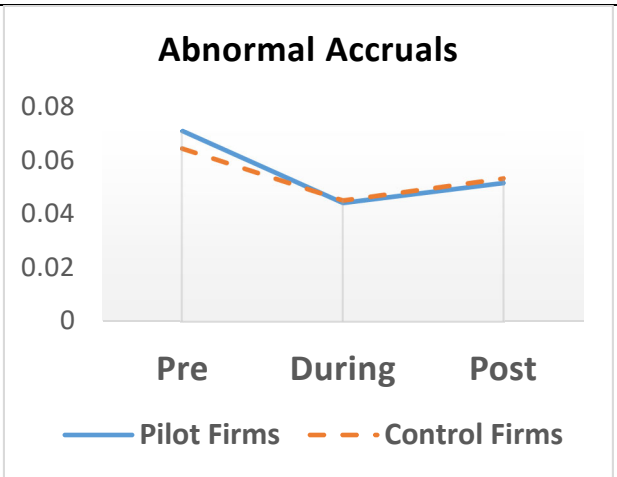
Means for each accruals measure, during the Pre, During, and Post periods, separately for pilot and control firms. Left hand figures use our unbalanced panel and sample periods. Right hand figures use a balanced panel and the FHK sample periods.

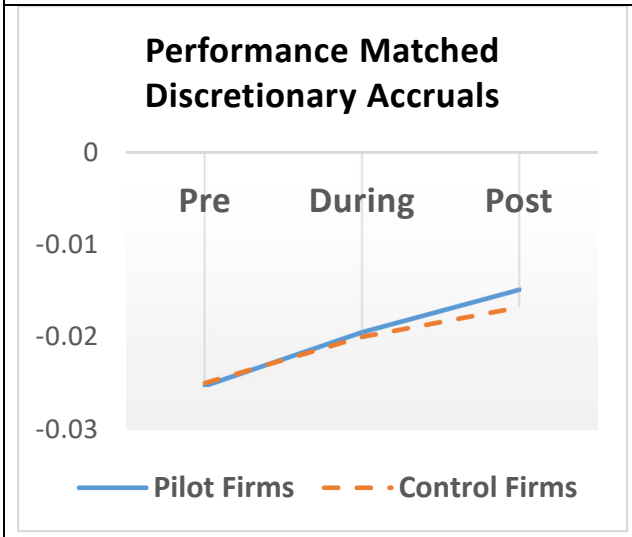
Our Sample Periods, Unbalanced Panel	FHK Sample Periods, Balanced Panel
	


Total Accruals
0.04
0.02
0
-0.02
-0.04
Pre
During
Post
Pilot Firms
Control Firms


Total Accruals
0.06
0.04
0.02
0
-0.02
Pre
During
Post
Pilot Firms
Control Firms


Abnormal Accruals
0.08
0.06
0.04
0.02
0
Pre
During
Post
Pilot Firms
Control Firms


Abnormal Accruals
0.08
0.06
0.04
0.02
0
Pre
During
Post
Pilot Firms
Control Firms


Performance Matched Discretionary Accruals
0
-0.01
-0.02
-0.03
Pre
During
Post
Pilot Firms
Control Firms

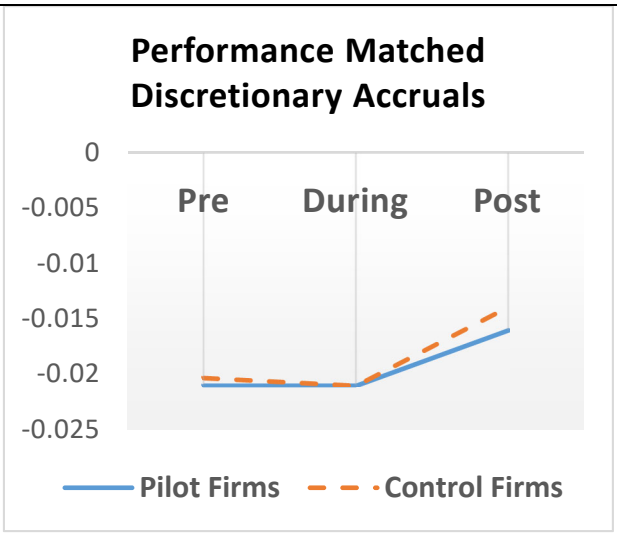

Performance Matched Discretionary Accruals
0
-0.005
-0.01
-0.015
-0.02
-0.025
Pre
During
Post
Pilot Firms
Control Firms

Figure 2. Leads-and-Lags Graphs for Accruals Measures: Our Specification; Unbalanced Panel

Leads and lags regressions of indicated accruals measures, from regressions following eqn. (5), with firm and year fixed effects, over fiscal years 1998-2010. Variables are the same as in Table II. y-axis shows coefficients on the lead and lag dummies; vertical bars show 95% CIs around coefficients, using standard errors clustered on firm. Coefficient for 2004 is set to zero. Vertical lines show start and end of the experiment period.

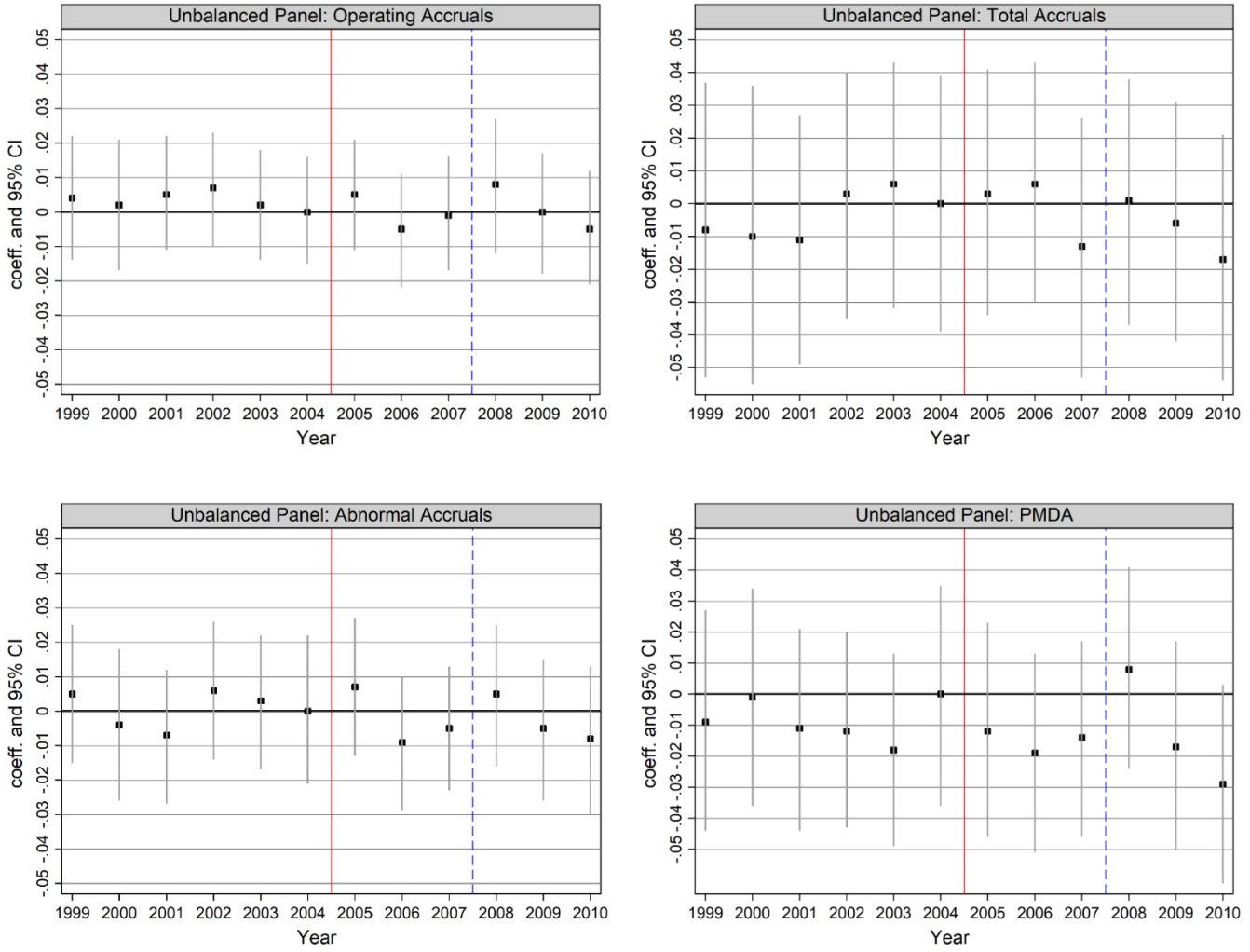
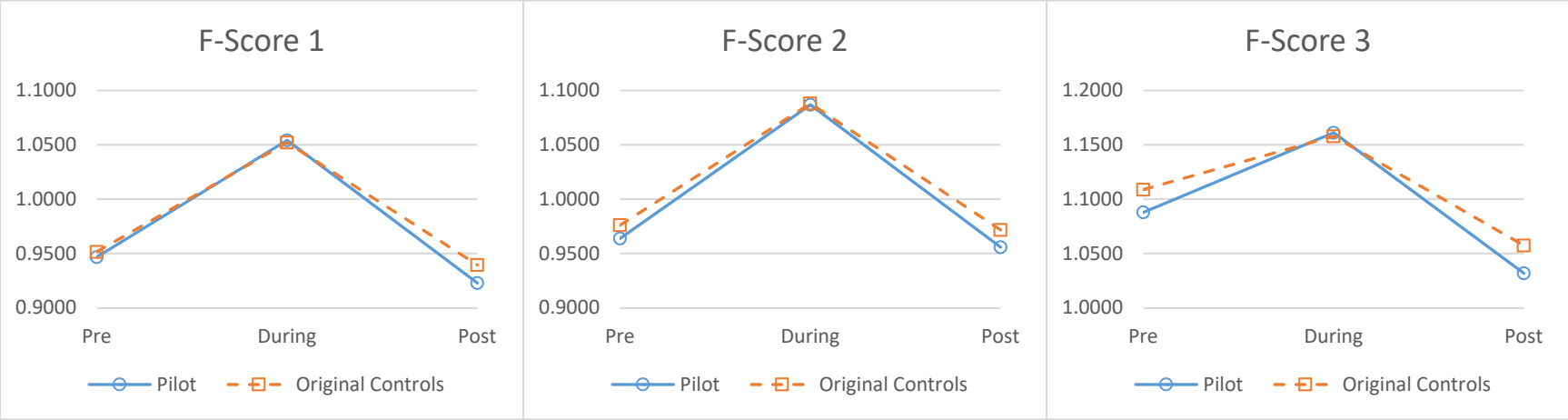


Figure 3. Univariate Means – Pre/During/Post

Panel A. Comparable to Table VIII: F-score, for Balanced Panel, FHK Sample Periods

Means for each F-score measure, during the Pre, During, and Post periods, separately for pilot and control firms, using a balanced panel of firms with data on all F-score measures for all sample years, and the FHK sample periods.



Panel B. Comparable to Table IX: HF Measure, for Balanced Panel and FHK Sample Periods

Means for each HF measure, during the Pre, During, and Post periods, separately for pilot and control firms. Sample and sample periods are same as Panel A.

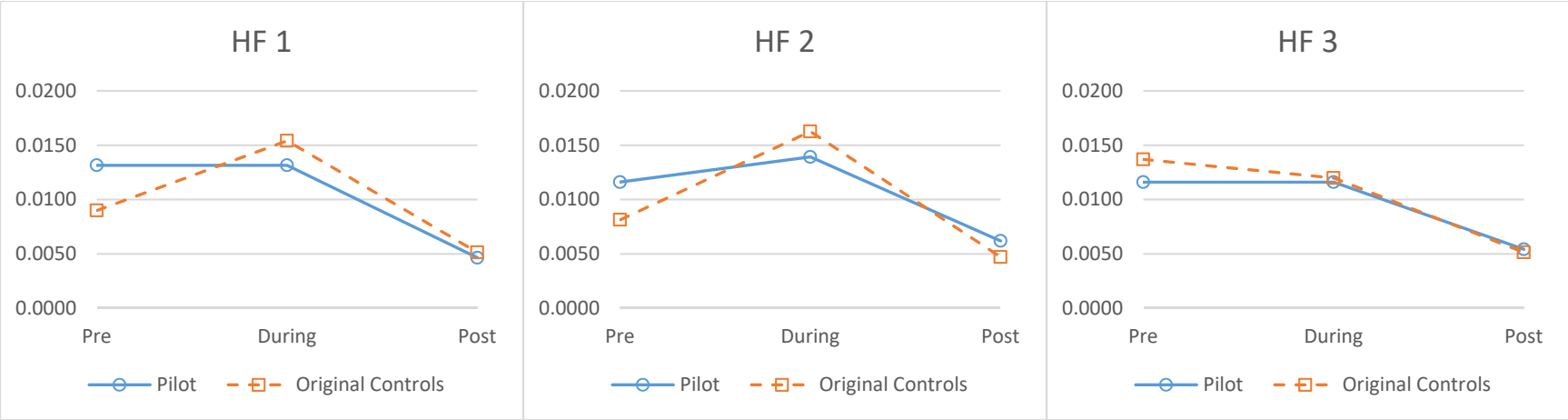
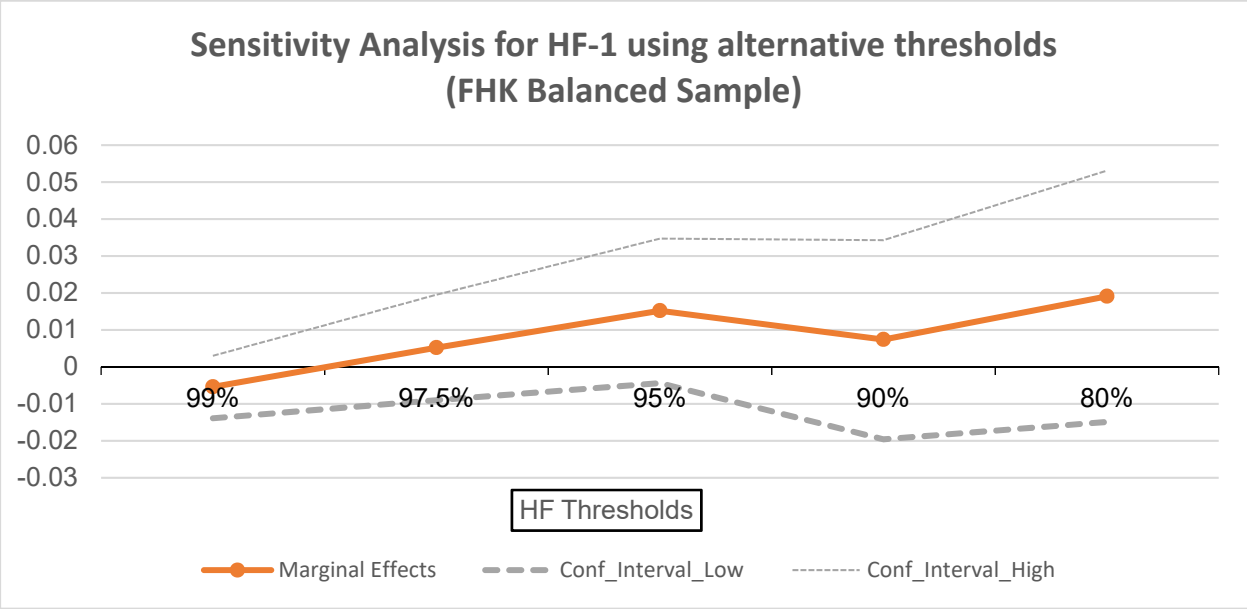
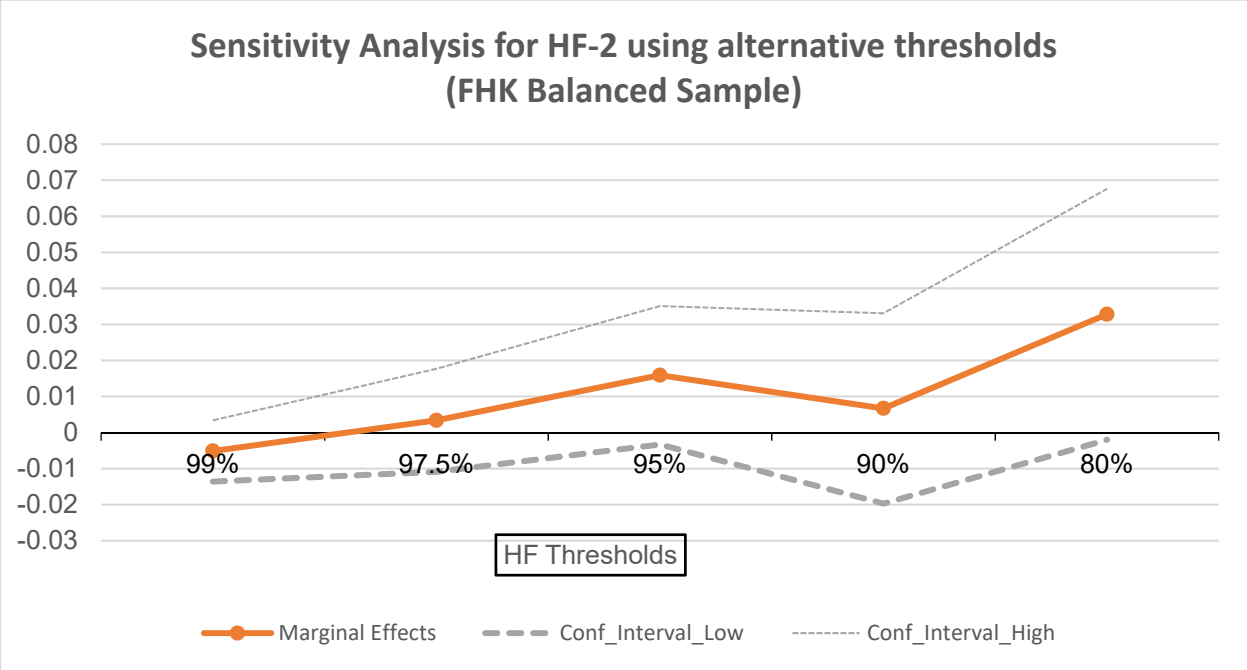


Figure 4. Sensitivity Analysis using Alternative HF Thresholds

Graph shows average marginal effects on *Pilot*During* from probit regression with variations on the FHK HF-1 and HF-2 measures, using indicated thresholds, as the dependent variables, respectively, using FHK sample and specification, without covariates (similar to regression 1 in Table IX). Circles show point estimates. Upper and lower dashed lines show top and bottom of the 95% confidence interval, using standard errors clustered on firm. Threshold percentiles for setting HF=1 values are 99% (used by FHK), 97.5%, 95%, 90%, and 80%.





Appendix 1. News Coverage of the Short-Sale Experiment

The short-sale experiment attracted very little press attention. We searched the business press for news stories and other information about the experiment, both during the launch period (June 2004 through April 2005) and the experiment period (May 2005-May 2007), and present below the results of that search in. The SEC's announcement on June 23, 2004, of its plans to conduct the experiment, and the formal experiment announcement on July 28, 2004, attracted no news attention whatsoever, at least none that we could find. Most of the comments the SEC received supported the experiment.[24] The very limited news coverage of the experiment between the announcement and the actual launch contained only technical explanations of how the experiment would work.

We found no evidence of concern or political opposition from firm managers. The SEC's July 2004 announcement of the pilot was not covered in any of the standard business news sources (including the Dow Jones News Service (DJNS), Bloomberg, the Wall Street Journal (WSJ), and the New York Times (NYT)).[25] The experiment launch, in May 2005, was noted in a DJNS story a few days earlier, with a WSJ summary the next day. In 2006, the SEC extended the experiment, originally scheduled for one-year, with minimal press attention and no apparent controversy. We found one DNJS story about the extension, with a WSJ summary the next day. Neither WSJ story was long enough to warrant a byline.[26]

The SEC's proposal to repeal the short-sale rule, announced in December 2006, also attracted no apparent opposition that we could find. An NYT story about the repeal explains:[27]

> You may not have read of this proposal. It was virtually ignored by the news media, and if any companies are upset about it, they have not made themselves known. A pilot program that exempted some companies from the so-called uptick rule starting in 2005 drew little attention.

---

[24] See Announcement Release, reprinted in 69 Federal Register (Aug. 6, 2004), at 48,008, 48012.

[25] Appendix A lists all of the business press stories we found. The first story was on 30 November 2004, and explained that the SEC was delaying the experiment launch to give the exchanges time to make programming changes so they could implement the experiment. This was not even a separate story about the short sale experiment; instead, this story was appended to a main story about another SEC rule. Judith Burns, SEC Delays Short-Sale Pilot, Seeks NMS Comment, *Dow Jones News Service* (Nov. 30, 2004).

[26] SEC Pilot Program To Halt `Uptick' Rule, *Wall Street Journal* (April 29, 2005); SEC to Extend Test On Short-Sale Rules, *Wall Street Journal* (April 22, 2006).

[27] Floyd Norris, 70 Years Later, A Scapegoat Gets a Break, *New York Times* (Dec. 8, 2006). The NYT also published an op-ed article in October 2006, supporting repeal of the short-sale rule. Richard Sauer, Bring on the Bears, *New York Times* (Oct, 6, 2006).

This is the *only* NYT story about the experiment we found.

The SEC formally approved repeal in June 2007. A *Wall Street Journal* story on the repeal explained that the rule had become "more of an annoyance than a hindrance" to short-sellers, discussed researchers' view that "the uptick rule's usefulness has disappeared", and did not mention any opposition to repeal.[28]

New Stories about the Experiment

To capture news stories about the SEC Experiment we searched the Factiva database, for the following sources using the following search terms, over July 1, 2004 through year-end 2007: Dow Jones News Service (DJNS); New York Times (NYT); Wall Street Journal (WSJ); Bloomberg; PR Newswire; Forbes; Bloomberg BusinessWeek. We used all combinations of two or more of the following search terms: SEC (also S.E.C.); Regulation (also Reg) SHO; Pilot Program; Short Sale(s). It was infeasible to search solely for "short sale" or for "pilot program" due to a large number of largely irrelevant matches.

Date	Source	Title	Authors	Notes and excerpts
11/30/2004	DJNS	SEC Delays Short-Sale Pilot, Seeks NMS Comment	Judith Burns	Main story is SEC seeking comment on Regulation NMS (for national market system); story notes that "Separately, the SEC announced that an experiment in lifting short-sale restrictions for some stocks, set to begin in January, will be delayed until May. [SEC spokesperson] Nazareth said the additional time will allow exchanges to complete the necessary computer programming modifications. . . ."
04/26/2005	DJNS	SEC Encourages Researchers To Study Short-Sale Data	Judith Burns	**"**SEC staffers on Tuesday invited outside research on whether market quality would change if short-sale restrictions such as the "bid" and "tick" test were eliminated, and whether a uniform bid test should be extended to smaller stocks. The SEC staff said it has arranged for markets to publicly release monthly short-sale data, and urged researchers to submit studies on the short-sale experiment to the SEC.**"**
04/28/2005	DJNS	SEC To Suspend 'Uptick' Rule In Reg SHO Pilot Program	Karen Talley	Describes experiment, based on SEC press release: "The suspension of the uptick rule is aimed at allowing the agency to obtain data to help assess whether it should be removed in part or in whole. . . . The pilot program is seen by the SEC as being in the public interest because the uptick rule may actually harm market quality by inhibiting free movement of prices. The rule is about 70 years old and there is some sentiment it may have outlived its purpose, the SEC said."
04/29/2005	WSJ	SEC Pilot Program To Halt `Uptick' Rule	none	Summary of DJNS story from prior day.

[28] Spencer Jakab, Short-Sellers May Owe ETFs Some Thanks --- Dropping of 'Uptick' Rule By SEC Comes as Growth Of Stock Baskets Is Soaring, *Wall Street Journal* (June 15, 2007).

Date	Source	Title	Authors	Notes and excerpts
11/17/2005	DJNS Int'l; (based on PR Newswire)	State Regulators Set Forum On Abusive Naked Short-Selling	John Connor	North American Securities Administrators Association NASAA plans meeting to discuss naked short selling. Also, ""the SEC launched a pilot program in May to determine the effectiveness of [Regulation SHO]. 'Since the pilot program has reached the midway point, we believe this forum offers a timely opportunity for a thorough discussion of the effectiveness of Regulation SHO from a variety of perspectives," NASAA president Struck said.
04/21/2006	DJNS	SEC OKs Extending Short-Sale Pilot To August 2007	Judith Burns	Based on SEC Press Release: "The Securities and Exchange Commission announced Friday that it will extend an experiment lifting short-sale restrictions for some stocks. The one-year pilot program, set to end April 28, will be extended until Aug. 6, 2007. . . . The SEC said the extension will spare markets the expense of changing computer systems to restore restrictions on stocks in the pilot while it analyzes results from the experiment."
04/22/2006	WSJ	SEC to Extend Test On Short-Sale Rules	none	Summary of DJNS story from prior day.
06/13/2006	DJNS	SEC To Host Sept. 15 Discussion Of Short-Sale Experiment	Judith Burns	SEC to hold roundtable on Sept. 15 to discuss research on short-sale experiment: "The SEC said the session on its Regulation SHO will discuss results of an experiment that lifted restrictions on short sales of some stocks and hear findings from academic researchers who have studied it."
06/28/2006	WSJ Europe	Lawsuits set focus on short selling	Randall Smith	Story about naked short selling; brief mention of Pilot Program: "Some short sellers say they can't knock down stock prices because of "uptick" rules limiting such sales when prices are falling. However, the SEC has a pilot program exempting about 1,000 stocks from the rules, which also don't apply to some trades off the stocks' exchanges."
09/15/2006	DJNS	SEC Meets To Revisit Short-Sale Restrictions	Judith Burns	Review of research on short-sale rule presented at SEC roundtable; that research persuaded the SEC to rescind the short-sale rule.
10/06/2006	NYT	Bring On the Bears (Editorial)	Richard Sauer	"In an unusual (and laudable) effort to measure whether a long-lived regulation actually works, the Securities and Exchange Commission recently completed a pilot program to suspend the uptick rule for a third of the stocks on the Russell 3000 index and compare their performance to stocks still subject to the rule."
12/08/2006	NYT	70 Years Later, A Scapegoat Gets a Break	Floyd Norris	"You may not have read of that proposal. It was virtually ignored by the news media, and if any companies are upset about it, they have not made themselves known. A pilot program that exempted some companies from the so-called uptick rule starting in 2005 drew little attention."
03/23/2007	DJNS	SEC Seeks To 'Modernize' Short-Selling Regs - Official	Daisy Maxey	Story on remarks by James Brigagliano, associate director in SEC Division of Market Regulation division on short-sale rules, at educational seminar on regulatory issues for hedge fund managers.
06/14/2007	Dow Jones Commodities	Reforms Of Short-Sale Rules Sensible	Spencer Jakab	Discusses SEC approval of rule repeal and research finding limited effect of short-sale rules.

Date	Source	Title	Authors	Notes and excerpts
	Service	(So Far)		
06/15/2007	WSJ	Short-Sellers May Owe ETFs Some Thanks --- Dropping of 'Uptick' Rule By SEC Comes as Growth Of Stock Baskets Is Soaring	Spencer Jakab	Similar to previous day story in Dow Jones Commodities Service.

First news story:

The first story we found in the business press came four months after the experiment was announced. On 30 November 2004, the *Dow Jones News Service* published a short summary of a delay in launching the experiment, to let the exchanges make programming changes. This was *not* a separate story about the experiment; it was appended to a main story about another SEC rule. The story, in full:

> **Short-Sale Experiment Delay**
>
> Separately, the SEC announced that an experiment in lifting short-sale restrictions for some stocks, set to begin in January, will be delayed until May.
>
> [SEC market regulation division director Annette] Nazareth said the additional time will allow exchanges to complete the necessary computer programming modifications. She said the SEC originally planned to have individual brokerage firms assume responsibility for that task, but exchanges volunteered to step in when brokerages balked at the timetable. "You want to make sure the programming is right," said [Securities Industry Association President Marc] Lackritz, who praised the SEC's temporary delay in the experiment.
>
> The SEC approved the short-sale pilot program in July, as part of Regulation SHO, a broader package of reforms involving short sales. The one-year pilot program was to begin Jan. 3, but under an order approved Tuesday, the SEC agreed to delay it until May 2, and have it run through April 2006. All other terms of the pilot project remain the same, and all other provisions of Regulation SHO will take effect Jan. 3, as planned, Nazareth stressed.
>
> Short selling involves sales of borrowed stock. Short sellers must replace the shares at a later date and profit if the stock price declines in the interim. Although short selling is legal, it is subject to restrictions to prevent market manipulation such as "bear raids" that drive stock prices sharply lower.
>
> The SEC's pilot program will suspend short-sale restrictions for about 1,000 stocks in the Russell 3000 Index. Most are listed on the New York Stock Exchange or the Nasdaq Stock Market, with about 2% listed on the American Stock Exchange.
>
> Stocks included in the pilot project include Corning Inc. (GLW), Kohl's Corp. (KSS), Marsh & McLennan Cos. (MMC), Oracle Corp. (ORCL), Peet's Coffee & Tea Inc. (PEET) and Walt Disney Co. (DIS).
>
> Regulators hope the experiment will allow them to study the effect of trading without restrictions on short sales, possibly paving the way to lift short-selling restrictions, in whole or in part, for most actively traded stocks.

Internet Appendix

for

The Reg SHO Reanalysis Project: Reconsidering Fang, Huang and Karpoff (2016) on Earnings Management

Bernard Black
Northwestern University, Pritzker School of Law and Kellogg School of Management

Hemang Desai
Southern Methodist University, Cox School of Business

Kate Litvak
Northwestern University, Pritzker School of Law

Woongsun Yoo
Saginaw Valley State University, College of Business and Management

Jeff Yu
University of Arizona, School of Accountancy

(draft September 2019)

Appendix A. Summary of Papers Studying the Reg SHO Experiment

The table below summarizes the papers we found by August 2019, studying the short-sale experiment. The list includes 5 early papers, 33 recent papers, plus Boehmer, Jones and Zhang (2018), a market microstructure paper which does not fall within either category. For papers which study multiple questions, we exercised judgment in specifying the primary question.

No.	Papers	Research Question	Channels	Summary of Results	Sample size	Exchange
1	Albertus, Bird, Karolyi, and Ruchti (2017), WP	Effect of Reg. SHO exp't on intrafirm capital allocation	Short selling threats discipline managers' capital allocation decisions.	No effect on multinationals' total investment, but 30% higher capital allocation to foreign subsidiaries with strong recent performance; subsidiaries that receive additional capital show no decrease in productivity.	376 multinationals with 5,575 subsidiaries.	All
2	Alexander and Peterson (2008), J Fin. Markets	Effect of Reg. SHO exp't on trader behavior and market quality	Removal of price tests lets short sellers to place orders that receive quicker execution.	For NYSE firms, removal of uptick test lets short sellers trade more aggressively: treated firms have smaller short trade sizes, more short trades, more short volume, and smaller ask depths. No evidence that removal of uptick rule affects liquidity, price volatility or price efficiency. Smaller, usually insignificant effects on Nasdaq. The study quotes prior work (Ferri et al., 2006) to suggest that the bid test was not effective.	223 matched NYSE pairs and 195 matched Nasdaq pairs	NYSE and Nasdaq, studied separately
3	Bai (2008), Rutgers Bus. Law J.	Effect of Reg. SHO exp't on intraday stock prices and short selling after negative earnings surprises	Negative earnings surprises can cause share price changes and higher short sale volume. If uptick rule matters, one should find greater effect on price declines and short sale volume immediately after negative earnings surprises.	No evidence that treated firms experience faster price declines or higher short sale volume than control firms. For NYSE firms, short orders with immediate execution barred by the uptick rule were executed within 15 minutes. Nasdaq bid test was usually not binding.	Neg. earnings surprises: 311 to pilot firms (170 NYSE; 255 distinct firms); 634 to control firms (341 NYSE; 525 distinct firms	NYSE and Nasdaq, studied separately

No.	Papers	Research Question	Channels	Summary of Results	Sample size	Exchange
4	Bai, Lee, and Zhang (2018), WP	Effect of Reg. SHO exp't on firms' investment in workplace safety (measured by work-related injury and illness rates)	*Performance Pressure Hypothesis*: an increase in short selling pressure leads managers to be more myopic and shift away from long-term investments such as those in workplace safety.	Work-related injury and illness rates increase significantly for treatment firms after Reg SHO compared to control firms. The effect is stronger for firms that are in more competitive industries, are more financially constrained, have disadvantaged labor force and have poor corporate governance mechanism.	286 pilot and 532 control firms	All
5	Bennett and Wang (2018), WP	Effect of Reg. SHO exp't on forced CEO turnover	Revelation channel: Short selling accelerates price discovery, which makes firms more likely to fire their CEOs based on the new negative information revealed in stock price. Manipulation channel: Price manipulation by uninformed short sellers can mislead CEOs into making poor decisions and in turn lead to worse CEO performance. Meanwhile, boards can be misled by the manipulated lower prices and blame CEO for the lower market valuation.	Increased likelihood of CEO forced turnover for pilot firms. Consistent with revelation channel, effects are stronger when firms have more earnings management, less informative stock prices and less competitive product markets. Consistent with manipulation channel, the effects are also stronger when firms have more growth opportunities, fewer blockholders and less volatile stock.	425 pilot and 756 control firms	All
6	Bhattacharya, Christensen, Li and Ouyang, 2019.	Effect of Reg SHO exp't on pro-forma reporting.	Increased threat of short selling due to suspension of uptick rule/bid test can discipline pro-forma reporting by pilot firms.	Results suggest that increased threat of short selling significantly curbs aggressive non-GAAP pro-forma reporting by Pilot firms.	225 pilot firms and 507 control firms.	All

No.	Papers	Research Question	Channels	Summary of Results	Sample size	Exchange
7	Billett, Liu, and Tian (2018), WP	Effect of Reg SHO exp't on information spillover and cross monitoring between stock market and loan market	Reg. SHO affects information production and monitoring by short sellers in the stock market, which may have a spillover effect to the syndicated loan market and influence loan contracting terms. Conversely, bank monitoring also provides information and/or certification to the stock market	While firms without bank monitors exhibit a significant decline in stock price upon the announcement of SHO, firms with bank monitors do not react. Firms affected by SHO enjoy a 21 basis point lower loan spread that increases to 36 basis points for bank-dependent firms. SHO does not affect non-price loan terms such as maturity, amount, collateral and covenants.	527 pilot and 1012 control firms.	All
8	Boehmer, Jones and Zhang (2018), WP	Effect of 2007 repeal of uptick rule on arbitrage trading.	The 2007 full uptick repeal makes synchronous portfolio trading such as index arbitrage easier and less costly to execute. The 2005 partial uptick repeal incentivize aggressive short sellers to shift toward pilot firms and away from control firms.	Short activity increases substantially for all firms after the 2007 full uptick repeal, even for pilot firms for which the uptick rule was suspended since 2005. Shorting activity on non-pilot firms co-move more with pilot firms after the full repeal. Short selling aggressiveness among the non-pilot firms decreases after 2005 partial repeal; also reduced comovement between shorting activity on pilot and control firms.	NYSE firms: 360 pilot and 728 control	NYSE
9	Cai and Guo (2018), WP	Effect of Reg. SHO exp't on real earnings management	The threat of short selling could discipline managers and constrain real earnings management	Pilot firms experience a reduction of real earnings management after SHO relative to the non-pilot firms. The effect is stronger in firms where managers are more entrenched and firms with lower institutional ownership and higher analyst coverage.	736 pilot and 1503 control firms	All

No.	Papers	Research Question	Channels	Summary of Results	Sample size	Exchange
10	Chang, Huang, Su, and Tseng (2018), WP	Effect of Reg SHO exp't on compensation contracts.	Relaxation of short sale constraint decreases the speculative component in stock prices, making short-termism less attractive to current shareholders, thereby reducing short-term incentives in managerial compensation contracts.	Pilot firms have longer CEO pay duration relative to non-pilot firms during SHO, but this difference becomes insignificant after SHO concludes. This result is concentrated among firms with pre-existing high investor disagreement. Removing short-sale constraints also leads to longer CEO investment horizon and fewer stock repurchases.	254 pilot and 496 control firms	All
11	Chen, Zhu, and Chang (2017), Accounting and Finance.	Effect of Reg. SHO exp't on corporate payouts	Firms adjust corporate payout policies to signal their good quality to counteract intensified short-selling pressures after SHO.	Small pilot firms increase their cash dividends during SHO relative to control firms, and continue to pay higher cash dividends even after the pilot program ended in 2007. However, share repurchase activities of pilot firms remain unchanged before, during and after SHO.	616 pilot and 1275 control firms	All
12	Chen, Cheng, Luo, and Yue (2014), WP	Effect of Reg. SHO exp't on corporate disclosures	Reduction in short sale constraints incentivize managers to disclose good news in a timely manner to boost the confidence of the stakeholders in the firm and to deter short sellers.	Pilot firms are more likely to issue good news management forecasts without changing the issuance of bad news forecasts. Pilot firms are more likely to bundle bad news forecasts with good news earnings announcements, but management forecast optimistic bias does not increase.	768 pilot and 1484 control firms	All
13	Chen and Wu (2019), WP	Effect of Reg. SHO exp't on real activity manipulation	Short selling threat leads pilot firms to reduce real earnings management activities relative to control firms.	Pilot firms reduces real activity manipulation by increasing abnormal cash flows and decreasing abnormal production activities.	487 pilot and 989 control firms	All
14	Cheng and Zhang (2019), WP	Effect of Reg. SHO exp't on credit rating informativeness	Short sellers discipline credit rating agencies and also provide additional information to credit rating agencies.	Credit ratings of pilot firms become more informative than those of control firms during the pilot program.	276 pilot and 510 control firms	All

No.	Papers	Research Question	Channels	Summary of Results	Sample size	Exchange
15	Choi (2018), Finance Research Letters	Effect of Reg SHO exp't on analyst forecast precision.	Reg. SHO reduced SS constraints for Pilot firms. The paper assumes that this increased SS activity for Pilot firms. The increased activity by passive investors in turn causes analysts to expend less effort in producing their forecasts. This is evidenced a greater incidence of rounded forecasts for Pilot firms relative to control firms.	In a regression of ROUNDING on firm characteristics, the coefficient on Pilot*During is positive indicating that analysts' tendency to issue rounded forecasts increased during the period of Reg SHO for Pilot firms relative to Control firms.	497 pilot and 993 control firms.	All
16	Chu, Hirshleifer and Ma (2017)	Effect of Reg. SHO exp't on share return anomalies	Use the Reg. SHO to examine the impact of lower arbitrage constraints on returns to anomalies.	Monthly returns reduce on average by -0.69% for Pilot firms during the test period. The return difference between Pilot and Control firms vanishes after the experiment ends. No effect is observed for Nasdaq firms.	515 pilot and 1025 control firms	NYSE
17	Clinch, Li, and Zhang (2018), WP	Effect of Reg. SHO exp't on bad news disclosure	The paper conjectures that lifting of SS restrictions on pilot firms reduces incentives and potentially increases litigation costs for managers to withhold bad news and hence expects an increase in voluntary bad news forecasts for Pilot firms relative to control firms.	Documents a significant increase in likelihood of bad news management forecasts for Pilot firms relative to control firms during the experiment period. No corresponding result is observed in the Pre or the Post period or for good news management forecasts.	919 pilot and 1888 control firms	All
18	De Angelis, Grullon, and Michenaud (2017), RFS	Effect of Reg. SHO exp't on CEO compensation	Reg SHO increases the threat of bear raids for Pilot firms causing Pilot managers to be more risk averse. The firm can then compensate for the increased risk aversion by increasing the convexity of pay and also by providing more downside protection to managers' pay.	Pilot firms increase convexity of payoffs (Vega) and increase stock option grants relative to share grants	472 pilot and 878 control firms. Balanced sample: 388 pilot and 702 control firms	All

No.	Papers	Research Question	Channels	Summary of Results	Sample size	Exchange
19	Deng, Gao, and Kim (2019), JCF, forthcoming	Effect of Reg SHO exp't on stock price crash risk (negative skewness of returns, down to up volatility)	Lifting SS constraints for Pilot firms reduces managers' incentives for withholding bad news which in turn reduces stock price crash risk for pilot firms. The investment efficiency of Pilot firms also improves.	The stock price crash risk decreases for the Pilot firms during the experiment period relative to Control firms.	776 pilot and 1,734 control firms.	All
20	Diether, Lee, and Werner (2009), JF.	Effect of Reg SHO exp't on short-selling activity, daily returns and volatility, and market quality. Outcomes: Share returns, short sale volume, spreads, and volatility		Higher short sale volume. No effect on daily volatility or share returns. Small increase in spreads and intraday volatility for NYSE firms but not NASDAQ firms	NYSE: 448 pilot and 904 control firms Nasdaq: 376 pilot and 757 control firms	NYSE and Nasdaq, studied separately
21	FHK: Fang, Huang, and Karpoff (2016), Journal of Finance	Effect of Reg SHO on earnings management during the experiment period. The measure of EM is Performance matched discretionary accruals	Increased threat of SS causes Pilot firms to decrease EM relative to control firms during the experiment period relative to control firms.	PMDA of Pilot firms are shown to decline during the experiment period relative to control firms. The difference reverses in the Post experiment period.	388 pilot and 709 control firms	All
22	Francis, Samuel, and Wu (2017), WP	Effect of Reg SHO exp't on dividends.	Increased SS can improve price efficiency for Pilot firms which in turn decreases the need for dividend payout. On the other hand, increased SS increases monitoring of managers of Pilot and hence managers become risk averse and increase payout instead of risky investments.	Pilot firms are more likely to increase dividends than control firms.	590 pilot Firms and 1166 control firms. Balanced sample: 448 pilot and 834 control firms.	All

No.	Papers	Research Question	Channels	Summary of Results	Sample size	Exchange
23	GMW: Grullon, Michenaud, and Weston (2015), RFS	Effect of Reg SHO exp't on capital expenditures, R&D, assets, and equity and debt issuance.	For overvalued firms, removing short-selling constraints can lead to lower share prices and reduce overinvestment. Channel: Managers learn from share prices when making investment decisions.	1) Short interest increases for pilot firms after experiment announcement (July 28, 2004) 2) Share prices of pilot firms fall after SEC approves experiment (June 23, 2004) but before list of treated firms is known. 3) Pilot firms invest less during experiment period. (lower CAPEX and asset growth).. 4) Pilot firms reduce equity issuance. Results (2)-(4) larger for small firms.	651 pilot and 1,279 control firms	All (large and small separately)
24	He and Tian (2016), WP	Effect of Reg SHO exp't on managerial myopia.	Tests whether increased threat of SS due to Reg SHO for Pilot firms exacerbates or mitigates managerial myopia. The proxy for myopia is firm innovation measured by the number of patents and patent citations.	Quality, value and originality of patents generated by the Pilot firms improve relative to control firms suggesting that short sellers mitigate myopia.	592 pilot and 1,158 control firms.	All
25	HHZ: Hope, Hu, and Zhao (2016), JAE	Effect of Reg SHO exp't on audit fees.	Auditors of Pilot firms might face higher litigation risk due to higher ex-ante threats of short-selling activities. Thus auditors might increase their audit effort to reduce the probability of making mistakes and ask for higher fees (audit effort channel), or pass the burden of increased risk onto shareholders by charging higher fees without increasing their audit effort (risk-premium channel).	Treated firms incur higher audit fees. The effect is more pronounced for (1) firms with higher loan default risk and (2) firms whose managers have high levels of in-the-money options. Default risk is measured using Bharath and Shumway (2008)'s measure. High level of in-the-money options is relative to SIC 4 digit industry median in at least four out of five years from 2000-2004.	538 pilot firms and 1,072 control firms	All

No.	Papers	Research Question	Channels	Summary of Results	Sample size	Exchange
26	Hu, Li, Liao and Zhang (2019)	Effect of Reg SHO exp't on loan loss provisions and risk taking by bank holding companies	Reduction in short sale constraints for pilot BHCs increases the fear of bank runs and hence incentivizes pilot firms to look better and healthier.	Pilot BHCs are shown to delay loan loss recognition and the effect is concentrated in smaller BHCs and those with lower regulatory capital.	bank holding companies: 72 pilot and 175 control	
27	Kan and Gong (2018), Int'l Review Fin.	Effect of Reg SHO exp't on price informativeness and return synchronicity	Relaxing short sale constraints will lead to improved price discovery and thus higher return synchronicity	Stock return synchronicity is positively associated with price informativeness The price informativeness and the stock return synchronicity of pilot firms increase relative to those of non-pilot firms during the experiment period, but such difference disappeared after the experiment ended.	986 pilot firms, matched to control firms (without replacement)	All
28	Ke, Lo, Sheng, and Zhang (2018), WP	Effect of Reg SHO exp't on quality of analysts' forecast bias and accuracy.	Reg. SHO exp't, leads analysts to expend more effort to uncover bad news; improves price efficiency making it easier for analysts to provide better forecasts	There is a decline forecast bias and in forecast inaccuracy for Pilot firms during the 8 quarters after Reg SHO was implemented.	Balanced sample has 448 Pilot firms and 879 Control firms.	All
29	Kim and Park (2015), WP	Effect of Reg SHO exp't on sensitivity of investment sensitivity to share price	Relaxing short sale constraints leads managers to learn less from share price (why not explained), hence reduces sensitivity of investment decisions to price.	Pilot firms show lower sensitivity of investment to price; lower future operating performance ("theory": pilot firms' managers learn less from their firms' stock prices due to lower PIN)	Balanced panel: 339 pilot and 647 control firms	All
30	Kunzmann and Meier (2018), WP	Effect of Reg SHO exp't on forced CEO turnover	Higher short interest predicts shareholder activism; positive relation between short interest and forced turnover grows stronger with higher activism.	No effect of removing short sale constraints on turnover probability for small firms; higher probability of forced turnover for large firms. Large firm turnover is correlated with short interest.	Small firms: 235 treated; 454 control. Large firms: 252 treated; 493 control	All (large and small separately)
31	Li and Zhang (2015), JAR	Effect of Reg SHO exp't on voluntary disclosure using as proxies management forecast precision and readability of bad news annual reports	Managers of pilot firms respond to increased short-sale threat by reducing disclosure precision, to maintain share price.	Price sensitivity of pilot firms to bad news forecasts increases; no change in sensitivity to good news forecasts.	346 pilot and 711 control firms.	All

No.	Papers	Research Question	Channels	Summary of Results	Sample size	Exchange
32	Lu and Peng (2018), WP	Effect of Reg SHO exp't on measures of corporate tax avoidance	Increased short-sale threat of SS increases monitoring and disciplines managers who respond by reducing corporate tax avoidance.	Pilot firms engage in less tax avoidance and have higher effective tax rate.	579 pilot and 1194 control firms.	All
33	Rusinova, Wernicke, and Bansal (2018), WP	Effect of Reg SHO exp't on corporate social responsibility (CSR), measured by KLD index	To be protected from downward pressure on price by short selling, pilot firms strengthen relations with stakeholders by improving CSR.	1) Pilot firms increase their CSR scores; effect is smaller for firms with higher percentage of transient institutional investors and financially constrained firms	1,682 firms (They do not indicate how many treated firms and controls are among these 1,682 firms)	All
34	Securities and Exchange Commission, Office of Economic Analysis (2007)	Effect of Reg SHO exp't on market quality. Outcomes: Short selling, liquidity (depth and spread), volatility, market efficiency (return reversal and auto-regression), and extreme price changes		1) Pilot firms: higher short sale volume; but no change in short interest or option trading volume 2) Quoted ask depth decreases for NYSE listed pilot firms 3) Pilot firms: higher intraday volatility; 5-minute return reversals and semi-auto-regression, but no change in extreme price movements	NYSE: 504 pilot and 973 control firms. Nasdaq: 439 pilot and 917 control firms	NYSE and Nasdaq, treated separately
35	Sun and Xu (2018), WP	Effect of Reg SHO exp't on readability of 10-K annual reports	Pilot firms reduce readability to reduce short-selling pressure.	Pilot firm 10-K decreases during experiment period; The readability decreases in post exp' when uptick rule was repealed.	630 pilot and 1269 control firms. Balanced sample: 382 pilot and 674 control firms.	All
36	Wang (2018), WP	Effect of Reg SHO exp't on corporate cash holdings	Short-selling pressure reduces share prices to drop and makes external financing more difficult. To avoid being financially constrained, pilot firms hold more cash..	Pilot firms hold more cash; effect more pronounced for firms with more financial constraints, more liquid shares, more short-term investors, and stronger product market competition.	Not stated in the paper	All

No.	Papers	Research Question	Channels	Summary of Results	Sample size	Exchange
37	Wei and Zhang (2018), WP	Effect of Reg. SHO exp't on SEO underpricing.	Stronger short selling threat exposes firms to greater price downside risk and thus induces lower risk-taking, which mitigates negative market reaction to SEOs and reduces SEO underpricing	1) Pilot firms experience smaller SEO underpricing; effect is stronger for riskier firms and firms with higher executive risk-taking incentives, higher default risk, and lower M/B ratio. 2) Pilot firms have lower discretionary accruals, capital expenditure, leverage, and default risk. 3) Pilot firms are more likely to issue seasoned equity and issue more equity. 4) Control firms have smaller SEO underpricing after experiment ends.	221 pilot and 502 control firms conducting SEOs	All
38	Young (2016), Finance Research Letters	Effect of Reg SHO exp't on conditional conservatism	Pilot firms decrease conditional conservatism by delaying recognition of bad earnings because their share prices become more sensitive to bad news.	Pilot firms decrease conditional conservatism during experiment period.	1,905 firms (number of treated and control firms not stated)	All
39	Zhang (2018), WP	Effect of Reg SHO exp't on mispricing and liquidity.	Relaxing short-sale restrictions can facilitate arbitrage, which could correct mispricing or provide liquidity.. Outcomes: measures of abnormal returns	Pilot firms have reduced abnormal returns and increased liquidity; effect more pronounced for firms more likely to have binding short-sale constraints	742 pilot and 1,482 control firms	All

Appendix B. Additional Results for Accruals, F-score, and HF measures

This appendix contains additional results, graphical evidence as well as regressions for various other specifications that are not reported in the paper.

Table A-I: Additional covariate balance results, similar to Table I, but for balanced panel of firms with data on covariates and accruals for the full sample period ("accruals balanced panel").

Table A-II: Summary of results for accruals measures, using balanced panel and FHK sample periods and specification (similar to Table VI), but using different winsorization options.

Table A-III: Analysis of accruals using balanced panel and FHK sample periods and specification, similar to Table VI, except that the sample is "double balanced" and requires data on covariates, accruals, and Fm-score to be available throughout the sample period.

Table A-IV: Analysis of F-score, similar to Table VIII, but using unbalanced panel and our sample periods and specification.

Table A-V: Analysis of F-score, intermediate between Table A-IV and Table VIII, using unbalanced panel but otherwise using FHK sample periods and specification.

Table A-VI: Analysis of HF measure, similar to Table IX, but using unbalanced panel, our sample periods, and otherwise using FHK specification.

Table A-VII: Analysis of HF measure, similar to Table IX, except that the sample is the accruals balanced panel rather than (as in Table IX) a balanced panel based on firms with data on F-score and covariates for all sample years ("F-score balanced panel").

Figure A-1: Love plot showing covariate balance for the accruals balanced panel.

Figure A-2: Annual means for accruals measures, for pilot and control firms over 1998-2010, using unbalanced panel.

Figure A-3: Leads-and-lags graphs for F-score, using F-score balanced panel, FHK sample periods, and FHK specification.

Figure A-4: Leads-and-lags graphs for F-score, using unbalanced panel, our sample periods, and specification.

Figure A-5: Annual means, separately for pilot and control firms, for F-score and HF measures, using unbalanced panel.

Figure A-6: Univariate Means – Pre/During/Post, for pilot and control firms, for F-score and HF measures, similar to Figure 4, but using unbalanced panel.

Figure A-7. Sensitivity analysis for HF-1 and HF-2 measures, using different HF percentile thresholds, similar to Figure 4, but using unbalanced sample and our sample periods.

Some notes on particular Appendix tables and figure are presented below

Annual Means for Accruals Measures

Figure A-2 shows annual means, separately for pilot and control firms, for each of the accruals measures. The pilot and control means closely track each other for operating accruals, total accruals and AA. There is more annual variation for PMDA, consistent with the lower power of this measure, but no apparent trends, and differences during the experiment period are small. The pre-experiment variation provides additional reason to consider PMDA a less-preferred measure than the other accruals measures.

Winsorization Variations for Accruals Measures

In Table A-II, we present results for the four accruals measures, for both unbalanced and balanced panels, using the FHK sample periods and specification, with different winsorization choices. Most of these choices affect AA and PMDA, but not Operating or Total Accruals.

Some observations about PMDA. First, it is sensitive to the decision to winsorize or not. This can be seen in the extremely large coefficients on Pilot*During in Panel B (Do not winsorize when estimating AA and PMDA; but do winsorize in DiD regressions) and Panel D (do not winsorize at all). PMDA is also more sensitive than other measures to the choice between unbalanced and balanced panel, especially without winsorization.

Note on Computation of F-score2 and F-score3.

Dechow et al. (2011) provide precise definitions for most of the variables that enter their F-score measures. An exception is *leasedum*, an off-balance-sheet measure used in F-score2 and F-score3. Dechow et al. (2011) define this variable as existence of operating leases, but do not provide a specific Compustat data item or items to draw this information from. We define *leasedum* as an indicator variable taking value 1 if either MRC1 (firm has operating leases obligations in the next year) *or* MRCT (firm has operating lease obligations during the next five years) is greater than 0, and zero otherwise. We use both because some firms with positive values for MRC1 have missing values for MRCT.

HF Measures: Which "Balanced Panel" to Use?

Table A-VII is similar to text Table IX. The sole difference is that in Table IX, we used an "F-score balanced panel" which includes only firms with data on all F-scores for all years, while in Table A-VII, we use the same "accruals balanced panel" we used to study accruals. If we use the accruals balanced panel, not all firms have all F-scores for all years, so the sample is slightly unbalanced. This minor change meaningfully affects the coefficients on the HF measures. For example, for HF-1 with covariates, the average marginal effect for Pilot*During is -0.0054 in Table IX, but is smaller in magnitude, at -0.0038, in Table A-VII.

More troubling, if one compares the standard errors for Pilot*During with two-way clustering in Table X to those in Table A-VII, Panel B, there is a huge difference between the two. For the HF-1measure with covariates, the s.e. with the F-score balanced panel is .0209, versus .0788 with the accruals balanced panel. This huge difference, from a modest change in sample, casts further doubt on whether one can reliably use two-way clustered s.e.'s in a short panel setting. By contrast, s.e.'s clustered on firm are stable, across samples, both with and without covariates.

Table A-I: Additional Covariate Balance Results

Table is similar to Table I in the text, but for balanced panel with data on covariates and accruals measures for full sample period ("accruals balanced panel"). Sample size for most variables is indicated in the table; sample is slightly smaller for some variables due to missing data on Compustat**.** *, **, *** indicates statistical significance at the 10%, 5%, and 1% levels.

	Fully Treated		**Original Controls**		**Norm. Diff.**	***t*-test for Means**	**Rank-sum test**
Number of firms	**491**		**917**				
	Mean	**Median**	**Mean**	**Median**			
ln(Assets, $M)	6.963	6.785	6.937	6.757	0.017	0.30	0.48
Assets, $M	3,852	884	4,194	860	-0.034	-0.59	0.48
ln(Sales, $M)	6.785	6.822	6.746	6.729	0.022	0.39	0.40
Sales, $M	3,786	909	3,696	828	0.010	0.19	0.40
ln(Market Cap, $M)	7.165	6.916	7.117	6.920	0.033	0.58	0.66
Market Cap, $M	5,009	1,008	4,738	1,012	0.022	0.39	0.66
Tobin's q	2.309	1.828	2.288	1.837	0.016	0.28	-0.22
Short Interest (%) (05/2004-04/2005)	4.783	3.107	4.778	3.251	0.001	0.02	0.33
Capex/Assets	0.048	0.035	0.050	0.033	-0.047	-0.82	0.13
R&D/Sales (win at 1.00)	0.077	0.002	0.081	0.006	-0.023	-0.42	-1.70*
ROA	0.127	0.134	0.118	0.126	0.076	1.34	0.99
Leverage	0.274	0.240	0.280	0.240	-0.026	-0.46	-0.31
Book/Market Ratio	0.406	0.383	0.393	0.359	0.052	0.94	0.93
Trading Volume (win at 99%)	0.186	0.146	0.196	0.152	-0.069	-1.22	-0.92
Beta	1.432	1.373	1.429	1.323	0.005	0.09	0.09
Share Returns (07/2003-06/2004)	0.487	0.347	0.484	0.341	0.005	0.08	-0.08
Share Returns (07/2004-04/2005)	-0.008	-0.038	-0.014	-0.025	0.020	0.35	-0.63
Operating Accruals	-0.054	-0.048	-0.061	-0.052	0.084	1.46	1.54
Total Accruals	0.067	0.030	0.057	0.023	0.052	0.94	0.85
Abnormal Accruals (AA)	0.051	0.026	0.047	0.023	0.022	0.40	0.28
PMDA	-0.029	-0.012	-0.034	-0.014	0.018	0.32	0.56

Table A-II: Summary of Results with Different Options for Winsorization and AA Comparison Group

Regressions are same as in text Table VI, except that we use the indicated choices for winsorization and for the peer group used to compute AA and identify performance matched firms. All winsorization is at 1%/99%, and applies to all continuous variables (both dependent and independent). All regressions include same covariates as in Table VI, pilot dummy, and Pre, During and Post dummies; coefficients on these variables are suppressed. *t*-statistics, using standard errors clustered on firm, are in parentheses. *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Significant results, at 5% level or better, in **boldface**.

		Accruals Type	***Operating Accruals***		***Total Accruals***		***AA***		***PMDA***	
Panel	**Winsorization approach**	Sample	Unbal	Balanced	Unbal	Balanced	Unbal	Balanced	Unbal	Balanced
A	*"Full winsorization (same as table 6)*: (i) winsorize by year, and use all Compustat firms when estimating AA and PMDA; (ii) winsorize in DiD regressions	*Pilot * During*	-0.0045	-0.0065	-0.0010	0.0074	-0.0018	-0.0070	0.0019	0.0005
			(-1.27)	(-1.60)	(-0.13)	(0.98)	(-0.33)	(-1.06)	(0.22)	(0.05)
		*Pilot * Post*	0.0004	-0.0044	-0.0052	-0.0032	-0.0007	-0.0079	0.0021	-0.0016
			(0.11)	(-1.03)	(-0.77)	(-0.43)	(-0.12)	(-1.12)	(0.24)	(-0.16)
		Sign reversal	0.0049	0.0021	-0.0042	-0.0106	0.0011	-0.0009	0.0002	-0.0021
			(1.25)	(0.55)	(-0.57)	(-1.44)	(0.19)	(-0.15)	(0.02)	(-0.21)
B	Do not winsorize when estimating AA and PMDA; but do winsorize in DiD regressions	*Pilot * During*	Same as Panel A				0.0111	0.0415	0.1200	0.1141
							(0.13)	(0.55)	(1.45)	(1.58)
		*Pilot * Post*					-0.1383	-0.0907	0.0341	0.0127
							(-1.80)	(-1.27)	(0.41)	(0.18)
		Sign reversal					-0.1494	-0.1322	-0.0859	-0.1014
							(-1.76)	(-1.48)	(-0.99)	(-1.25)
C	Winsorize when estimating AA and PMDA, do not winsorize AA and PMDA again in DiD regressions (but still winsorize covariates)	*Pilot * During*	Same as Panel A				-0.0014	-0.0050	0.0056	0.0064
							(-0.21)	(-0.57)	(0.51)	(0.45)
		*Pilot * Post*					0.0111	0.0029	0.0110	0.0092
							(1.03)	(0.24)	(0.82)	(0.58)
		Sign reversal					0.0125	0.0079	0.0054	0.0028
							(1.18)	(0.72)	(0.43)	(0.20)
D	No winsorization	*Pilot * During*	0.0019	0.0001	0.0019	0.0116	-0.0236	0.0919	0.1408	**0.2879****
			(0.35)	(0.02)	(0.19)	(0.98)	(-0.22)	(0.98)	(0.99)	**(2.13)**
		*Pilot * Post*	0.0289	0.0228	0.0212	0.0241	-0.1210	0.0032	0.0831	0.1962
			(1.96)	(1.26)	(1.38)	(1.25)	(-1.05)	(0.03)	(0.50)	(1.17)
		Sign reversal	0.0270	0.0227	0.0193	0.0125	-0.0974	-0.0887	-0.0577	-0.0917
			(1.77)	(1.31)	(1.14)	(0.65)	(-0.83)	(-0.69)	(-0.37)	(-0.51)
E	Full winsorization. Same as Panel A, but estimate AA and PMDA using	*Pilot * During*	Same as Panel A				-0.0047	-0.0048	-0.0054	-0.0045
							(-1.47)	(-1.24)	(-1.25)	(-0.85)

	firms in sample, instead of all firms in Compustat.	*Pilot * Post*					-0.0011	-0.0046	-0.0027	-0.0012
							(-0.33)	(-1.13)	(-0.58)	(-0.21)
		Sign reversal					0.0036	0.0002	0.0027	0.0033
							(1.03)	(0.04)	(0.56)	(0.65)
F	Partial winsorization. Same as Panel B, but estimate AA and PMDA using firms in sample, instead of all firms in Compustat.	*Pilot * During*	Same as Panel A				-0.0036	-0.0042	-0.0064	-0.0060
							(-1.04)	(-1.01)	(-1.29)	(-1.01)
		*Pilot * Post*					-0.0003	-0.0041	-0.0041	-0.0024
							(-0.07)	(-0.84)	(-0.69)	(-0.34)
		Sign reversal					0.0033	0.0001	0.0024	0.0036
							(0.80)	(0.03)	(0.40)	(0.58)
G	Partial winsorization. Same as Panel C, but estimate AA and PMDA using firms in sample, instead of all firms in Compustat.	*Pilot * During*	Same as Panel A				-0.0051	-0.0046	-0.0055	-0.0030
							(-1.49)	(-1.03)	(-1.20)	(-0.51)
		*Pilot * Post*					-0.0014	-0.0044	-0.0023	0.0003
							(-0.37)	(-0.92)	(-0.47)	(0.05)
		Sign reversal					0.0037	0.0002	0.0031	0.0033
							(1.01)	(0.06)	(0.63)	(0.62)
H	No winsorization. Same as Panel D, but estimate AA and PMDA using firms in sample, instead of all firms in Compustat.	*Pilot * During*	Same as Panel D				-0.0036	-0.0042	-0.0071	-0.0028
							(-0.82)	(-0.68)	(-1.17)	(-0.35)
		*Pilot * Post*					0.0145	0.0162	0.0183	0.0293
							(1.03)	(0.97)	(0.90)	(1.20)
		Sign reversal					0.0181	0.0204	0.0254	0.0322
							(1.25)	(1.18)	(1.23)	(1.29)

Table A-III: Accruals Measures with "Double Balanced" Sample

Regression specification is same as text Table VI, except the sample is limited to firms with data for full sample period on accruals, covariates, and F-score measures. See Table II and III notes for detailed variable definitions. *t*-statistics, using standard errors clustered on firm, are in parentheses. *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Significant results, at 5% level or better, **in boldface.**

Accruals Type	**Operating**		**Total**		**AA**		**PMDA**	
*Pilot * During*	-0.0057	-0.0055	0.0074	0.0102	-0.0068	-0.0062	-0.0020	-0.0023
	(-1.39)	(-1.37)	(0.87)	(1.32)	(-1.00)	(-0.92)	(-0.19)	(-0.22)
*Pilot * Post*	-0.0043	-0.0041	-0.0074	-0.0061	-0.0090	-0.0088	-0.0042	-0.0044
	(-0.98)	(-0.96)	(-0.93)	(-0.82)	(-1.23)	(-1.23)	(-0.41)	(-0.43)
ln(assets)		**0.0042*** **		0.0004		-0.0028*		0.0000
		(4.45)		(0.31)		(-1.93)		(0.02)
market-to-book		0.0001		**0.0021*** **		**0.0022*** **		-0.0009
		(0.27)		**(2.61)**		**(3.80)**		(-1.03)
ROA		**0.0610*** **		**0.3815*** **		**0.0784*** **		-0.0243
		(4.72)		**(16.04)**		**(4.91)**		(-1.01)
Leverage		**-0.0276*** **		0.0005		**-0.0236*** **		0.0016
		(-4.93)		(0.06)		**(-3.13)**		(0.17)
Pilot dummy	0.0061	0.0056	-0.0006	-0.0018	0.0069	0.0063	0.0020	0.0021
	(1.63)	(1.50)	(-0.09)	(-0.29)	(1.16)	(1.07)	(0.25)	(0.27)
During	**0.0251*** **	**0.0216*** **	**0.0174*** **	0.0083*	**-0.0183*** **	**-0.0196*** **	-0.0005	0.0003
	(10.56)	**(9.10)**	**(3.50)**	(1.83)	**(-4.58)**	**(-4.91)**	(-0.07)	(0.05)
Post	0.0032	0.0016	**-0.0265*** **	**-0.0252*** **	**-0.0105** **	-0.0071	0.0066	0.0059
	(1.20)	(0.63)	**(-5.34)**	**(-5.63)**	**(-2.43)**	(-1.63)	(1.03)	(0.91)
Firm FE	NO							
Year FE	NO							
Firm-Year Obs.	10,638							
Pilot (Control) Firms	419 (763)							
Adjusted R^2	1.5%	3.9%	1.5%	12.3%	0.4%	1.7%	-0.0%	-0.0%

Table A-IV: F-score, Our Sample Periods, Unbalanced Panel, Our Specification

Regression specification is similar to Table VIII in the paper, except that here we use the unbalanced panel, with our specification and sample periods, while Table VIII uses a balanced panel, with the FHK specification and their sample periods. Regression with firm and year FE of indicated dependent variables (F-score1 through F-score3) on Pilot*During, Pilot*Post, and for even numbered regressions, indicated covariates, winsorized at 1%/99%. Sample is based on 2005 Financial Analysis Sample over fiscal years 2001-2010. Regression specification follows eqn. (3), except that the outcome variable is F-score instead of accruals. *t*-statistics, using standard errors clustered on firm, are in parentheses. *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Significant results, at the 5% level or better, in **boldface**.

Dep. variable	***F-score1***		***F-score2***		***F-score3***	
	(1)	(20	(3)	(4)	(5)	(6)
*Pilot * During*	0.0096	0.0046	0.0105	0.0057	0.0051	0.0034
	(0.54)	(0.27)	(0.54)	(0.31)	(0.18)	(0.12)
*Pilot * Post*	-0.0060	0.0003	-0.0034	0.0034	-0.0209	-0.0077
	(-0.30)	(0.02)	(-0.16)	(0.16)	(-0.66)	(-0.25)
Sample mean	0.9874	0.9870	1.0172	1.0168	1.1159	1.1159
ln(assets)		**0.2261*****		**0.2602*****		**0.2668*****
		(14.14)		**(14.67)**		**(10.76)**
market-to-book		-0.0010		-0.0011		0.0028
		(-0.92)		(-1.01)		(1.22)
ROA		**0.6321*****		**0.6592*****		**1.0283*****
		(9.93)		**(9.91)**		**(8.38)**
Leverage		0.0146		0.0437		0.0157
		(0.59)		(1.61)		(0.43)
Firm and Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm-Year Obs.	17,996	17,862	17,739	17,630	17,462	17,435
Pilot (Control) firms	687 (1,380)	687 (1,379)	687 (1,380)	687 (1,379)	686 (1,378)	686 (1,378)
Adjusted R^2	54.2%	58.1%	54.1%	58.3%	46.0%	49.5%

Table A-V: F-score, Our Sample Periods, Unbalanced Panel, Otherwise FHK Specification

This table is an intermediate table between Table A-IV which uses our specification, unbalanced panel, and our sample periods, and text Table VIII, which uses the FHK specification, including their sample periods and balanced panel. In this table, we use the unbalanced panel and our sample periods, but otherwise use the FHK specification, including replacing firm FE with pilot dummy, and replacing year FE with Pre, During, and Post dummies.

Regression of indicated dependent variables (F-score1 through F-score3) on Pilot*During, Pilot*Post, and for even numbered regressions, indicated covariates, winsorized at 1%/99%. Sample is based on 2005 Financial Analysis Sample over fiscal years 2001-2010. Regression specification is the same as Table A-III, except that: (1) we drop firm fixed effects and add Pilot dummy (which would otherwise be absorbed by the firm FE); and (2) we replace year fixed effects with Pre, During, and Post dummy variables.. *t*-statistics, using standard errors clustered on firm, are in parentheses. *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Significant results, at 5% level or better, in **boldface**.

Dep. variable	***F-score1***		***F-score2***		***F-score3***	
*Pilot * During*	-0.0091	-0.0034	-0.0116	-0.0046	-0.0152	-0.0058
	(-0.47)	(-0.18)	(-0.54)	(-0.22)	(-0.53)	(-0.21)
*Pilot * Post*	-0.0151	-0.0056	-0.0134	-0.0022	-0.0258	-0.0125
	(-0.75)	(-0.28)	(-0.62)	(-0.10)	(-0.84)	(-0.42)
ln(assets)		**0.0198*****		**0.0238*****		0.0131
		(3.12)		**(3.54)**		(1.54)
market-to-book		**-0.0046*****		**-0.0055*****		-0.0036
		(-3.44)		**(-3.95)**		(-1.54)
ROA		**0.6278*****		**0.6367*****		**0.8526*****
		(13.52)		**(12.90)**		**(12.20)**
Leverage		0.0062		0.0212		0.0272
		(0.24)		(0.78)		(0.82)
Pilot dummy	0.0085	0.0009	0.0062	-0.0029	0.0169	0.0043
	(0.38)	(0.04)	(0.27)	(-0.13)	(0.53)	(0.14)
During	**0.0657*****	**0.0525*****	**0.0798*****	**0.0631*****	0.0209	0.0074
	(5.74)	**(4.74)**	**(6.52)**	**(5.30)**	(1.31)	(0.48)
Post	**-0.0520*****	**-0.0729*****	**-0.0434*****	**-0.0682*****	**-0.0824*****	**-0.0990*****
	(-4.32)	**(-6.21)**	**(-3.42)**	**(-5.42)**	**(-4.84)**	**(-5.87)**
Intercept	**0.9829*****	**0.7965*****	**1.0074*****	**0.7928*****	**1.1303*****	**0.9541*****
	(76.40)	**(20.09)**	**(74.77)**	**(18.84)**	**(64.62)**	**(17.94)**
Firm-Year Obs.	17,996	17,862	17,739	17,630	17,462	17,435
Pilot (Control) firms	687 (1,380)	687 (1,379)	687 (1,380)	687 (1,379)	686 (1,378)	686 (1,378)
Adjusted R^2	0.6%	5.2%	0.6%	5.0%	0.3%	4.0%

Table A-VI: HF-Measure, Our Sample Periods, Unbalanced Panel, Otherwise FHK Specification

Table A-VI is similar to Table IX in the paper, except that it uses unbalanced panel and our sample periods, while Table IX uses the FHK specification, sample periods and balanced panel. Average marginal effects from probit regressions, for dependent variables HF-1 through HF-3, respectively, on Pilot*During, Pilot*Post, and for even numbered regressions, indicated covariates, winsorized at 1%/99%. Sample is based on 2005 Financial Analysis Sample over fiscal years 2001-2010. *z*-statistics, using standard errors clustered on firm, are in parentheses. *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Significant results, at 5% level or better, in **boldface**.

Dep. variable	***HF-1***		***HF-2***		***HF-3***	
*Pilot * During*	-0.0005	-0.0004	0.0006	0.0007	-0.0016	-0.0013
	(-0.13)	(-0.13)	(0.18)	(0.21)	(-0.41)	(-0.35)
*Pilot * Post*	0.0061	0.0069	0.0022	0.0020	-0.0003	0.0008
	(1.35)	(1.52)	(0.49)	(0.46)	(-0.06)	(0.17)
ln(assets)		-0.0009		-0.0005		-0.0015**
		(-1.39)		(-0.70)		(-2.16)
market-to-book		-0.0003		-0.0003		-0.0001
		(-1.31)		(-1.30)		(-0.53)
ROA		0.0394***		0.0484***		0.0379***
		(4.02)		(4.73)		(4.25)
Leverage		-0.0023		0.0020		-0.0008
		(-0.69)		(0.62)		(-0.23)
Pilot dummy	-0.0013	-0.0015	-0.0000	-0.0003	0.0004	-0.0001
	(-0.57)	(-0.66)	(-0.00)	(-0.13)	(0.16)	(-0.02)
During	-0.0007	-0.0003	-0.0008	-0.0008	-0.0032	-0.0027
	(-0.34)	(-0.18)	(-0.38)	(-0.39)	(-1.45)	(-1.31)
Post	-0.0134***	-0.0131***	-0.0113***	-0.0104***	-0.0097***	-0.0091***
	(-4.21)	(-4.17)	(-3.69)	(-3.50)	(-3.05)	(-2.91)
Firm-Year Obs.	17,996	17,862	17,739	17,630	17,462	17,435
Pilot (Control) firms	687 (1,380)	687 (1,379)	687 (1,380)	687 (1,379)	686 (1,378)	686 (1,378)
Pseudo R^2	1.6%	5.0%	1.4%	5.5%	1.2%	4.5%

Table A-VII: HF Measure, Using Balanced Panel for Accruals Measures, FHK Sample Period and FHK Specification

Panel A. Specification is similar to Table IX, except that the sample is the balanced panel we used for studying accruals (firms with data on PMDA and covariates for all sample years), rather than (as in Table IX) a balanced panel based on firms with data on F-score and covariates for all sample years. As a result, the sample is slightly unbalanced for F-score and the HF measure. Table reports average marginal effects from probit regressions in lieu of probit coefficients. *z*-statistics, using standard errors clustered on firm, are in parentheses. *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Significant results, at 5% level or better, in **boldface**. ,

	HF-1		***HF-2***		***HF-3***	
*Pilot * During*	-0.0039	-0.0038	-0.0042	-0.0043	0.0009	0.0006
	(-0.93)	(-0.93)	(-0.99)	(-1.02)	(0.21)	(0.15)
*Pilot * Post*	-0.0039	-0.0039	-0.0008	-0.0012	0.0013	0.0009
	(-0.76)	(-0.75)	(-0.16)	(-0.23)	(0.24)	(0.17)
ln(assets)		0.0002		0.0007		-0.0008
		(0.35)		(1.02)		(-1.00)
market-to-book		**-0.0009*****		-0.0007*		-0.0002
		(-2.70)		(-1.90)		(-0.70)
ROA		**0.0397*****		**0.0495*****		**0.0424*****
		(3.50)		**(3.94)**		**(3.70)**
Leverage		-0.0064		-0.0013		-0.0000
		(-1.55)		(-0.32)		(-0.00)
Pilot dummy	0.0046	0.0043	0.0038	0.0037	-0.0008	-0.0006
	(1.38)	(1.30)	(1.07)	(1.04)	(-0.24)	(-0.18)
During	0.0049*	0.0043*	0.0064**	0.0056**	-0.0011	-0.0012
	(1.91)	(1.70)	(2.34)	(2.10)	(-0.45)	(-0.48)
Post	-0.0058*	-0.0060*	-0.0049	-0.0048	-0.0082**	-0.0072**
	(-1.76)	(-1.81)	(-1.42)	(-1.41)	(-2.32)	(-2.06)
Firm-Year Obs.	11,605	11,605	11,491	11,491	11,430	11,430
Pilot (Control) firms	455 (842)	455 (842)	455 (842)	455 (842)	455 (842)	455 (842)
Pseudo R^2	1.7%	4.3%	1.6%	5.2%	0.9%	3.8%

Panel B. One-way vs. two-way clustered standard errors

Probit coefficients for indicated variables from same probit regressions as in Table A-VI (which reports average marginal effects), except that regression are run with both (i) standard errors clustered on firm, and (ii) standard errors clustered separately on firm and year. Coefficients on covariates are suppressed. *, **, *** indicates significance at the 10%, 5%, and 1% level, respectively. Significant results, at 5% level or better, in boldface.

	HF-1		***HF-2***		***HF-3***	
*Pilot * During*	-0.1463	-0.1478	-0.1593	-0.1672	0.0343	0.0249
s.e (cluster on firm)	[0.1580]	[0.1605]	(0.1616)	(0.1649)	(0.1648)	(0.1675)
s.e. (2-way cluster)	[0.0829]*	[0.0788]*	(0.0872)*	**(0.0794)****	(0.1208)	(0.1341)
*Pilot * Post*	-0.1490	-0.1498	-0.0313	-0.0454	0.0477	0.0339
s.e (cluster on firm)	[0.1974]	[0.2002]	(0.1929)	(0.1967)	(0.1989)	(0.2024)
s.e. (2-way cluster)	[0.1328]	[0.1233]	(0.1303)	(0.1241)	(0.1069)	(0.1144)
Covariates	NO	YES	NO	YES	NO	YES

Figure A-1: Love Plot for the Balanced Sample

Figure provides a graphical overview of covariate balance for the balanced sample. It shows t-statistics for differences between treated and original control firms, for the variables listed in Panel A of Table I in the paper. Vertical lines indicate t-statistics of at -1.96, 0, and +1.96.

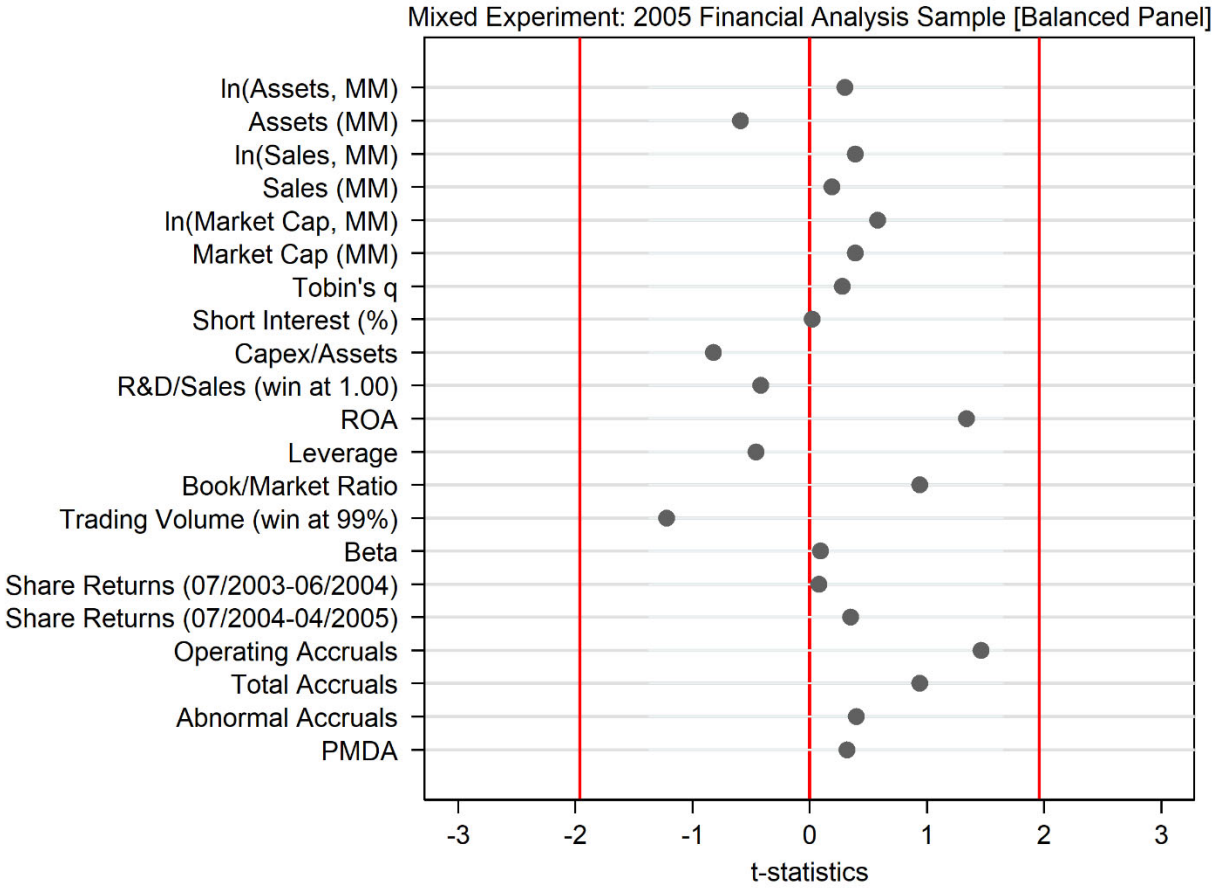


Figure A-2: Univariate Means for Pilot and Control Firms

Figures show annual means, separately for pilot and control firms, for each accruals measure, over 1998-2010. Sample is same as in Table II (unbalanced panel); sample period is extended back to 1998.

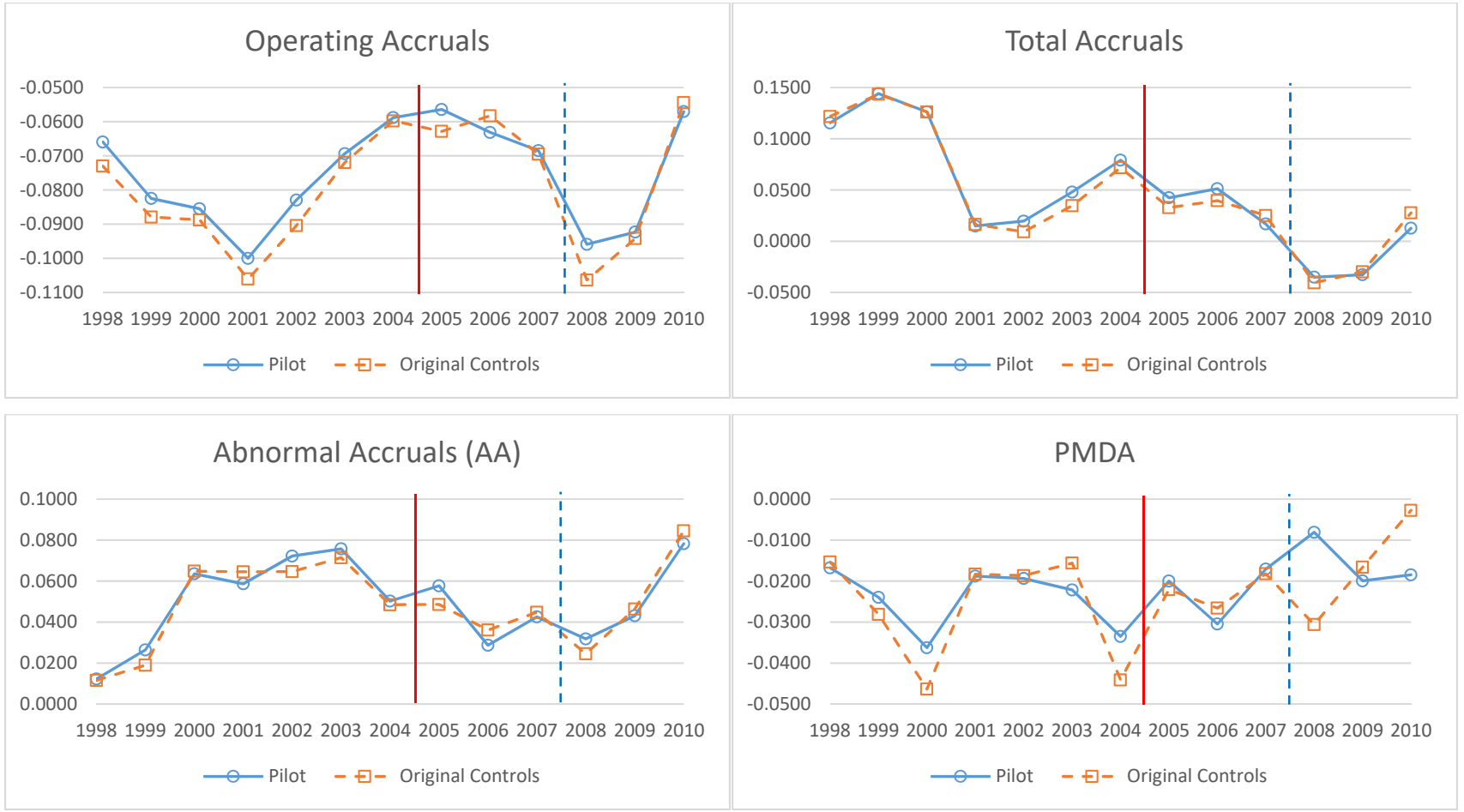


Figure A-3: Lead and Lag graphs for F-score, Balanced Panel

Sample and Specification Similar to Table VIII: FHK F-score Balanced Panel, Sample Periods, and Specification

Leads and lags regressions of F-score, from regressions following eqn. (5), with firm and year fixed effects, over FHK sample period (calendar years 2001-2010, omitting 2004). Variables and sample are same as in Table VIII. y-axis shows coefficients on the lead and lag dummies; vertical bars show 95% CIs around coefficients, using standard errors clustered on firm. Coefficient for 2003 is set to zero.

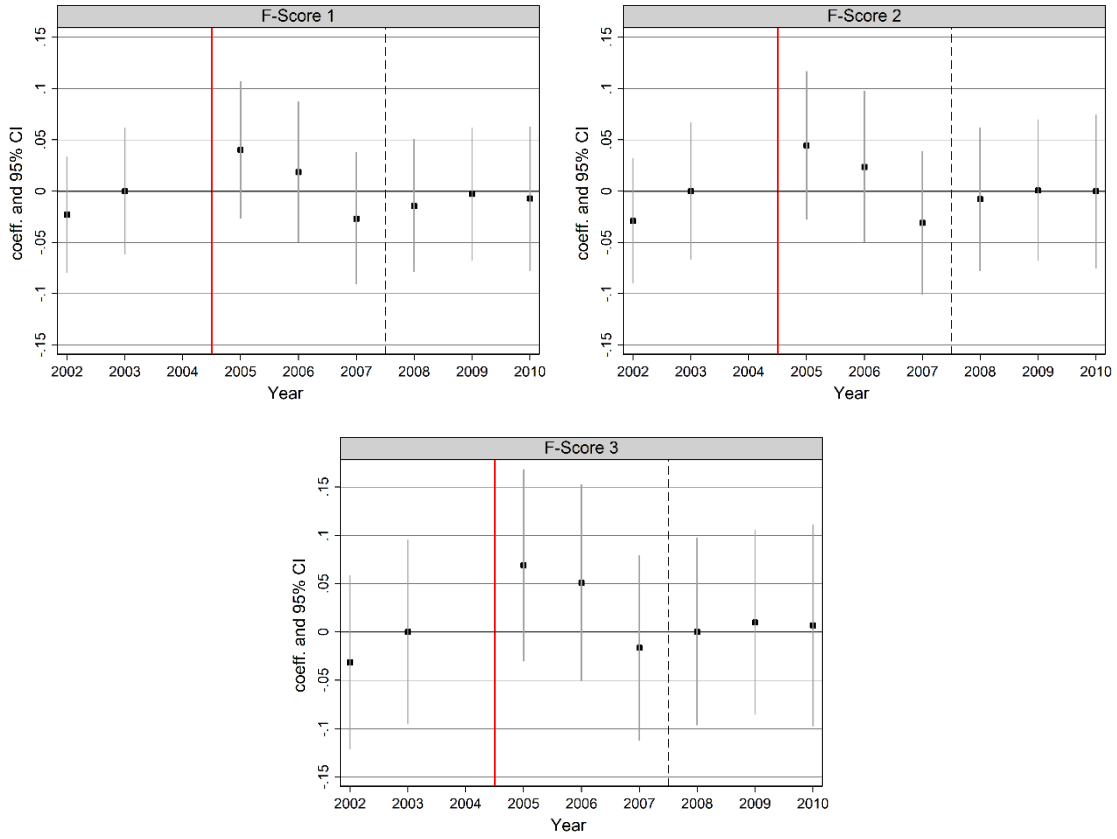


Figure A-4: F-score, Leads-and-Lags

Sample and Specification Comparable to Table A-IV: Unbalanced Panel, Our Sample Periods and Specification

Leads and lags regressions of F-score, from regressions following eqn. (5), with firm and year fixed effects, over fiscal years 2001-2010. Variables and sample are same as in Table A-IV. y-axis shows coefficients on the lead and lag dummies; vertical bars show 95% CIs around coefficients, using standard errors clustered on firm. Coefficient for 2004 is set to zero. Vertical lines indicate start and end of experiment period.

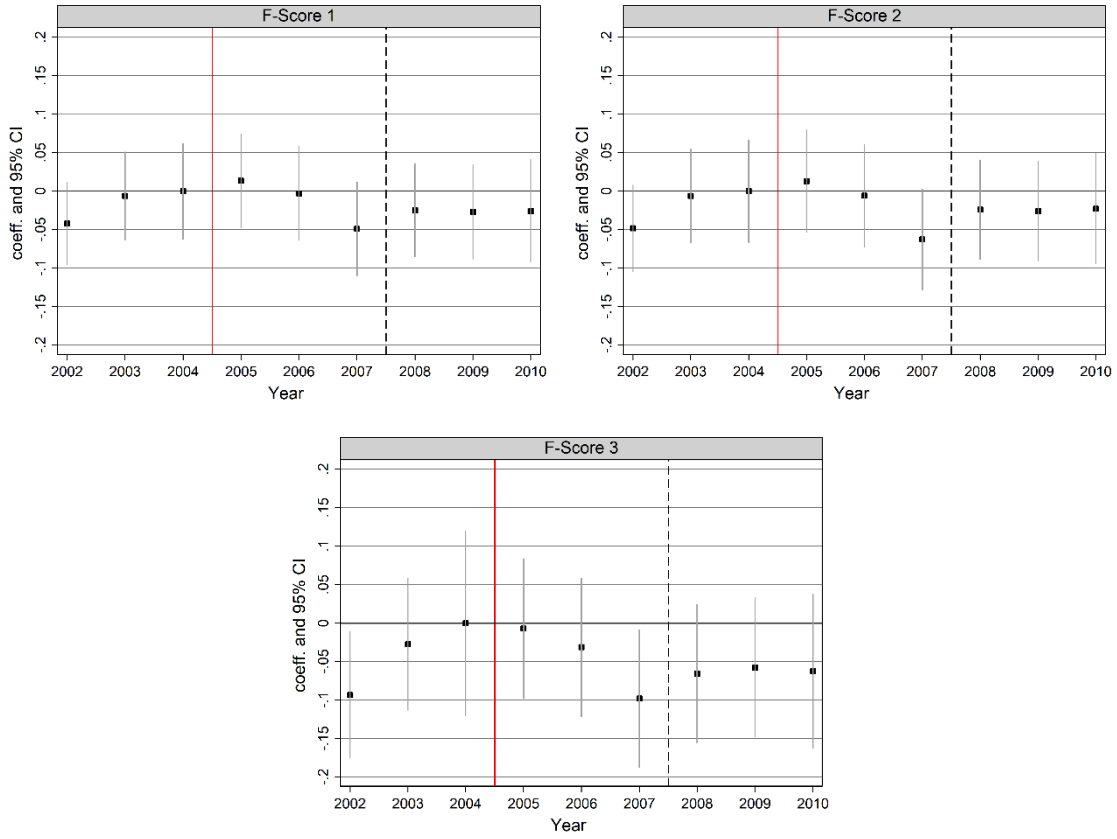
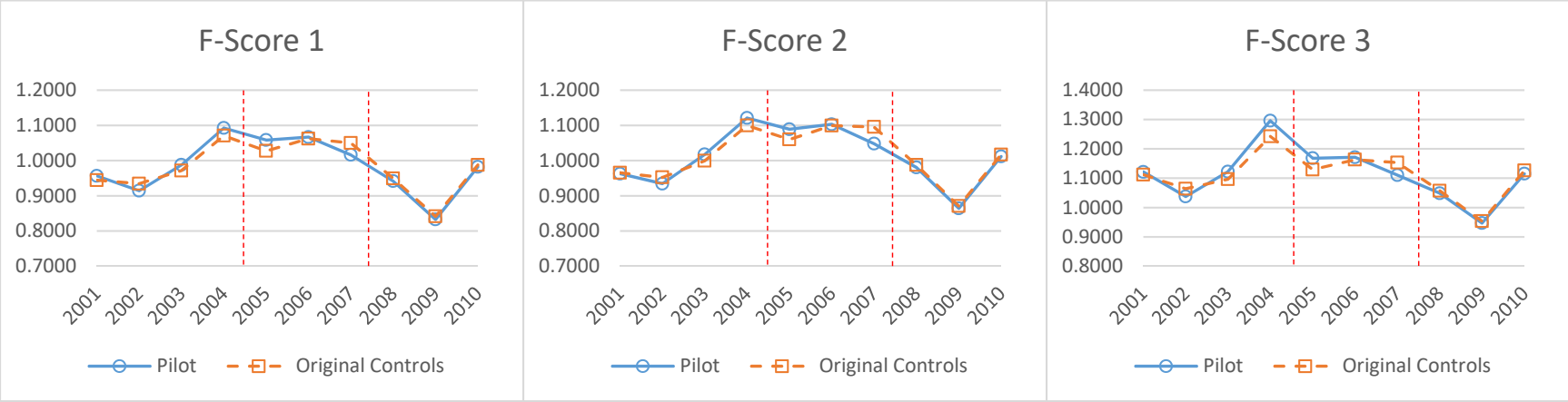


Figure A-5: Univariate Means – Annual

Figures show annual means, separately for pilot and control firms, for each F-score and HF measure, over 2001-2010, using unbalanced panel.**Panel A. Comparable to Table A-IV: F-score, Our Sample Periods, Unbalanced Panel, Our Specification**



Panel B. Comparable to Table A-VI: HF-Score, Our Sample Periods, Unbalanced Panel, Otherwise FHK Specification

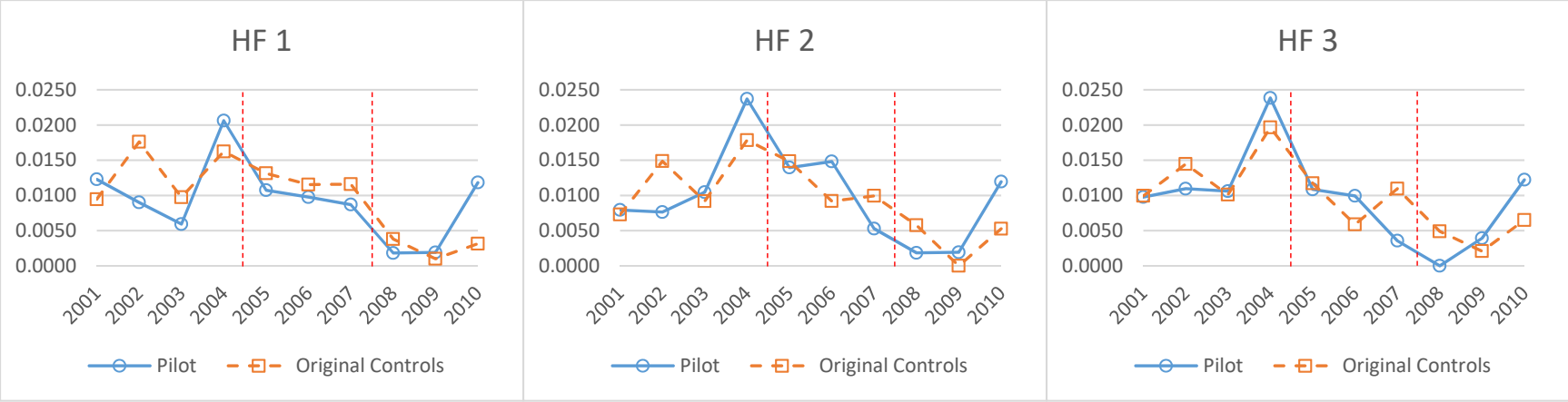
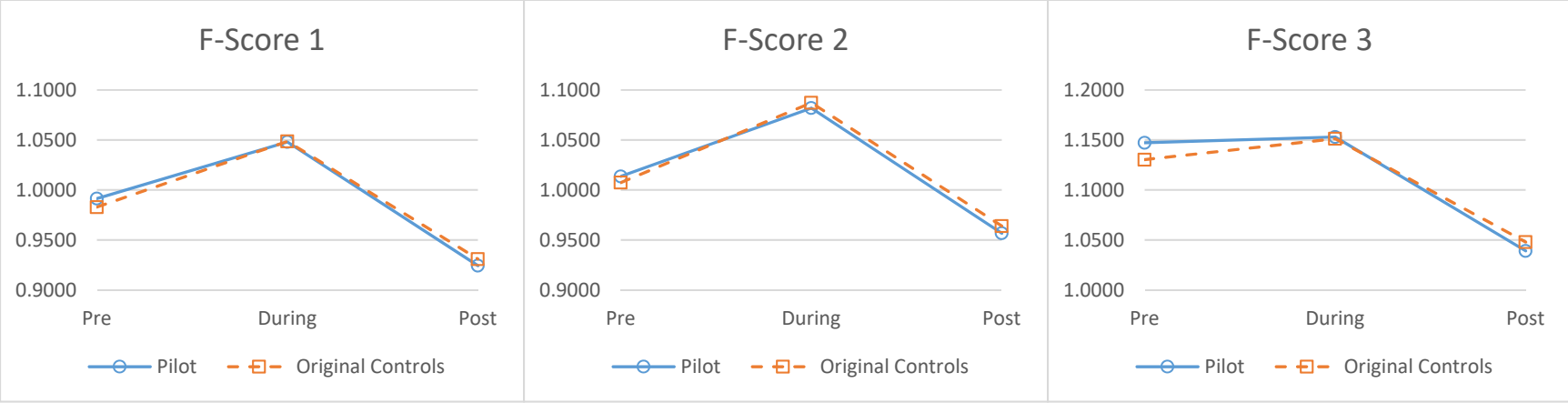


Figure A-6: Univariate Means – Pre/During/Post

Panel A. Comparable to Table A-IV: F-score, Our Sample Periods, Unbalanced Panel, Our Specification

Figures are similar to text Figure 3, Panel A, and show means for each F-score measure, during the Pre, During, and Post periods, separately for pilot and control firms, except using unbalanced panel, our sample periods, and specification.



Panel B. Comparable to Table A-VI: HF-Score, Our Sample Periods, Unbalanced Panel, Otherwise FHK Specification

Figures are similar to text Figure 3, Panel B, and show means for each HF measure, during the Pre, During, and Post periods, separately for pilot and control firms, except using unbalanced panel and our sample periods.

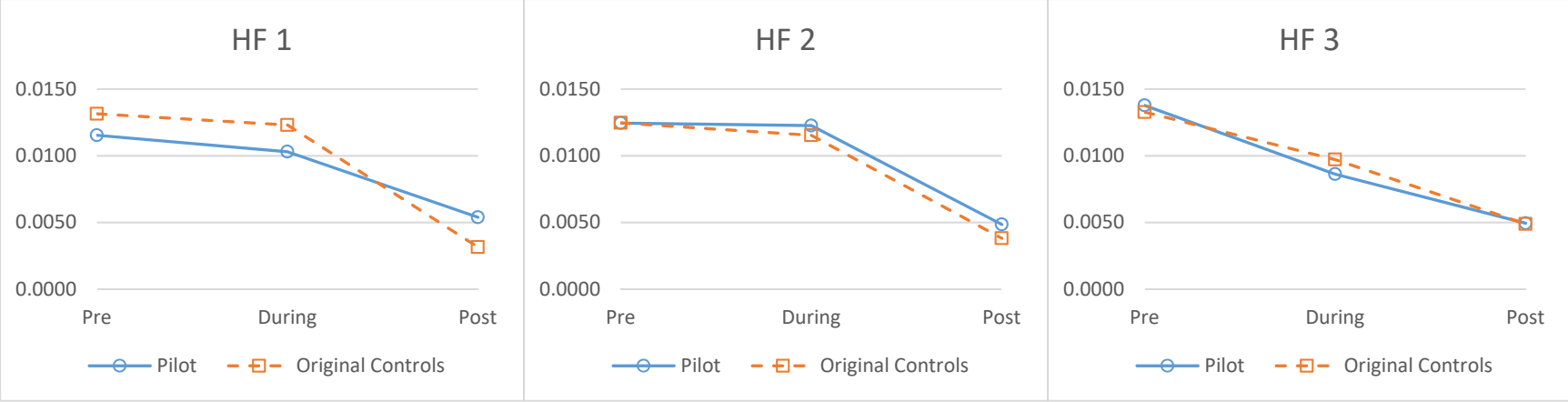
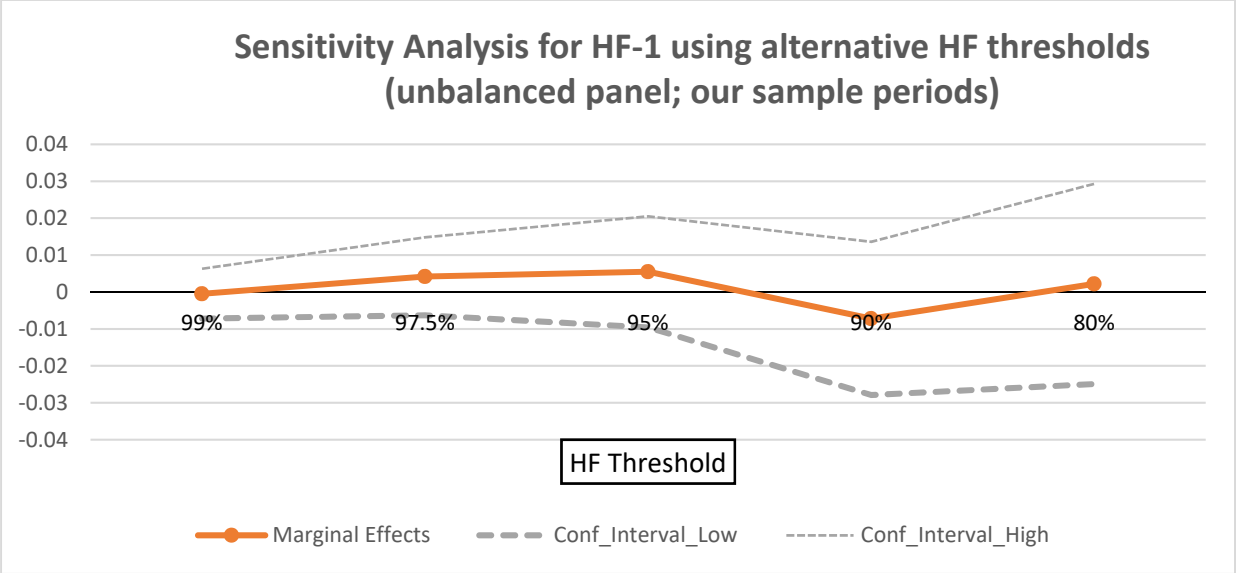
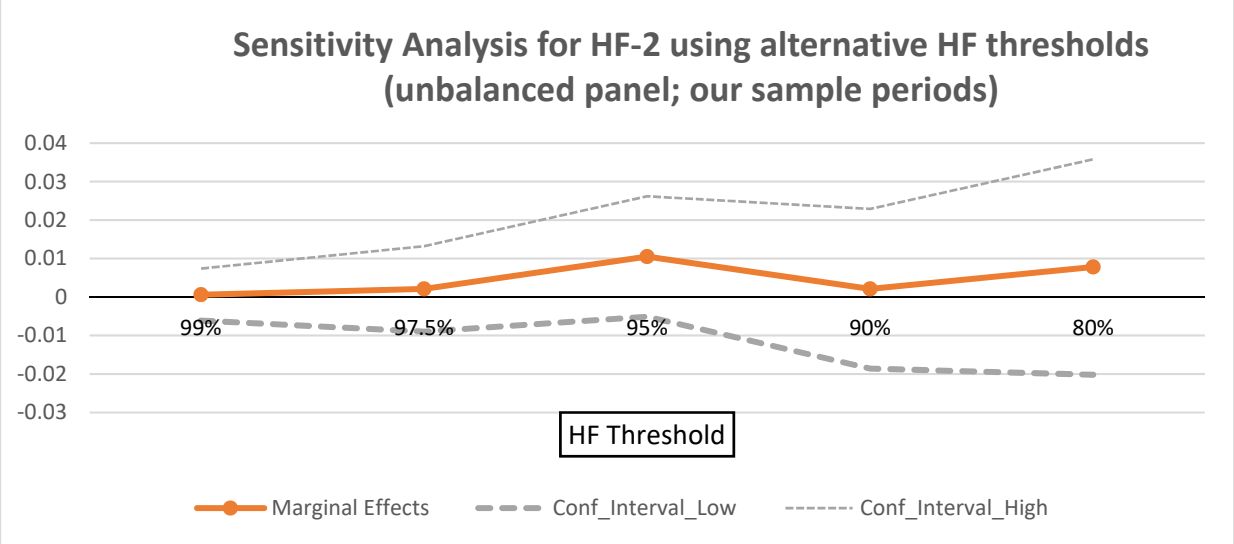


Figure A-7: Sensitivity Analysis using alternative HF thresholds

Figure is similar to Figure 4 in the paper, except it uses the unbalanced panel and our sample periods, instead of the FHK balanced panel and sample periods. Graph shows average marginal effects on *Pilot*During* from probit regression without covariates, with variations on the HF-1 and HF-2 measures using indicated thresholds as the dependent variables. Circles show point estimates. Upper dotted and lower dashed line show 95% confidence interval. Threshold percentiles for F-score1, for setting HF=1 values are 99% (used by FHK), 97.5%, 95%, 90%, and 80% respectively.





Paper Included in Appendix A

Albertus, James F., Andrew Bird, Stephen A. Karolyi, and Thomas Ruchti (2017), Short selling governance and intrafirm resource allocation, working paper, at https://business.uoregon.edu/files/media/web-finance2017-albertus-bird-karolyi-ruchti.pdf.

Alexander, Gordon J. and Mark A. Peterson (2008), The Effect of Price Tests on Trader Behavior and Market Quality: An Analysis of Regulation SHO, *Journal of Financial Markets* 11, 84-111.

Bai, John J., Eunju Lee, and Chi Zhang (2018), Capital Market Frictions and Human Capital Investment: Evidence from Workplace Safety around Regulation SHO, working paper, at https://ssrn.com/abstract=3187686.

Bai, Lynn (2008), The Uptick Rule of Short Sale Regulation—Can it Alleviate Downward Price Pressure from Negative Earnings Shocks, *Rutgers Business Law Journal* 5, 1-63.

Bennett, Benjamin, and Zexi Wang (2018), Do Short Sellers Cause CEOs to Be Fired: Evidence from a Randomized Experiment, working paper, at http://ssrn.com\abstract=3135483.

Bhattacharya, Neil, Theodore E. Christensen, Qunfeng Liao, and Bo Ouyang (2019), Can Short Sellers Constrain Opportunistic Non-GAAP Earnings Reporting?, working paper, at https://ssrn.com/abstract=2724935.

Billett, Matthew T., Fangzhou Liu, and Xuan Tian (2018), Information Spillovers and Cross Monitoring between the Stock Market and Loan Market, *Kelley School of Business Research Paper* 16-20, at https://ssrn.com/abstract=2732784.

Boehmer, Ekkehart, Charles M. Jones, and Ziaoyan Zhang (2018), Potential Pilot Problems: Treatment Spillovers in Financial Regulatory Experiments, working paper, at https://ssrn.com/abstract=2621598.

Cai, Tianyu, and Lixiong Guo (2018), Short Selling and Real Earnings Management, working paper, at http://fmaconferences.org/SanDiego/Papers/Short%20Selling%20and%20REM.pdf

Chang, Yen-Cheng, Minjie Huang, Yu-Siang Su, and Kevin Tseng (2018), Short-Termist CEO Compensation in Speculative Markets: A Controlled Experiment, working paper, at https://ssrn.com/abstract=3002525

Chen, Hang, Yushu Zhu, and Liang Chang (2017), Short-selling constraints and corporate payout policy, *Accounting & Finance* 1-33.

Chen, Huimin and Qiang Wu (2019), Short Selling Threat and Real Activity Manipulation: Evidence from a Natural Experiment, working paper.

Cheng, Mei and Eliza Xia Zhang (2019), Short Sellers and Informativeness of Credit Ratings: Evidence from a Natural Experiment, working paper.

Cheng, Qiang, Xia Chen, Ting Luo, and Heng Yue (2014), Short sellers and corporate disclosures, working paper, at https://ink.library.smu.edu.sg/soa_research/1306.

Choi, Hae Mi (2018), Short selling and the rounding of analysts' forecasts, *Finance Research Letters* 25, 47-54.

Chu, Yongqiang, David A. Hirshleifer, and Liang Ma (2017), The Causal Effect of Limits to Arbitrage on Asset Pricing Anomalies, working paper, at http://ssrn.com/abstract=3092992.

Clinch, Greg, Wei Li, and Yunyan Zhang (2018), Short selling and firms' disclosure of bad news: Evidence from Regulation SHO, *Available at SSRN 2594717*.

De Angelis, David, Gustavo Grullon, and Sebastien Michenaud (2017), The Effects of Short-Selling Threats on Incentive Contracts: Evidence from an Experiment, *Review of Financial* Studies 30, 1627-1659.

Diether, Karl, Kuan-Hui Lee, and Ingrid Werner (2009), It's SHO Time! Short-Sale Price-Tests and Market Quality, *Journal of Finance* 64, 37-73.

Fang, Vivien W., Allen Huang, and Jonathan Karpoff (2016), Short Selling and Earnings Management: A Controlled Experiment, *Journal of Finance* 71, 1251-1293.

Francis, Bill B., Gilna Samuel, and Qiang Wu (2017), The Impact of Financial Markets on Payout Policy: Evidence from Short Selling, working paper, at http://ssrn.com/abstract=3038300.

Grullon, Gustavo, Sebastien Michenaud, and James Weston (2015), The Real Effects of Short-Selling Constraints, *Review of Financial Studies* 28, 1737-1767.

He, Jie (Jack), and Xuan Tian (2016), Do Short Sellers Exacerbate or Mitigate Managerial Myopia? Evidence from Patenting Activities, working paper, at http://ssrn.com/abstract=2380352.

Hope, Ole-Kristian, Danqi Hu, and Wuyang Zhao (2016), Third-Party Consequences of Short-Selling Threats: The Case of Auditor Behavior, *Journal of Accounting and Economics* 63, 479-498.

Hribar, Paul, and D. Craig Nichols (2007), The Use of Unsigned Earnings Quality Measures in Tests of Earnings Management, *Journal of Accounting Research* 45, 1017-1053.

Hu, Danqi, Wei Li, Scott Liao, and YunYan Zhang (2019), The Impact of Short Selling on Bank Holding Companies' Loss Recognition and Risk Taking, working paper.

Kan, Shuo, and Stephen Gong (2018), Does High Stock return synchronicity Indicate High or Low Price Informativeness? Evidence from a regulatory experiment, *International Review of Finance* 18, no. 4: 523-546.

Ke, Yun, Kin Lo, Jinfei Sheng, and Jenny Li Zhang (2018), Does Short Selling Improve Analyst Forecast Quality?, working paper.

Kim, Pureum, and You-il (Chris) Park (2015), Regulatory Short Selling Constraints and the Sensitivity of Corporate Investment to Price, working paper.

Kunzmann, Anja, and Kristina Meier (2017) Does the Board Learn from Short Sellers? Evidence from CEO Turnovers, working paper.

Li, Yinghua and Liandong Zhang (2015), Short Selling Pressure, Stock Price Behavior, and Management Forecast Precision: Evidence from a Natural Experiment, *Journal of Accounting Research* 53, 79-117.

Lu, Yao, and Zhang Peng (2018), Monitoring from Capital Market and Corporate Tax Avoidance: Evidence from Short Selling Pilot Program, working paper.

Rusinova, Vanya, Georg Wernicke, and Tima Bansal (2018), Short-selling and firm performance on Corporate Social Responsibility: Evidence from a natural experiment, working paper.

Securities and Exchange Commission, Office of Economic Analysis (2007), Economic Analysis of the Short Sale Price Restrictions under the Regulation SHO Pilot, at https://www.sec.gov/news/studies/2007/regshopilot020607.pdf.

Sun, Minxing, and Weike Xu (2018), Short Selling and Readability in Financial Disclosures: A Controlled Experiment, working paper.

Wang, Zexi (2018), Short sellers, institutional investors, and corporate cash holdings, working paper, at https://ssrn.com/abstract=2410239.

Wei, Siqi, and Jun Zhang (2018), Firms' Risk Incentives and SEO Underpricing: Evidence from a Natural Experiment, working paper.

Young, Alex (2016), Capital market frictions and conservative reporting: Evidence from short selling constraints, *Finance Research Letters* 17, 227-234.

Zhang, Liang (2018), Can Hedge Funds Correct Mispricing and Provide Liquidity? Evidence from Reg SHO, working paper.