

**Internet Appendix for  
“Do short sellers affect corporate innovation? Evidence from a policy experiment”**

**Table A1: Randomization tests based on 5,000 simulated samples for the baseline DiD analysis**

This table reports results for randomization tests based on 5,000 simulated samples for the baseline DiD analysis. For each simulation, we draw a random sample of 432 “pilot” firms from the pool of actual pilot and non-pilot firms in the event year (2004), and then treat the rest of the pool (793 of them) as “non-pilot” firms. We then perform the DiD tests as in Columns (2), (4), (6) of Table 3 Panel A on this simulated sample. We repeat the simulation process 5,000 times and summarize the distributions of the coefficients and t-stats for the main variable of interest, *Pilot\*Post*.

Variable	Mean	P25	Median	P75	S.D.	N
<i>NSCitePat</i>						
Coefficient before <i>Pilot*Post</i>	0.000	-0.024	0.000	0.025	0.035	5000
T-stat for <i>Pilot*Post</i>	0.016	-0.685	-0.008	0.712	0.995	5000
<i>Value</i>						
Coefficient before <i>Pilot*Post</i>	-0.001	-0.022	0.000	0.020	0.031	5000
T-stat for <i>Pilot*Post</i>	-0.015	-0.697	0.000	0.656	1.006	5000
<i>NSCiteRD</i>						
Coefficient before <i>Pilot*Post</i>	0.000	-0.024	0.000	0.025	0.037	5000
T-stat for <i>Pilot*Post</i>	0.015	-0.656	0.005	0.690	0.998	5000

**Table A2: Cross-sectional tests based on transient institutional ownership**

This table reports the results of the difference-in-differences (DiD) tests in the two subsamples based on whether a firm's ex-ante transient institutional ownership (measured in the year 2003) is above or below the sample median. Definitions of variables are listed in the Appendix. The same set of control variables used in Table 3 are included in all regressions. Each regression includes a separate intercept as well as firm and year fixed effects. Standard errors clustered by firm are displayed in parentheses. The last two rows report the Chi-squared ( $\chi^2$ ) test statistics and the corresponding p-values for the difference in the DiD estimators between the subsamples with high and low transient institutional ownership. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent levels, respectively.

<i>Dep. Var.</i>	<i>NSCitePat<sub>t</sub></i>		<i>Value<sub>t</sub></i>		<i>NSCiteRD<sub>t</sub></i>	
	High Transient	Low Transient	High Transient	Low Transient	High Transient	Low Transient
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Pilot*Post</i>	0.110** (0.048)	0.053 (0.038)	0.097** (0.043)	0.015 (0.038)	0.075** (0.031)	0.056 (0.045)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	4,296	4,289	4,296	4,289	4,296	4,289
R-sq.	0.800	0.830	0.800	0.867	0.790	0.800
$\chi^2$	2.560		4.254		0.479	
P-value	0.110		0.039		0.489	

**Table A3: Summary statistics for litigation risk and stock price informativeness variables**

This table reports the summary statistics of the litigation risk and stock price informativeness variables. *PatentCase\_SS* (*PatentCase\_All*) is a dummy variable that equals one if a firm experiences any patenting-related lawsuits filed by potential short-sellers (all types of plaintiffs) in a given year, and zero otherwise. These two variables are summarized using the Compustat universe (with available information) over the pre-event period (i.e., three years before Regulation SHO). *LitigationRisk\_SS* (*LitigationRisk\_All*) is the fitted probability (estimated following Cohen, Gurun, and Kominers (2019)) that a firm will face a patenting-related lawsuit filed by potential short-sellers (all types of plaintiffs) in a given year.  $1 - R^2$  is one minus the r-squared from regressing a firm's daily stock returns on a constant, the CRSP value-weighted market returns, and the returns of the three-digit SIC industry portfolio in a given year. *PIN* is a firm's probability of informed trading in a given year provided by Brown and Hillegeist (2007).

Panel A: Summary statistics for litigation risk variables

	Mean	S.D.	P25	Median	P75	N
<i>PatentCase_SS</i>	0.001	0.038	0.000	0.000	0.000	15,834
<i>PatentCase_All</i>	0.019	0.135	0.000	0.000	0.000	15,834
<i>LitigationRisk_SS</i>	0.424	0.287	0.190	0.369	0.650	8557
<i>LitigationRisk_All</i>	4.747	3.459	1.833	3.741	7.823	8557

Panel B: Summary statistics for stock price informativeness variables

	Mean	S.D.	P25	Median	P75	N
$1 - R^2$	0.672	0.242	0.541	0.731	0.857	8,977
<i>PIN</i>	0.146	0.078	0.097	0.129	0.179	8,942

**Table A4: Testable hypotheses and the expected effects on firms' patent quality, quantity, and innovation investment**

This table outlines the three testable hypotheses and their predictions regarding the effects of Regulation SHO on pilot firms' patent quality, quantity, and innovation investment. The *disciplining hypothesis* predicts an increase in patent quality but has no clear implications for patent quantity or innovation investment. The *strategic patenting* argument predicts an increase in patent quality, a decrease in patent quantity, and has no clear implications for innovation investment. The *managerial learning* argument predicts an increase in patent quality, but a decreases in patent quantity and innovation investment.

Hypothesis	Patent quality	Patent quantity	Innovation investment
<i>Disciplining hypothesis</i>	+	N/A	N/A
<i>Strategic patenting hypothesis</i>	+	-	N/A
<i>Managerial learning hypothesis</i>	+	-	-

**Table A5: DiD tests for new product announcements**

This table reports the results of the difference-in-differences (DiD) tests on a firm's new product announcement. *AnnCAR* is a firm's sum of all positive cumulative abnormal returns (CAR) around product announcements in a given year. *MajorAnn* is a firm's log number of announcements with the CAR above the 75<sup>th</sup> percentile in a given year. Definitions of other variables are listed in the Appendix. The same set of control variables used in Table 3 are included in all regressions. Each regression includes a separate intercept as well as firm and year fixed effects. Standard errors clustered by firm are displayed in parentheses. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent levels, respectively.

<i>Dep. Var.</i>	<i>AnnCAR<sub>t</sub></i>		<i>MajorAnn<sub>t</sub></i>	
	(1)	(2)	(3)	(4)
<i>Pilot*Post</i>	0.005** (0.002)	0.004* (0.002)	0.022* (0.013)	0.019 (0.013)
Controls	No	Yes	No	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	8,718	8,585	8,718	8,585
R-squared	0.537	0.546	0.515	0.521

**Table A6: DiD tests for stock price informativeness**

This table reports the results of the difference-in-differences (DiD) tests on firms' stock price informativeness measures.  $1 - R_t^2$  is one minus the r-squared from regressing a firm's daily stock returns on a constant, the CRSP value-weighted market returns, and the returns of the three-digit SIC industry portfolio in year  $t$ .  $PIN_t$  is a firm's probability of informed trading in year  $t$  provided by Brown and Hillegeist (2007). Definitions of other variables are listed in the Appendix. The same set of control variables used in Table 3 are included in all regressions. Each regression includes a separate intercept as well as firm and year fixed effects. Standard errors clustered by firm are displayed in parentheses. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent levels, respectively.

<i>Dep. Var.</i>	$1 - R_t^2$		$PIN_t$	
	(1)	(2)	(3)	(4)
<i>Pilot*Post</i>	-0.015 (0.010)	-0.013 (0.009)	-0.020*** (0.003)	-0.019*** (0.003)
Controls	No	Yes	No	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	8,977	8,848	8,942	8,806
R-squared	0.776	0.804	0.616	0.676

**Table A7: Baseline difference-in-differences (DiD) test with additional control variables**

This table reports the results of the difference-in-differences (DiD) test of innovation quality, value, and efficiency around Regulation SHO with additional control variables measured in the same year as the innovation outcome variable. *AcqCost* is acquisition expense scaled by total assets. *DA* is discretionary accruals constructed using the modified Jones model (Jones, 1991). *ForecastBad* is a dummy variable that equals one if a firm has at least one management forecast below the median of outstanding analyst forecast in a given year, and zero otherwise. *InsiderSell* is a dummy variable that equals one if a firm has at least one open market sell made by any insiders as recorded in Form 4 in a given year, and zero otherwise. *SEO* is a dummy variable that equals one if a firm has issued seasoned equity offerings in a given year, and zero otherwise. *LogVega* is the natural logarithm of the convexity of a CEO's compensation payoffs, provided by Coles, Daniel, and Naveen (2006). Definitions of other variables are listed in the Appendix. We include *LogVega* only in columns (4) to (6) because the construction of it requires a firm to be covered by Execucomp, which significantly reduces sample size. Other control variables used in Table 3 Panel A, lagged by one year, are also included but not reported. Each regression includes a separate intercept as well as firm and year fixed effects. Standard errors clustered by firm are displayed in parentheses. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent levels, respectively.

<i>Dep. Var.</i>	<i>NSCitePat<sub>t</sub></i>	<i>Value<sub>t</sub></i>	<i>NSCiteRD<sub>t</sub></i>	<i>NSCitePat<sub>t</sub></i>	<i>Value<sub>t</sub></i>	<i>NSCiteRD<sub>t</sub></i>
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Pilot*Post</i>	0.077** (0.033)	0.052* (0.029)	0.068** (0.031)	0.050* (0.030)	0.067* (0.038)	0.086** (0.038)
<i>AcqCost<sub>t</sub></i>	0.067 (0.114)	-0.051 (0.089)	0.082 (0.091)	0.145 (0.129)	-0.083 (0.115)	0.036 (0.112)
<i>DA<sub>t</sub></i>	-0.006 (0.004)	-0.003 (0.003)	-0.001 (0.003)	-0.003 (0.004)	-0.002 (0.005)	-0.001 (0.004)
<i>ForecastBad<sub>t</sub></i>	-0.011 (0.020)	-0.035** (0.018)	-0.015 (0.017)	-0.011 (0.021)	-0.030 (0.023)	-0.023 (0.019)
<i>InsiderSell<sub>t</sub></i>	0.017 (0.030)	0.014 (0.025)	0.008 (0.025)	-0.010 (0.037)	0.062 (0.047)	-0.007 (0.039)
<i>SEO<sub>t</sub></i>	-0.051 (0.048)	-0.002 (0.051)	-0.026 (0.033)	0.046 (0.063)	0.096 (0.073)	0.032 (0.043)
<i>LogVega<sub>t</sub></i>				-0.017* (0.009)	-0.030*** (0.010)	-0.016* (0.009)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,483	8,483	8,483	5,265	5,265	5,265
R-squared	0.802	0.837	0.797	0.841	0.855	0.830