NBER WORKING PAPER SERIES

THE CAUSAL EFFECT OF LIMITS TO ARBITRAGE ON ASSET PRICING ANOMALIES

Yongqiang Chu
David Hirshleifer
Liang Ma

Working Paper 24144
http://www.nber.org/papers/w24144

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2017, Revised October 2018

For helpful comments and suggestions, we thank Wei Xiong (the Editor), an anonymous Associate Editor, two anonymous referees, Karl Diether, Lukasz Pomorski, Jeffrey Pontiff, Lin Sun, and participants at the 2016 Rodney L. White Center for Financial Research Conference on Financial Decisions and Asset Markets at Wharton and the 2017 American Finance Association Meetings. This research has received financial support from the Moore School Research Grant Program. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Causal Effect of Limits to Arbitrage on Asset Pricing Anomalies
Yongqiang Chu, David Hirshleifer, and Liang Ma
NBER Working Paper No. 24144
December 2017, Revised October 2018
JEL No. G12,G18,G4

ABSTRACT

We examine the causal effect of limits to arbitrage on 11 well-known asset pricing anomalies using the pilot program of Regulation SHO, which relaxed short-sale constraints for a quasi-random set of pilot stocks, as a natural experiment. We find that the anomalies became weaker on portfolios constructed with pilot stocks during the pilot period. The pilot program reduced the combined anomaly long-short portfolio returns by 72 basis points per month, a difference that survives risk adjustment with standard factor models. The effect comes only from the short legs of the anomaly portfolios.

Yongqiang Chu
University of South Carolina
1014 Greene Street
Columbia, SC 29208
yongqiang.chu@moore.sc.edu

David Hirshleifer
The Paul Merage School of Business
University of California, Irvine 4291
Pereira Drive
Irvine, CA 92697
and NBER
david.h@uci.edu

Liang Ma
University of South Carolina
Darla Moore School of Business
1014 Greene Street
Columbia, SC 29208
United States
liang.ma@moore.sc.edu

1 Introduction

Over the last several decades, finance researchers have discovered many cross-sectional asset pricing anomalies, wherein predetermined security characteristics predict future stock returns.[1] Such patterns can derive from either rational risk premia or market mispricing. The mispricing explanation goes hand-in-hand with the idea that there are limits to arbitrage which delay the flow of wealth from irrational to sophisticated investors (Shleifer and Vishny (1997)). In contrast, if return predictability is the result of rational risk premia for bearing factor risk, limits to arbitrage should not affect expected returns.

It is therefore interesting to ascertain whether return anomalies, to the extent that they reflect mispricing, are persistent because limits to arbitrage prevent sophisticated investors from trading profitably against them. However, it is empirically hard to measure pure variations in limits to arbitrage that exclude variations in other economic forces that might affect either risk premia or mispricing. In this paper, we study the *causal* effect of limits to arbitrage on 11 well-known asset pricing anomalies—namely, the momentum, gross profitability, asset growth, investment to assets, return on assets, net operating assets, accruals, net stock issuance, composite equity issuance, failure probability, and O-score anomalies. These 11 anomalies were the focus of Stambaugh, Yu, and Yuan (2012) in their study of sentiment and anomalies, and were chosen because of their survival after adjusting for the Fama-French three factors. Examining the causal effect of limits to arbitrage on these anomalies provides insight into whether, and to what extent, well-known return anomalies derive from risk versus mispricing.

It is challenging to identify the causal effect of limits to arbitrage, as we seldom directly observe them, or pure variations in them. The existing literature therefore often relies on firm characteristics, such as idiosyncratic volatility, size, and stock liquidity, as proxies for limits to arbitrage. However, these proxies are likely to be correlated with risk. For example,

[1] Harvey, Liu, and Zhu (2016) provide a comprehensive list of variables that can predict cross-sectional stock returns.

size has been offered as the basis for a risk factor in the three-factor model of Fama and French (1993), and volatility can be a risk measure in models with limited diversification such as settings with costs of trading or with asymmetric information. This raises the possibility that effects attributed to limits to arbitrage may actually be due to rational risk premia.

We offer here a pure test of the causal effect of limits to arbitrage on asset pricing anomalies. Short sale constraints are one of the most important limits of arbitrage (e.g., Jones and Lamont (2002), Lamont and Thaler (2003), Nagel (2005), Gromb and Vayanos (2010)). Research on the effect of short-sale constraints on asset prices relies mainly on indirect proxies such as breadth of ownership (Chen, Hong, and Stein (2002)), institutional ownership (Asquith, Pathak, and Ritter (2005), Nagel (2005), Hirshleifer, Teoh, and Yu (2011)), firm size (Ali and Trombley (2006), Israel and Moskowitz (2013)), short interest (Asquith, Pathak, and Ritter (2005)), and shorting cost estimated from stock borrowing and lending behavior (Jones and Lamont (2002), Geczy, Musto, and Reed (2002), Drechsler and Drechsler (2014)). Several of these proxies may be correlated across stocks or over time with variations in factor risk.

We exploit a natural experiment, the Rule 202T pilot program of Regulation SHO (hereafter the pilot program), to identify the *causal* effect of limits to arbitrage, and in particular short-sale constraints, on asset pricing anomalies. Regulation SHO was adopted by the Securities and Exchange Commission (SEC) in July 2004. Within stocks in the Russell 3000 index as of June 2004, the pilot program designated every third stock ranked by average daily trading volume (in the prior year) on each of NYSE, AMEX, and Nasdaq as pilot stocks. The pilot program then removed short sale price tests on this quasi-randomly selected group of pilot stocks. Prior to Regulation SHO, the specific form of short sale price tests differed across different stock markets. NYSE/AMEX imposed the uptick rule, which only allowed a short sale to be placed on a plus tick or a zero-plus tick. Nasdaq imposed the bid price test, which did not allow short sales at or below the (inside) bid when the inside bid was at or below the previous inside bid. From May 2, 2005 to August 6, 2007, the pilot stocks on NYSE/AMEX were exempted from the uptick rule and those on Nasdaq were exempted from the bid price test. The pilot program therefore made it easier to short sell pilot stocks relative to non-pilot stocks. Because the assignment of pilot and non-pilot firms is quasi-random, the program provides an ideal setting to examine the causal effect of short-sale constraints on asset pricing anomalies. It is known that (see Diether, Lee, and Werner (2009) and discussion in Section 2) the bid price test for Nasdaq stocks is not very restrictive, and a significant fraction of trading volume in Nasdaq-listed stocks is executed on ArcaEx and INET, which do not enforce the bid price test. We therefore exclude Nasdaq stocks and only include pilot and non-pilot stocks traded on NYSE/AMEX in our main analysis.

We examine two main hypotheses regarding the differential performance of pilot versus non-pilot anomaly portfolios, *during* the pilot period of Regulation SHO. The first is that the anomalies become weaker for pilot firms relative to non-pilot firms during the pilot period. During the pilot period, arbitrageurs could more easily short pilot stocks to construct arbitrage portfolios, which should reduce mispricing. It follows that the return spread of arbitrage portfolios should decline for pilot stocks relative to non-pilot stocks.

To test the first hypothesis, for each asset pricing anomaly, we construct long-short portfolios with pilot and non-pilot stocks separately. Specifically, we first sort all pilot stocks into deciles according to the return-predicting characteristic, and then calculate the anomaly returns as the return differences between the highest performing decile based on existing anomaly evidence (the long leg) and the lowest performing decile (the short leg). We then do the same with all non-pilot stocks. In a difference-in-differences framework, we find that the anomalies were much weaker in long-short portfolios constructed using pilot stocks during the pilot period. The effect is statistically significant in five of the 11 anomalies. When the 11 anomalies are combined in a joint test, the effect is both statistically and economically significant. The pilot program reduced the anomaly returns by 72 basis points per month, or 8.64% per year.

The second hypothesis is that the decrease in anomaly returns for pilot stocks during the pilot period comes mostly from the short leg portfolios. In general, anomaly returns can come from either overpriced short legs or underpriced long legs. A loosening of short-sale constraints should reduce profitability of short leg arbitrage portfolios. In the same difference-in-differences framework, we find that the returns of short leg portfolios constructed with pilot stocks were significantly and substantially higher during the pilot period, i.e., short strategies became less profitable. In contrast, there is no significant effect of the pilot program on long leg portfolios.

We consider two additional hypotheses. First, the difference in anomaly returns between pilot and non-pilot stocks should vanish after the ending of the pilot program, with the disappearance of the difference in short-sale restrictions between pilot and non-pilot stocks.[2] We find empirical evidence consistent with this hypothesis. Furthermore, we expect to observe return dynamics of short-leg portfolios at the beginning and the end of the pilot program. Specifically, pilot short legs should underperform non-pilot short legs at the beginning of the pilot program, right after the uptick rule was lifted for pilot stocks. Similarly, pilot short legs should outperform non-pilot short legs at the end of the pilot program, right after the uptick rule was lifted for non-pilot stocks. We find evidence supporting this hypothesis as well.

We provide a battery of robustness checks for our main results derived from testing the two main hypotheses. We first show that our main results are robust to different sample periods. Furthermore, we carry out a set of falsification tests. As a placebo test, we maintain the assignment of pilot and non-pilot firms but change the timing of the pilot period fictitiously to 2001-2003 and test whether this fictitious pilot program also affected the asset pricing anomalies during the 1980-2003 period.[3] We find that the fictitious pilot program had no effect on asset pricing anomalies, suggesting that the main results are indeed driven

[2] After the pilot program ended, the uptick rule was lifted for non-pilot stocks as well.

[3] We end the placebo test sample in year 2003 so that the actual pilot program does not affect the placebo test results.

by the pilot program. As another falsification test, we show that the difference in anomaly strength between pilot and non-pilot stocks during the pilot period was small and insignificant for Nasdaq stocks, which again confirms that our main results come from the relaxation of short-sale constraints.

Lastly, we perform several subsample analyses for our main results. As argued and documented by Diether, Lee, and Werner (2009), small and less liquid stocks were more affected by the suspension of the uptick rule (see the discussion in Section 5.4). Consistent with this, we find that the main effect (i.e. the effect of easier short selling on anomalies during the pilot period) is more pronounced among small and less liquid stocks. We further perform a more direct subsample analysis that splits stocks based on the extent to which they were restricted by the uptick rule before the pilot program. We find that our main effect is stronger among stocks that were more restricted by the uptick rule before the pilot program, which again corroborates our mechanism.

Collectively, these results show that limits to arbitrage, and in particular, short-sale constraints play an important role in generating the 11 well-known anomalies. These findings therefore suggest that these anomalies are, at a minimum, driven in substantial part by mispricing.

A potential alternative explanation for our main results is that the pilot program made pilot stocks more salient. Even though the pilot program is a quasi-experiment, it is not a double blind study; market participants were aware of the change. It is possible that the sheer fact that the list of pilot stocks was publicly known drew attention to these stocks, and that higher investor attention to pilot stocks weakened anomalies, driving our main results. The results on Nasdaq stocks indicate that this mechanism is less plausible than our proposed mechanism of change in limits to arbitrage, since under this story Nasdaq pilot stocks that experienced an increase in investor attention should also have weakened anomalies, which is not the case. In contrast our limits to short arbitrage hypothesis explains why the pilot effects are not present on Nasdaq. Furthermore, the results on shorting activity and subsample results based upon uptick rule restrictiveness probe further into whether short-sale constraints are the source of the effects of the pilot program, and lend further support to our proposed mechanism. However, we do not assert that our tests rule out the alternative mechanism completely.

The behavioral finance literature has long argued that limits to arbitrage help explain the persistence of asset pricing anomalies despite the incentives of sophisticated investors to trade profitably against such anomalies (Shleifer and Vishny (1997), Hirshleifer (2001), Barberis and Thaler (2003), Gromb and Vayanos (2010)). Empirical tests have examined the association between various proxies for limits to arbitrage and asset returns. These proxies for limits to arbitrage include stock price (Pontiff (1996), Mashruwala, Rajgopal, and Shevlin (2006)), size (Pontiff (1996), Ali, Hwang, and Trombley (2003), Israel and Moskowitz (2013)), idiosyncratic volatility (Ali, Hwang, and Trombley (2003), Mashruwala, Rajgopal, and Shevlin (2006)), transaction costs (Ali, Hwang, and Trombley (2003)), investor sophistication (Ali, Hwang, and Trombley (2003)), dollar trading volume (Mashruwala, Rajgopal, and Shevlin (2006)), and capital constraints of merger arbitrageurs (Baker and Savaşoglu (2002)) in the context of merger arbitrage.

Many of the proxies for limits to arbitrage used in existing literature may actually capture risk, which makes it hard to distinguish between risk-based and mispricing-based explanations of anomalies. As documented by Lam and Wei (2011), proxies for limits to arbitrage are often highly correlated with proxies for investment frictions (risk).[4]

Our paper is more closely related to the empirical literature on how short sale constraints or short sale costs affect asset prices and asset pricing anomalies.[5] One strand of this litera-

[4]Lam and Wei (2011) attempt to distinguish between mispricing-based and risk-based (q-theory with investment frictions) explanations of the asset growth anomaly. They examine a comprehensive list of proxies for limits to arbitrage: idiosyncratic volatility, the number of institutional shareholders, three measures of information uncertainty including analyst coverage, dispersion in analysts' earnings, and cash flow volatility, and five measures of transaction costs including stock price, effective bid-ask spread, institutional ownership, Amihud illiquidity, and dollar trading volume.

[5]See Reed (2013) and the references therein for more discussion on the role of short selling in financial markets.

ture employs indirect proxies for short-sale constraints (Chen, Hong, and Stein (2002), Nagel (2005), Ali and Trombley (2006), Asquith, Pathak, and Ritter (2005), Hirshleifer, Teoh, and Yu (2011), Israel and Moskowitz (2013)). The indirect proxies of short-sale constraints used in this strand of literature may, however, also capture variations in risk.

Another strand of this literature uses more direct proxies for short sale constraints or short sale costs, measured using data of stock borrowing and lending (D'Avolio (2002), Geczy, Musto, and Reed (2002), Jones and Lamont (2002), Cohen, Diether, and Malloy (2007), Cao, Dhaliwal, Kolasinski, and Reed (2007), Saffi and Sigurdsson (2010), Drechsler and Drechsler (2014), Beneish, Lee, and Nichols (2015), Engelberg, Reed, and Ringgenberg (2017)). These proxies are more direct in the sense that they are associated with the equity lending process, e.g. with stock loan fees and stock lending supply. Nevertheless, these proxies can still be correlated across stocks or over time with shifts in factor risk, so that the return effects can still be driven by risk premia. In contrast, the natural experiment in our paper focuses on a regulatory shift that only alters permitted short-selling behavior, and therefore is unlikely to be correlated with shifts in factor risk.

Our paper contributes to the existing literature by providing a clean and powerful test of the causal effect of limits to arbitrage in general and short-sale constraints in particular on asset pricing anomalies. In contrast to existing literature that mainly relies on proxies for limits to arbitrage and short-sale constraints (which may capture risk or be correlated with risk as discussed above), we use exogenous shocks to short-sale constraints generated by a natural experiment, the pilot program of Regulation SHO. The quasi-randomness of the assignment of pilot and non-pilot stocks makes a stock's assignment unlikely to be correlated with the loadings of stocks on risk factors. We are therefore able to conclude from our analysis whether limits of arbitrage and thereby mispricing actually affect asset pricing anomalies.

Our paper also adds to the literature that studies the impact of the pilot program of Regulation SHO. A few recent papers examine its effect on aspects related to stock prices. Diether, Lee, and Werner (2009) examine whether the suspension of short-sale price tests by the pilot program affects market quality. They find that short-selling activity increases for NYSE and Nasdaq pilot stocks and NYSE pilot stocks experience a slight increase in spreads and intraday volatility, while the effect on market quality for Nasdaq stocks is smaller. Grullon, Michenaud, and Weston (2015) find that the pilot program leads to an increase in short-selling activity and a decline in prices for pilot stocks and this effect is stronger for small firms, and these firms react by reducing equity issues and investment. Li and Zhang (2015) show that the pilot program increases price sensitivity to bad news and thereby makes managers more likely to reduce the precision of bad news forecasts. Fang, Huang, and Karpoff (2016) show that the pressure of short-selling on stock prices due to the pilot program curbs managers' willingness to manipulate earnings. To the best of our knowledge, our paper is the first to show that the pilot program affects the strength of well-known anomalies and reduces overpricing on short legs of these anomalies. Moreover, the implications of our results go beyond the effect of the pilot program *per se* and provide insight to a broad question in the asset pricing literature, whether these anomalies reflect mispricing or compensation for risk.

2 Background on Regulation SHO

Following the stock market crash of 1929, short-selling restrictions (price tests) were introduced in the 1930s on stocks traded in the United States. Prior to Regulation SHO, the specific form of short selling price tests differed across different stock markets. NYSE/AMEX imposed the uptick rule, which only allowed a short sale to be placed on a plus tick or a zero-plus tick. The zero-plus tick is a zero tick where the most recent price change preceding the trade was a plus tick. Nasdaq imposed the bid price test. Short sales on Nasdaq stocks were not allowed at or below the (inside) bid when the inside bid was at or below the previous inside bid.

Regulation SHO was designed by the SEC to investigate whether the uptick rule imposed by NYSE (and also AMEX) and the bid price test imposed by Nasdaq affect market quality, and to develop potential uniform price tests if these rules are necessary. It was announced by the SEC on July 28, 2004. For stocks in the Russell 3000 index as of June 2004, the pilot program of Regulation SHO designated every third stock ranked by average daily trading volume (in the prior year) on each of NYSE, AMEX, and Nasdaq as pilot stocks. The Russell 3000 stocks were chosen by the SEC as they represent a broad cross-section of U.S. stocks. The volume-ranking design was adopted to provide a quasi-random assignment of pilot versus non-pilot stocks. The pilot program removed the uptick rule and the bid price test on this quasi-randomly selected group of pilot stocks. From May 2, 2005 to August 6, 2007, the pilot stocks on NYSE/AMEX were exempted from the uptick rule and those on Nasdaq were exempted from the bid price test.

The pilot program ended on August 6, 2007. Slightly before the ending of the pilot program, on July 6, 2007, the SEC eliminated short-sale price tests for all exchange-listed stocks. Therefore, the pilot program effectively ran from May 2, 2005 to July 6, 2007. As discussed in Fang, Huang, and Karpoff (2016), the elimination of the short-sale price tests for all exchange-listed stocks received extensive criticism from managers and politicians. This led the SEC to partially restore a modified uptick rule on February 24, 2010. Under the modified rule, short-sale price tests were imposed when a stock's price declines by 10% or more from its closing price on the previous trading day.

As discussed in details in Diether, Lee, and Werner (2009), the bid price test is not very restrictive (compared to the uptick rule).[6] Furthermore, a significant fraction of trading volume in Nasdaq-listed stocks is executed on ArcaEx and INET, which do not enforce the bid price test. As a result, the effect of the pilot program on Nasdaq-listed stocks in terms of relaxing short-sale arbitrage constraints should be minimal. We therefore exclude Nasdaq

[6] Here, we quote the example given on page 44 of Diether, Lee, and Werner (2009) to illustrate why the bid price test is less restrictive than the uptick rule. Suppose for a stock the last sale price is $28.05 on a plus tick, and the quotes are $28.00 to $28.05. To comply with the bid price test, a short sale can be placed on Nasdaq at a marketable limit sell order at $28.00, as long as the most recent bid was $28.00 or below. A NYSE short seller, however, has to place the short-sale order at $28.05, which is 4 cents higher, in order to comply with the uptick rule.

stocks and only include pilot and non-pilot stocks traded on NYSE/AMEX in our main analysis.

3 Data and Anomalies

3.1 Sample

Starting with the June 2004 Russell 3000 index, we follow the procedure described in SEC's first pilot order of Regulation SHO (Securities Exchange Act Release No. 50104) to build our sample of pilot and non-pilot stocks. We exclude stocks that were not listed on the NYSE, AMEX, or Nasdaq NM, and stocks that went public or had spin-offs after April 30, 2004. The initial sample consists of 986 pilot stocks (based on the list published in the SEC's pilot order[7]) and 1,966 non-pilot stocks. We then merge this initial sample with the Center for Research in Security Prices (CRSP) and Compustat (both annual and quarterly data) data sets to form portfolios and analyze portfolio returns of the 11 anomalies. As discussed in the introduction and Section 2, the bid price test for Nasdaq-listed stocks is likely to have minimal effect. Our final sample therefore consists of pilot and non-pilot stocks in the pilot program that are listed on NYSE or AMEX at portfolio formation. Within the initial sample of pilot and non-pilot stocks of the pilot program, 1,025 non-pilot stocks and 515 pilot stocks are included in our final sample, among which 1,477 stocks are traded on NYSE and 63 stocks are traded on AMEX. The ratio of non-pilot stocks to pilot stocks is roughly 2:1.[8] The sample period for our main empirical analysis is from January 1980 to June 2007, after which the pilot program of Regulation SHO ended.[9]

[7]https://www.sec.gov/rules/other/34-50104.htm

[8]In untabulated analysis, we examine the robustness of our results when we set the number of pilot and non-pilot firms to be equal, by randomly removing half of the non-pilot firms with simulation. We show that our main results are robust in this aspect.

[9]The pilot program officially ended on August 6, 2007. However, on July 6, 2007, all exchange-listed stocks were exempted from short-sale price tests, which effectively ended the pilot program.

Our sample of pilot and non-pilot stocks is selected at the end of June 2004. For most of the pre-pilot period in our difference-in-differences analysis, the sample is selected (in terms of selecting pilot versus non-pilot stocks) using information not available yet. This, however, is not an issue for our analysis. The reason is in the pre-pilot period, the information of a stock being pilot or not is only used to classify it into different groups of comparison. Furthermore, Table 4 shows that our main results can also be identified using the pilot period *per se*, when the information of a stock being pilot or not is available.

3.2 Anomalies

We focus on the 11 anomalies studied by Stambaugh, Yu, and Yuan (2012), which they select based on survival after adjustment for the Fama-French three factors.

Below we briefly describe each anomaly, relegating details of variable construction to the Appendix. For each anomaly, there is a corresponding long-short trading strategy that goes long in the stocks that earn high returns (the long leg) and goes short in those that earn low returns (the short leg). The relationship between the subsequent stock performance and the ranking variable is positive for some anomalies and negative for others. For example, stocks with high past returns outperform those with low past returns for the momentum anomaly, while stocks with low asset growth rate outperform those with high asset growth rate for the anomaly of asset growth. Table 1 summarizes the characteristics of stocks in the long and short legs for each anomaly.

Anomaly 1: Momentum. The momentum effect in stock returns was first documented by Jegadeesh and Titman (1993), and is one of the most prominent anomalies in asset pricing. It refers to the phenomenon that stocks with higher past recent returns continue to outperform stocks with lower past recent returns. We employ the conventional 11-1-1 momentum strategy to construct our momentum portfolios. The ranking period is 11-month from $t-12$

to $t-2$. The holding period is month t. Month $t-1$ is skipped to eliminate the short-run reversal effect.

Anomaly 2: Gross profitability. As documented by Novy-Marx (2013), stocks with high gross profitability on average earn higher returns than stocks with low gross profitability. He further shows that the profitability premium becomes more pronounced after controlling for the value premium. Following Novy-Marx (2013), we measure gross profitability as total revenue minus cost of goods sold, scaled by total assets.

Anomaly 3: Asset growth. Cooper, Gulen, and Schill (2008) find that stocks with a high growth rate in their total assets earn low subsequent returns. A possible explanation for this phenomenon is that investors tend to overreact to growth rates in total assets. We measure asset growth as the change in total assets, scaled by lagged total assets.

Anomaly 4: Investment to assets. Titman, Wei, and Xie (2004) find that firms increasing capital investments earn negative benchmark-adjusted returns subsequently. They propose that this phenomenon is consistent with the hypothesis that investors underreact to the empire building implications of increased investment expenditures. We measure investment to assets as the annual change in gross property, plant, and equipment plus the annual change in inventories, scaled by lagged total assets.

Anomaly 5: Return on assets. Fama and French (2006) document that in Fama-MacBeth cross-sectional regressions, earnings can predict stock returns. Chen, Novy-Marx, and Zhang (2011) and Stambaugh, Yu, and Yuan (2012) find that return on assets, measured on a quarterly basis, can predict subsequent stock returns. A higher past return on assets leads to higher subsequent stock returns. We measure return on assets as quarterly earnings scaled by quarterly total assets.

Anomaly 6: Net operating assets. Hirshleifer, Hou, Teoh, and Zhang (2004) find that firms with higher net operating assets earn lower subsequent returns. They attribute this phenomenon to investor limited attention. Net operating assets capture cumulative differences between operating income and free cash flow. Investors with limited attention may not process all information thoroughly and therefore may focus on accounting profitability without sufficiently taking into account cash profitability information, leading to overvaluation of firms with higher net operating assets. We measure net operating assets as the difference between all operating assets and all operating liabilities on the balance sheet, scaled by lagged total assets.

Anomaly 7: Accruals. As documented by Sloan (1996), firms with higher accruals on average earn lower subsequent returns. This suggests that stock prices fail to fully reflect information contained in the accruals and cash flow components of current earnings, which is consistent with investors having limited attention. We measure operating accruals as changes in non-cash working capital minus depreciation expense, scaled by lagged total assets.

Anomaly 8: Net stock issues. As documented by Loughran and Ritter (1995) and Pontiff and Woodgate (2008), net share issues negatively predict stock returns in the cross section. One explanation for this phenomenon in the literature is that firms issue stocks when they are overvalued and retire stocks when they are undervalued. We measure net stock issues on the annual basis as the change in the natural logarithm of a firm's adjusted shares over the last year.

Anomaly 9: Composite equity issues. Daniel and Titman (2006) find that an alternative measure of equity issuance, the composite equity issuance, is also a negative predictor of stock returns in the cross section. They propose that this measure is related to the "intangible" component of past returns. Measured as the part of growth rate in market equity not attributable to stock returns, composite equity issuance captures the amount of equity a firm issues (or retires) in exchange for cash or services. As a result, this measure increases with seasoned equity issuance, employee stock option plans, and share-based acquisitions, and decreases with share repurchases, dividends, and other actions that take cash out of the firm.

Anomaly 10: Failure probability. We use the failure probability proposed by Campbell, Hilscher, and Szilagyi (2008) to measure financial distress, which is estimated from a dynamic logit model to match empirically observed default events, with both market and accounting information taken into account. Campbell, Hilscher, and Szilagyi (2008) show that with this measure, more distressed firms earn lower subsequent returns on average than less distressed firms, especially after 1981.

Anomaly 11: O-score. We also use an alternative measure of financial distress, the O-score proposed by Ohlson (1980). Dichev (1998) shows that with this measure, more distressed firms earn lower subsequent returns on average than less distressed firms.

3.3 Summary of Anomaly Returns in Our Sample

Before proceeding to the main empirical analysis, we first verify the existence of the 11 anomalies in our sample of pilot and non-pilot firms. For each anomaly, we sort stocks in our sample into deciles based on the corresponding ranking variables and calculate the gross-return-weighted anomaly returns as the return differences between the highest performing decile (the long leg) and the lowest performing decile (the short leg). In other words, the portfolio break points we use are the decile break points in our pilot and non-pilot samples (that contain only NYSE/AMEX stocks), respectively.

Equal-weighted portfolio returns can lead to various statistical and microstructure biases (Asparouhova, Bessembinder, and Kalcheva (2013)). On the other hand, it is useful for testing purposes to make use of the information in small firm returns, because small firms are especially informative in understanding the effects of limits to arbitrage. As discussed in Diether, Lee, and Werner (2009), the suspension of short-sale price tests is likely to affect smaller stocks more. So a test that makes use of small firm returns maximizes our power to test the relevant hypothesis, whether limits to short arbitrage (in the form of short-sale price tests) affect anomalies. Our main tests therefore use gross-return-weighted portfolio returns unless otherwise noted. The gross-return weight for stock i at each month t, is its gross return $R_{i,t-1}$ in the preceding month $t-1$. As discussed in Asparouhova, Bessembinder, and Kalcheva (2010, 2013), gross-return-weighting places similar weight in drawing inferences on the information provided by each stock in the sample while mitigating the statistical and microstructure biases associated with equal-weighted portfolio returns.

To evaluate how different the effects of the uptick rule are on small versus large stocks, we also report the main difference-in-differences results using value-weighted portfolio returns in Table A.1 of the Appendix, and find a small effect. Comparing with the results using gross-return-weighted portfolio returns in Table 4 indicates that the uptick rule is indeed more important for small stocks.

We use data from CRSP to construct portfolios of Anomalies 1 and 9, use Compustat annual data to construct portfolios for Anomalies 2, 3, 4, 6, 7, and 8, use Compustat quarterly data to construct portfolios for Anomaly 5, and use CRSP and Compustat quarterly data to construct portfolios for Anomalies 10 and 11. For anomalies that use annual Compustat data, we follow Fama and French (1992) to match the accounting data for all fiscal years ending in calendar year $t-1$ with the stock returns from July of year t to June of $t+1$. For anomalies that use quarterly Compustat data, we use accounting information lagged by one quarter to match with stock returns.

We examine the average of raw anomaly returns and benchmark-adjusted anomaly returns controlling for the Capital Asset Pricing Model (CAPM) and the Fama-French three factor model over the sample period of January 1980 to December 2004. We end the sample period in December 2004 to avoid overlap with the pilot program. The average of benchmark-adjusted returns is the alpha from regressing the time series of excess returns onto the time series of the appropriate factors (the market excess return for the CAPM, and two additional factors, the SMB and HML factors, for the Fama-French three-factor model). Table 2 reports these average returns.

Table 2 shows that the long-short portfolio returns for all 11 anomalies survive risk-adjustment with the Fama-French three-factor model. The average Fama-French-three-factor-adjusted anomaly returns are presented in the last column of Table 2 and they are positive and statistically significant for all 11 anomalies. These results are consistent with the evidence in Stambaugh, Yu, and Yuan (2012). We therefore confirm that these anomalies exist on our more restricted sample of stocks.

4 Empirical Analysis

As stated in the introduction, our two main hypotheses are:

Hypothesis 1. *The relaxation of short sale constraints caused by the pilot program of Regulation SHO reduces anomaly returns for pilot stocks relative to non-pilot stocks during the pilot period.*

Hypothesis 2. *This decrease in anomaly returns comes primarily from the short leg anomaly portfolios. Short legs of pilot stocks outperform those of non-pilot stocks during the pilot period.*

We also test two additional hypotheses:

Hypothesis 3. *The difference in anomaly returns between pilot and non-pilot stocks disappears after the ending of the pilot program.*

Hypothesis 4. *At the beginning of the pilot program, short legs of pilot stocks underperform those of non-pilot stocks. At the end of pilot program, short legs of pilot stocks outperform those of non-pilot stocks.*

4.1 Verifying the Quasi-Randomness of Pilot Stock Assignment

As discussed in Section 2, the pilot firms were assigned in a quasi-random experiment (every third firm in a sorting of firms by trading volume on NYSE and, separately, on AMEX).

In our context, we further confirm that the pilot firms were in fact quasi-randomly assigned, with respect to firm characteristics associated with the 11 anomalies. To do so, we compare these firm characteristics between pilot and non-pilot firms at the end of year 2003, before the announcement of the pilot program (July 2004). We calculate the mean of these anomaly variables for pilot and non-pilot firms at the end of year 2003, and calculate their differences and the robust t-statistics of the differences. All variables are winsorized at the 1st and 99th percentiles of all firm-month observations to limit the effect of outliers. The results are reported in Panel A of Table 3. Except for the measure of gross profitability, for which the difference is only significant at the 10% level, other anomaly predictors show no statistically significant differences between pilot and non-pilot firms. Furthermore, the difference in gross profitability between pilot and non-pilot firms is small in magnitude compared with the two sample means.[10] As our empirical analysis focuses on the long legs and short legs of anomalies, in Panels B and C of Table 3, we also compare the mean of these anomaly variables for pilot and non-pilot stocks that fall in these two legs. The differences are again mostly small and statistically insignificant. Collectively, these results suggest that there was no significant difference between pilot and non-pilot firms prior to the announcement of the pilot program.

4.2 Main Difference-in-Differences Results

We now test the two main hypotheses, Hypotheses 1 and 2. We explore whether the pilot program led to differences in anomaly returns for the pilot stock sample relative to

[10] In untabulated results, we confirm that the difference in size and book-to-market ratio between pilot and non-pilot firms is also small and statistically insignificant.

the non-pilot stock sample, during the pilot period. We first construct anomaly portfolios based on pilot and non-pilot firms separately. Specifically, we sort all pilot stocks into deciles according to the predictors of the anomalies, and then calculate the returns of the highest performing decile (the long leg returns), the returns of the lowest performing decile (the short leg returns), and the differences between the two (the long-short returns). We then do the same on all non-pilot firms. We then examine whether returns of pilot portfolios are different from returns of non-pilot portfolios during the pilot period (relative to the pre-pilot period), using a difference-in-differences approach. Throughout the paper, we use the terms difference-in-differences and DiD interchangeably.

The main difference-in-differences test employs the following specification:

$$r_{it} = \gamma_t + \beta Pilot_i \times During_t + \beta_1 Pilot_i + \epsilon_{it}, \tag{1}$$

where r_{it} is the gross-return-weighted monthly return of portfolio i, which can be the long leg, the short leg, or the long-short portfolio of an anomaly, in month t; γ_t is the time fixed effects; $Pilot_i$ is a dummy variable which is equal to one if portfolio i is formed on pilot firms, and zero otherwise; $During_t$ is a dummy variable, which equals one if month t is between July 2005 and June 2007, i.e. during the pilot period of Regulation SHO. Since $During_t$ is subsumed by the time fixed effects, it is dropped from the regression. The time fixed effects γ_t capture the common factors and/or common macroeconomic variables that drive the portfolio returns for both pilot and non-pilot portfolios. In all analysis involving equation (1), we drop the two months (May and June 2005) at the beginning of the pilot program, to avoid capturing price movement over this short time window. In these regressions, the unit of analysis is a portfolio month observation. We estimate equation (1) for the long leg, the short leg, and the long-short portfolios separately.[11] The difference-in-differences coefficient

[11] We also explore an alternative empirical design to capture the effect of the pilot program on anomalies. Specifically, for each portfolio (the long/short leg or the long-short portfolio) of an anomaly, we take the return difference between the pilot portfolio and the non-pilot portfolio and denote the time series of this difference as r^d_{it}. By doing so, we isolate the cross-sectional difference between pilot and non-pilot portfolios.

β in equation (1) is the main coefficient of interest. It captures the difference in anomaly portfolio returns between pilot stocks and non-pilot stocks during the pilot period, relative to that in the pre-pilot period. We run the regression in equation (1) for each individual anomaly and also for all 11 anomalies combined. In the aggregate analysis, we replace the time fixed effects by the anomaly-time fixed effects, i.e. the fixed effects associated with each pair of anomaly and time. The aggregate analysis enhances the power of our test and produces the average effect of the pilot program across all 11 anomalies. The results are reported in Table 4.

Hypothesis 1 predicts that β is negative for anomaly long-short returns. Hypothesis 2 predicts that β is positive for short-leg returns, and will be close to zero for long-leg returns. The results support these two hypotheses. For the long-short returns (the last two columns of Panel A of Table 4), β's are consistently negative for all 11 anomalies and are statistically significant for five of them. When the 11 anomalies are combined, i.e. in the aggregate analysis where equation (1) is estimated for all 11 anomalies together with the time fixed effects replaced by the anomaly-time fixed effects, β is -0.72% with a t-statistic of -4.37. In other words, the pilot program reduced the monthly anomaly returns by 72 basis points per month, or 8.64 percentage points per year, on average. These results are consistent with Hypothesis 1.

In addition to the difference-in-differences coefficient β, we report the mean anomaly returns for non-pilot and pilot stocks, in the pre-pilot and during-pilot periods. They are presented in the first eight columns of Table 4, together with the differences in anomaly returns between pilot and non-pilot stocks in the pre-pilot and during-pilot periods. The results in Panel A of Table 4 show that the difference between pilot and non-pilot anomaly

We expect that this difference will only predict returns when $During_t = 1$. Therefore, we run the regression $r^d_{it} = \alpha + \beta During_t + \epsilon_{it}$, where the coefficient β captures the difference in portfolio returns between pilot and non-pilot stocks during the pilot period (relative to that in the pre-pilot period). It can be shown that this specification and the specification in equation (1) are mathematically equivalent, and we confirm in untabulated results that they produce identical estimates of β's. In the rest of the paper, we focus our discussion on results from the specification in equation (1).

returns does not exist in the pre-pilot period: it is small and insignificant for all 11 anomalies. Instead, the effect of the pilot program on anomaly returns comes mainly from the during-pilot period. The average difference in anomaly returns between pilot and non-pilot stocks in the during-pilot period is -0.65% with a t-statistic of -4.36. This evidence again suggests that our results are driven by the pilot program.

The results in Panels B and C indicate that the decreases in anomaly returns come almost entirely from the short legs. For short leg portfolios (Panel C), β's are consistently positive for all 11 anomalies and are statistically significant for five of them. When the 11 anomalies are combined, β is 0.63% with a t-statistic of 5.31. Again, the first eight columns of Panel C show that the effect comes mainly from the during-pilot period. In contrast, for long leg portfolios (Panel B), β's are close to zero and statistically insignificant for most anomalies. When the 11 anomalies are combined, β is still close to zero and statistically insignificant. These results support Hypothesis 2.

In the Appendix, we show that using benchmark-adjusted (CAPM- and FF-three-factor-model-adjusted) returns as the dependent variable in equation (1) delivers almost identical results as those in Table 4. This is expected, because the loadings on the benchmark factors of pilot versus non-pilot firms should also be similar, as the selection of pilot firms is quasi-random.

Our main tests for Hypotheses 1 and 2 based on equation (1) use a relatively short (two years) pilot period to estimate the effect of relaxing short-sale constraints on anomaly returns. This immediately raises the question of whether the sample size generates enough power to distinguish hypotheses. In the results presented in Table 4, we do indeed obtain statistically significant effects on some individual anomalies, especially on the short legs, as well as strong significance for the results that aggregate across the 11 anomalies. This is reassuring, and we discuss why our testing approach has sufficient power to distinguish hypotheses in the Appendix.

4.3 Post-Pilot-Program Results

After the ending of the pilot program, the difference in short-sale restrictions between pilot and non-pilot stocks disappeared. If our main results in Table 4 are indeed driven by the pilot program, we should expect that the difference in anomaly returns between pilot and non-pilot firms also vanish after the ending of the pilot program. This is formally stated in Hypothesis 3. To test this, we estimate a revised difference-in-differences specification as follows:

$$r_{it} = \gamma_t + \beta Pilot_i \times During_t + \beta_1 Pilot_i + \beta_2 Pilot_i \times Post_t + \epsilon_{it}, \tag{2}$$

where $Post_t$ is a dummy variable that represents the post-pilot period and equals 1 if month t is after August 2007 and zero otherwise, and other notations are defined exactly the same as in equation (1). The difference-in-differences coefficient β_2 in equation (2) is the coefficient of interest in this subsection and we expect it to be close to zero.

We run the regression in equation (2) for the whole sample period of January 1980 to December 2016. For this analysis, in addition to dropping the two months at the beginning of the pilot program (May and June 2005) as in the main analysis, we also drop the two months at the end of the pilot program (July and August 2007) to avoid capturing price movement over this short time window. The coefficients β (representing the effect of the pilot program during the pilot period) are identical to those reported in Table 4 (which can be shown mathematically) and therefore are not shown. Table 5 reports the coefficients β_2 for the 11 anomalies individually and in aggregate. All coefficients β_2 are statistically insignificant and close to zero when the 11 anomalies are combined. This confirms that as the difference in short-sale restrictions between pilot and non-pilot firms disappeared, the difference in anomaly returns between them also vanished, which is consistent with our main conclusion.

We now test between two possible scenarios for what may have happened after the ending of the pilot program. First, as we hypothesize, the non-pilot stocks became like the pilot stocks because they also experienced relaxation of short-sale constraints, which is also our main proposed mechanism. Second, the pilot stocks reverted back and became once again like non-pilot stocks, if they were subject to some unknown temporary influence other than the relaxation of short-sale constraints. We can have an insignificant β_2 in both cases. We now perform further tests to distinguish between these two possible scenarios.

To this end, we examine and compare the strength of anomaly returns in the during-pilot and post-pilot periods, for pilot versus non-pilot stocks. The average anomaly raw return across 11 anomalies in the during-pilot period is 0.51% ($t = 2.82$) for non-pilot stocks (see also Table A.3) and -0.14% ($t = -0.85$) for pilot stocks.

In Scenario 1 (our paper's main proposed mechanism), we should see anomaly returns also become *weaker* for non-pilot stocks right after the ending of the pilot program. In Scenario 2, we should see anomaly returns become *stronger* for pilot stocks right after the ending of the pilot program.

We find that in September 2007 to August 2009, which covers a time period of equal length to the pilot period right after the pilot program, the average anomaly raw return across 11 anomalies during the pilot period is -0.10% ($t = -0.12$) for non-pilot stocks and 0.04% ($t = 0.04$) for pilot stocks. This evidence is consistent with Scenario 1, not Scenario 2. This further supports our paper's main proposed mechanism and shows that easier arbitrage also reduced anomaly returns for non-pilot stocks right after the ending of the pilot program.

4.4 Return Dynamics of Short Legs

Our main results show that short-leg portfolios of pilot stocks outperformed those of non-pilot stocks *during* the pilot period. As stated in Hypothesis 4, there are additional implications of our mechanism regarding the price movement of short legs at the beginning and the end of the pilot program. Specifically, if the relaxation of short-sale constraints is indeed the underlying driving force of our main results, we should see a large price decrease (negative returns) at the beginning of the pilot program for pilot short legs (relative to non-pilot short legs), right after the uptick rule was lifted for pilot stocks. Similarly, at the end of the pilot program, when the uptick rule was lifted for non-pilot stocks as well, the relaxation of short-sale constraints would reduce overpricing for non-pilot short legs relative to pilot short legs. We therefore expect to see a price increase in the short legs of pilot stocks relative to those of non-pilot stocks.

We next test Hypothesis 4. We also provide a more complete view of return dynamics of pilot short-leg portfolios (relative to non-pilot short-leg portfolios), over time windows that cover the pre-pilot, during-pilot, and post-pilot periods.

Specifically, we construct a set of consecutive rolling time windows that begins in May 1999 to cover the three periods. Most of the time windows are two years long, in line with the length of the during-pilot period. At the beginning and the end of the pilot program, we use a short window of two months, to capture potential large price movements around these two dates. The exact timing of these windows is: May 1999 to April 2001, May 2001 to April 2003, May 2003 to April 2005, May 2005 to June 2005, July 2005 to June 2007, July 2007 to August 2007, September 2007 to August 2009, September 2009 to August 2011, and September 2011 to December 2013. The first three windows cover the pre-pilot period. The fourth (two-month) window covers the beginning of the pilot program. The fifth window covers the during-pilot period and corresponds to the dummy variable *During* in equation (1). The sixth (two-month) window covers the end of the pilot program. The last three windows cover the post-pilot period. The last rolling window is longer than two years to cover the full sample period ending in December 2013, and our results are similar if we instead use a two-year period, September 2011 to August 2013, for the last window.

We define a dummy variable for each of these rolling windows, $Window_j, j = 1, 2, ..., 9$, and then estimate the following difference-in-differences specification:

$$r_{it} = \gamma_t + \sum_{j=1}^{9} \beta_{wj} Pilot_i \times Window_{jt} + \beta_1 Pilot_i + \epsilon_{it}, \tag{3}$$

where r_{it} is the monthly return of short-leg portfolio i in month t; γ_t is the time fixed effects; $Pilot_i$ is a dummy variable that equals one if portfolio i is formed on pilot firms, and zero otherwise; $Window_{jt}$ is a dummy variable that equals one if month t is in $Window_j$. We carry out this analysis for all 11 anomalies in aggregate, to provide an overall view of the return dynamics on the short-leg portfolios and also to enhance the power of the test.[12] The sample period for this analysis is January 1980 to December 2013.

We obtain nine DiD coefficients β_{wj}, one for each window. Figure 1 plots these coefficients over time, together with their 90% confidence intervals.

In the pre-pilot period, the three DiD coefficients are close to zero and statistically insignificant, which indicates that there was no notable difference between pilot and non-pilot short-leg returns before the pilot program. In the during-pilot period, the DiD coefficient is positive and significant, which is consistent with our main DiD results in Table 4. In unreported results, we define a dummy variable for the transition period between the announcement and introduction of the pilot program (i.e. August 2004 to April 2005) and find that the DiD coefficient for this transition period is also close to zero and statistically insignificant. This suggests that the actual lifting of the uptick rule introduced by the pilot program was needed to reduce overpricing of short legs.

In the post-pilot period, the three DiD coefficients are close to zero and statistically insignificant. This suggests that the difference between pilot and non-pilot short-leg returns vanished in the post-pilot period, as all stocks were exempted from the uptick rule, which is

[12] This aggregation is especially needed for the short two-month windows at the beginning and the end of the pilot program, i.e. for testing Hypothesis 4.

consistent with the results in Table 5.

As for Hypothesis 4, at the beginning of the pilot program, we find that the DiD coefficient is large and negative. It is -1.30% ($t = -4.27$) per month over the two-month window of May to June 2005. This suggests that there was a large initial price decrease on pilot short legs once the uptick rule was lifted, which is consistent with Hypothesis 4 and provides direct evidence that overpricing is reduced once short-sale constraints are relaxed.

At the end of the pilot program, the uptick rule was lifted for non-pilot stocks as well. This relaxation of short-sale constraints should reduce overpricing for non-pilot short legs relative to pilot short legs. Hypothesis 4 predicts a price increase in the short legs of pilot stocks relative to those of non-pilot stocks, i.e. a positive DiD coefficient. Consistent with this, the DiD coefficient, 1.22% ($t = 1.99$), is positive and statistically significant and large in magnitude, over the two-month window of July to August 2007.

Overall, the results presented in Figure 1 support the conclusion that the relaxation of short-sale constraints reduced the mispricing associated with the asset pricing anomalies.

4.5 Shorting Activity

We hypothesize that the introduction of the pilot program will result in an increase in shorting for pilot short-leg portfolios relative to non-pilot short-leg portfolios, and that this increase will be sustained throughout the pilot period. At the end of the pilot program, the difference in shorting activity between pilot and non-pilot short-leg portfolios should decrease, as the uptick rule was also lifted for non-pilot stocks.[13]

[13] Rational arbitrageurs might potentially short preemptively, resulting in a relative increase in shorting on pilot short legs after the announcement but before the introduction of the pilot program. For two reasons, this effect may not be strong enough to leave clear tracks in the data. First, the expected benefit of preemptive shorting is not realized until the pilot program starts; during the interim, it is costly/risky to hold a short position from pilot program announcement to the start of the pilot program without commensurate incremental compensation. Second, since anomaly characteristics tend to revert to the mean over time, the shorting opportunity presented by an anomaly characteristic before the pilot period will on average be weaker by the time the pilot program starts, which reduces the expected amount of post-initiation shorting.

We focus our analysis on the window of January 2005 to October 2007. Over this window, we compare the difference in the amount of short selling between pilot and non-pilot short-leg portfolios using weekly Markit Securities Finance data. For each stock and during each week, we measure the amount of shorting selling as the ratio between the number of shares on loan and the number of shares outstanding. We calculate the portfolio-level amount of short selling as the average amount of short selling across all stocks in the short-leg portfolios. We take the difference in the amount of short selling between pilot and non-pilot short legs, and average this difference across the 11 anomalies. We normalize the difference to be zero at the beginning of the window.

Figure 2 plots the difference (together with its 90% confidence interval) and confirms the hypotheses above. There is no apparent increase in shorting on pilot short legs (relative to non-pilot short legs) until the beginning of the pilot program. The increase is then sustained throughout the pilot period. There is a decrease in the difference around the end of the pilot program, when the uptick rule was also lifted for non-pilot stocks. In unreported analysis, we find that the results using short interest from Compustat as a measure of shorting activity are qualitatively similar.

Overall, this set of results suggests that pilot short-leg portfolios indeed experience more short selling once short-sale constraints are relaxed, and is consistent with our argument that the relaxation of short-sale constraints reduces mispricing associated with the asset pricing anomalies.

5 Robustness Checks and Subsample Analyses

In this section, we present a battery of robustness checks and subsample analyses corresponding to the main results in Table 4.

5.1 Different Sample Periods

We conduct robustness checks with respect to the sample period for our main results. By doing so, we address the possible concern that the during-pilot period is shorter than the pre-pilot period and examine whether our choice of the sample period is critical. Specifically, we explore two shorter sample periods: January 1990 to June 2007 and January 2000 to June 2007, and estimate equation (1) over these two sample periods. The results are presented in Panels A and B of Table 6. The β estimates are qualitatively similar to those in Table 4. In aggregate, the β estimates are both statistically and economically significant for the short leg and the long-short portfolios with the two different sample periods. With all 11 anomalies combined, the coefficient β for the long-short portfolio is -0.53% with a t-statistic of -3.06 from 1990 to 2007 and -0.72% with a t-statistic of -3.46 from 2000 to 2007. The coefficient β for the short leg is 0.50% with a t-statistic of 3.98 from 1990 to 2007 and 0.74% with a t-statistic of 4.98 from 2000 to 2007.

5.2 Placebo Tests

In general, a potential problem with the difference-in-differences method is that the results can be driven by unobservable shocks that affect pilot and non-pilot firms differently, which may then undermine the causal inference of the main difference-in-differences results. The volume-ranking design used to choose pilot firms (in which every third firm in a sorting of firms by trading volume on NYSE and, separately, on AMEX was chosen as a pilot firm) makes the assignment of pilot and non-pilot firms quasi-random and unlikely to be highly correlated with unobserved shocks. Nonetheless, as a precaution, we still conduct a set of falsification tests.

As a first placebo test, we create a pseudo-event in year 2000 and perform a test as if the pseudo-event relaxed short-sale constraints for pilot firms.[14] To mimic the actual pilot

[14] We choose year 2000 for a pseudo-event as it is prior to the real event and the pseudo pilot program created accordingly does not overlap with the real pilot program. In untabulated results, we find similar

program closely, we assume that this pseudo pilot program was effective from May 2001 to July 2003. We then run the difference-in-differences regression as follows:

$$r_{it} = \gamma_t + \beta Pilot_i \times PseudoDuring_t + \beta_1 Pilot_i + \epsilon_{it}, \tag{4}$$

where $PseudoDuring_t$ is a dummy variable that equals one if month t is between July 2001 and June 2003, i.e., when the pseudo-event was effective, and other notations are defined exactly the same as in equation (1). Following closely the main analysis in Table 4, the sample period is from January 1980 to June 2003 and we drop the two months at the beginning of the pseudo pilot program (May and June 2001) from the sample.

The results of the placebo test are presented in Table 7. The coefficients on $Pilot_i \times PseudoDuring_t$ are mostly statistically insignificant and have mixed signs. When all 11 anomalies are combined, the coefficients are much smaller in magnitude than those in Table 4 for the short leg and the long-short portfolio and are statistically insignificant. The placebo test results therefore suggest that our main results are unlikely to be driven by unobserved shocks that affect pilot and non-pilot firms differently.

We further carry out a set of more systematic falsification tests, in which we recreate the stratified trading volume design the SEC used in the pilot program for rolling synthetic samples and examine how likely our main results can arise in a random two-year period.

Specifically, at the end of June of each year τ (from 1991 to 2000), we create the stratified sample following closely the trading volume design the SEC used for the pilot program. We select the largest (in terms of market capitalization at the end of June of year τ) 3,000 stocks. We then rank stocks on each of NYSE, AMEX, and Nasdaq based on their average daily trading volume during the prior year (June of year $\tau - 1$ through May of year τ) and select every third stock (beginning with the second one) as a pilot stock and the remainder as non-pilot stocks. We then re-do our main DiD analysis for these rolling stratified samples,

results when we create a pseudo-event in e.g. year 1998 or year 1999.

and we keep only NYSE and AMEX stocks to be consistent with the main analysis. For each sample created at the end of June of year τ, we define a pseudo pilot program that was effective from May of year $\tau+1$ to July of year $\tau+3$. We then run the following regression:

$$r_{it} = \gamma_t + \beta Pilot_i \times PseudoDuring_{\tau,t} + \beta_1 Pilot_i + \epsilon_{it}, \tag{5}$$

where $PseudoDuring_{\tau,t}$ is a dummy variable that equals one if month t is between July of year $\tau+1$ and June of year $\tau+3$, and other notations are defined exactly the same as in equations (1) and (4). The sample period is from January 1980 to June of year $\tau+3$. We drop the two months at the beginning of the pseudo pilot program (May and June of year $\tau+1$) from the sample.

None of these falsification tests generates results that are similar to our main results in Table 4. The most positive β for the short-leg portfolios (all 11 anomalies in aggregate) is 0.18% ($t = 0.54$), which is much weaker than our main result, 0.63% ($t = 5.31$). The most negative β for the long-short portfolios (all 11 anomalies in aggregate) is -0.10% ($t = -0.38$), which is much weaker than our main result, -0.72% ($t = -4.37$). This evidence further suggests that our main results are unlikely to arise by chance.

5.3 Results from Nasdaq Stocks

Our main empirical analysis is carried out on the sample of pilot and non-pilot stocks traded on NYSE/AMEX. In this subsection, we conduct a falsification test using the sample of pilot and non-pilot stocks traded on Nasdaq. As stated in the introduction and Section 2, the pilot program also removed the bid price test for pilot stocks traded on Nasdaq. However, the bid price test for Nasdaq stocks is not very restrictive, and a significant fraction of trading volume in Nasdaq-listed stocks is executed on ArcaEx and INET, which do not enforce the bid price test (see e.g. Diether, Lee, and Werner (2009)). We therefore expect at most a minimal effect of the pilot program on anomaly returns of Nasdaq-listed stocks.

This falsification test helps rule out a potential alternative explanation for our main results. Specifically, one may argue that the pilot program made pilot stocks more salient to investors, and the increase in investor attention to pilot stocks weakened anomalies, driving our main results. Since both Nasdaq pilot stocks and NYSE/AMEX pilot stocks were included in the pilot program, under this explanation, the same shift in investor attention would also occur for Nasdaq pilot stocks during the same pilot period. It follows that if this salience mechanism were driving our main results, we should observe an apparent effect of the pilot program on anomaly returns of Nasdaq stocks similar to that on anomaly returns of NYSE/AMEX stocks. On the other hand, if the relaxation of short-sale constraints drives our main results, we would expect to see a minimal effect on anomaly returns of Nasdaq stocks.

We repeat our main DiD analysis (equation (1)) on the sample of pilot and non-pilot stocks traded on Nasdaq National Market. The results are reported in Table 8. The DiD coefficient β is mostly statistically insignificant and has mixed signs across the 11 anomalies, for the long-leg, short-leg, and long-short returns. In aggregate, the coefficient is also small and insignificant for the long-leg, short-leg, and long-short returns. Overall, these results suggest that the pilot program had little effect on anomaly returns for Nasdaq stocks and confirm that our main results derive from the relaxation of short-sale constraints generated by the pilot program.

5.4 The Effect of Short-Sale Constraints on Mispricing of Different Kinds of Stocks

We next explore whether our main results documented in Table 4 are different among different classes of stocks. Diether, Lee, and Werner (2009) argue that small and less liquid stocks are likely to be more affected by the suspension of the uptick rule, i.e., the uptick rule impeded short selling more for small and less liquid stocks. The reason is that these stocks have wider spreads and therefore short sellers have to become liquidity providers to ensure compliance with the uptick rule, which makes short-sale orders more passive in the presence of the uptick rule. In addition, for small stocks, a penny tick may be a more significant impediment to shorting them. Consistent with this argument, Diether, Lee, and Werner (2009) find that the suspension of the uptick rule has a greater effect on spreads and some intraday volatility measures for small and less liquid stocks. In our context, we test whether the effect of the pilot program on asset pricing anomalies is more pronounced for small and less liquid stocks. Furthermore, we directly construct a stock-level measure that captures the restrictiveness of the uptick rule using TAQ data and test whether the effect of the pilot program on asset pricing anomalies is more pronounced for stocks that are more restricted by the uptick rule.

We first explore the difference in the effects of the pilot program on small versus large stocks. We split our main sample into two subsamples of small and large stocks based on their market capitalization at the end of April 2005, before the beginning of the pilot program. Large stocks are those with market capitalization above the median and small stocks are those with market capitalization below the median. We then form anomaly decile portfolios using pilot/non-pilot stocks in these subsamples, and repeat the main difference-in-differences analysis (equation (1)) for each subsample. The results are presented in Table 9.

Comparing Panels A and B of Table 9, the effect of the pilot program on asset pricing anomalies is indeed more pronounced among small stocks. When we aggregate over all 11 anomalies, the pilot program reduced the long-short portfolio return by 95 basis points for small stocks compared with 47 basis points for large stocks. The DiD estimate on the short legs is 88 basis points for small stocks compared with 37 basis points for large stocks. The results from subsample analysis split by liquidity are similar to those in Table 9 and are therefore omitted for brevity.

Next, we conduct subsample analysis split by the uptick-rule restrictiveness. We measure the uptick-rule restrictiveness in April 2005 (right before the beginning of the pilot program) using TAQ data as follows.

For each stock on each trading day of April 2005 and at a given time of day, we calculate the minimum shortable price that complies with the uptick rule. We then compare the minimum shortable price with the current bid. If the minimum shortable price is lower than or equal to the bid, then a short seller can potentially execute a short sale transaction at this price successfully. For each stock on each day of April 2005, we calculate the frequency of these potential short sale transactions under the restriction of the uptick rule. Taking the average of this frequency over all trading days in April 2005, we obtain the stock-level measure of uptick-rule restrictiveness. A stock for which short sales can be carried out more frequently has a lower degree of uptick-rule restrictiveness. We then split our main sample into two subsamples based on their uptick-rule restrictiveness in April 2005. More restricted stocks have above-median restrictiveness and less restricted stocks have below-median restrictiveness. For more restricted stocks, the average waiting time for a short seller before she can short is 1.62 days, which seems a meaningful restriction. The average magnitude of price change (return) during this waiting time is 1.90%. We form anomaly decile portfolios using pilot/non-pilot stocks in these subsamples, and repeat the main difference-in-differences analysis (equation (1)) for each subsample. The results are presented in Table 10.

Comparing Panels A and B of Table 10, the effect of the pilot program on asset pricing anomalies is indeed more pronounced among stocks that are more restricted by the uptick rule. It is also useful to note that the difference between the two subsamples is slightly larger than the difference between small and large stocks in Table 9. When we aggregate over all 11 anomalies, the pilot program reduced the long-short portfolio return by 109 basis points for more restricted stocks compared with 33 basis points for less restricted stocks. The DiD estimate on the short legs is 96 basis points for more restricted stocks compared with 30 basis points for less restricted stocks. The results from this subsample analysis further suggest that the lift of the uptick rule is the underlying driving force of our main results.

6 Conclusion

Using the pilot program of Regulation SHO, which relaxed short-sale constraints for a quasi-random set of stocks, we examine the causal effect of limits to arbitrage, and in particular short sale constraints, on 11 well-known asset pricing anomalies. We find that the long-short strategies for the 11 anomalies produced much smaller abnormal returns on portfolios constructed with pilot stocks during the pilot period. This suggests that these anomalies reflect mispricing, and that making arbitrage easier reduces such mispricing. The effect of the pilot program is only significant for the short legs of the anomaly long-short portfolios, which is consistent with the prediction that easy short arbitrage weakens the short side of the anomalies.

Furthermore, we show that the difference in anomaly returns between pilot and non-pilot stocks vanished after the pilot program ended, as the difference in short-sale constraints between pilot and non-pilot stocks disappeared. We also show that pilot short legs underperformed (outperformed) non-pilot short legs at the beginning (end) of the pilot program. Lastly, we find that the difference in anomaly portfolio returns between pilot and non-pilot stocks during the pilot period was more pronounced among small stocks and stocks that were more restricted by the uptick rule before the pilot program. Taken together, these findings provide strong and clear-cut confirmation that limits to arbitrage have a *causal* effect on the strength of well-known asset pricing anomalies and that these anomalies reflect to a large extent mispricing.

A Appendix

A.1 Definition of Anomaly Variables

The data to construct anomaly variables are from CRSP and annual and quarterly Compustat.

Anomaly 1: Momentum (RET). The past return RET_t for a stock is calculated as the compounded return over the 11-month ranking period $t-12$ to $t-2$.

Anomaly 2: Gross profitability (GP/A). The gross profitability measure GP/A_t for a firm is calculated as the difference between total revenue ($REVT_t$) and cost of goods sold ($COGS_t$), scaled by total assets AT_t.

Anomaly 3: Asset growth (AG). The asset growth measure AG_t for a firm is calculated as the change in total assets $AT_t - AT_{t-1}$, scaled by lagged total assets AT_{t-1}.

Anomaly 4: Investment to assets (IVA). Investment to assets IVA_t is defined as the annual change in gross property, plant, and equipment plus the annual change in inventories, $PPEGT_t - PPEGT_{t-1} + INVT_t - INVT_{t-1}$, scaled by lagged total assets AT_{t-1}.

Anomaly 5: Return on assets (ROA). Return on assets ROA_t is measured as the quarterly earnings, or income before extraordinary item, IBQ_t, scaled by quarterly total assets ATQ_t.

Anomaly 6: Net operating assets (NOA). Net operating assets are calculated as the difference between operating assets and operating liabilities, scaled by lagged total assets: $NOA_t = (Operating\ Assets_t - Operating\ Liabilities_t)/AT_{t-1}$, where $Operating\ Assets = Total\ Assets\ (AT) - Cash\ and\ Short\text{-}term\ Investment\ (CHE)$, and $Operating\ Liabilities = Total\ Assets\ (AT) - Short\text{-}term\ Debt\ (DLC) - Long\text{-}term\ Debt\ (DLTT)$ $- Minority\ Interests\ (MIB) - Preferred\ Stocks\ (PSTK) - Common\ Equity\ (CEQ)$.

Anomaly 7: Accruals (AC). Operating accruals are measured as the change in non-cash current assets, less the change in current liabilities (exclusive of short-term debt and taxes payable), and less depreciation and amortization, all divided by lagged total assets: $Accruals_t = [(\Delta Current\ Assets - \Delta Cash) - (\Delta Current\ Liabilities - \Delta Short\text{-}term\ Debt - \Delta Taxes\ Payable) - Depreciation\ and\ Amortization\ Expense)]/AT_{t-1}$. In terms of Compustat item notations, *Current Assets* is ACT, *Cash* is CHE, *Current Liabilities* is LCT, *Short-term Debt* is DLC, *Taxes Payable* is TXP, and *Depreciation and Amortization* is DP.

Anomaly 8: Net stock issues (NSI). Net stock issues on the annual basis are measured as the change in the natural logarithm of a firm's split-adjusted shares over the last year, $NSI_t = Ln(Adjusted\ Shares_t) - Ln(Adjusted\ Shares_{t-1})$, where $Adjusted\ Shares_t$ is the product of the common shares outstanding ($CSHO_t$) and the adjustment factor ($AJEX_t$).

Anomaly 9: Composite equity issues (CEI). Composite equity issues are measured over the past five-year window and are defined as the part of the growth rate in market equity not attributable to stock returns, $CEI_t = Ln(ME_t/ME_{t-5}) - r(t-5, t)$. In June of year t, for example, ME_t is the market equity at the end of June of year t and ME_{t-5} is the market equity at the end of June of year $t-5$, while $r(t-5, t)$ is the cumulative log return of the stock from the end of June of year $t-5$ to the end of June of year t.

Anomaly 10: Failure probability (FP). Following Campbell, Hilscher, and Szilagyi (2008), we use coefficients in Column 4 of their Table IV to construct the measure of FP_t, which is related to the failure probability through a monotonic transformation. Specifically, FP_t is measured as:

$$\begin{aligned} FP_t = & -9.16 - 20.26NIMTAAVG_t + 1.42TLMTA_t - 7.13EXRETAVG_t \\ & + 1.41SIGMA_t - 0.045RSIZE_t - 2.13CASHMTA_t + 0.075MB_t - 0.058PRICE_t, \end{aligned}$$

where the details of variables can be found in Campbell, Hilscher, and Szilagyi (2008).

Anomaly 11: O-score (OS). Following Ohlson (1980) and Chen, Novy-Marx, and Zhang (2011), we construct the O-score as:

$$OS_t = -1.32 - 0.407log(ADJASSET_t) + 6.03TLTA_t - 1.43WCTA_t + 0.076CLCA_t$$

$$- 1.72 OENEG_t - 2.37 NITA_t - 1.83 FUTL_t + 0.285 INTWO_t - 0.521 CHIN_t,$$

where $ADJASSET$ is the adjusted total assets and equals total assets (ATQ)+0.1× (market equity-book equity). $TLTA$ is the leverage ratio and equals the book value of debt (DLCQ plus DLTTQ) divided by $ADJASSET$. $WCTA$ is working capital (ACTQ minus LCTQ) divided by $ADJASSET$. $CLCA$ is current liabilities (LCTQ) divided by current assets (ACTQ). $OENEG$ is one if total liabilities (LTQ) exceeds total assets (ATQ) and is zero otherwise. $NITA$ is net income (NIQ) divided by $ADJASSET$. $FUTL$ is the fund provided by operations (PIQ) divided by liabilities (LTQ). $INTWO$ is equal to one if net income (NIQ) is negative for the last two quarters and zero otherwise. $CHIN$ is $(NI_t - NI_{t-1})/(|NI_t| + |NI_{t-1}|)$, where NI is net income (NIQ).[15]

A.2 Benchmark-Adjusted Returns

In the main empirical analysis, we use raw portfolio returns as the dependent variable in equation (1). In this subsection, we test the validity of our main results when we use benchmark-adjusted portfolio returns as the dependent variable in equation (1). If the mean return premia of the Fama-French three factors represent rational risk premia (an issue on which we do not take a stand here), then this analysis would verify whether relaxation of short sale constraints reduces mispricing measured against this benchmark.

To obtain benchmark-adjusted returns, we first regress the time series of excess returns onto the time series of appropriate factors (the market excess return for the CAPM, and two additional factors, the SMB and HML factors, for the Fama-French three-factor model). We then obtain the time series of benchmark-adjusted returns as the constant plus the residuals from the regression.

[15]Further details of these variables can be found in the Appendix of Chen, Novy-Marx, and Zhang (2011). They have $-0.407 log(ADJASSET_t/CPI_t)$, where CPI is the consumer price index, as the second term of O-score. Since scaling $ADJASSET$ by CPI or not does not affect the cross-sectional sorting of stocks based on O-score, we drop CPI in our calculation of O-score.

We would not expect this factor adjustment to change the results much, since the selection of pilot firms is quasi-random, implying similar loadings on the benchmark factors of pilot versus non-pilot firms. The results for benchmark-adjusted returns are presented in Table A.2. Consistent with this intuition, all the β estimates are similar to those in Table 4 of main results.

A.3 Power of the Main Tests

Our main tests for Hypotheses 1 and 2 based on equation (1) use a relatively short (two years) pilot period to estimate the effect of relaxing short-sale constraints on anomaly returns. Here, we address the issue of power explicitly. Intuitively, our main tests using equation (1) gain power by two means. First is the aggregation across 11 anomalies. Second, even for a single anomaly during the pilot period, what is relevant for our test is not the raw strength of that anomaly, it is the *difference* in strength of an anomaly between pilot versus non-pilot firms over the same time period.[16] This differencing effectively hedges away much of the factor volatility of returns, greatly increasing the precision of the test. To see this in a very simple way, suppose that the momentum return of Portfolio A were equal to the return of Portfolio B plus a constant. Then even if both portfolios were highly volatile, the difference in returns would be a constant, implying that the difference would be significant with an infinite t statistic. Of course a constant difference is unrealistic, but this example illustrates that testing for a difference filters out a large amount of variability from the test.

Consistent with this point, in untabulated results, we show that taking differences between portfolios constructed with pilot and non-pilot stocks substantially reduces return volatility. Monthly standard deviations of return differences between long-leg/short-leg portfolios constructed with pilot and non-pilot stocks are much smaller than those of returns on long-leg/short-leg portfolios themselves. For example, averaged across the 11 anomalies, the

[16] Econometrically, this is achieved by including anomaly-time fixed effects in our regression specification. Also, our tests actually examine the difference in this difference between the pilot and non-pilot periods, but this is not crucial for our argument.

monthly standard deviation of return differences between pilot and non-pilot stocks for the short leg is 1.71%, while the monthly standard deviation of short leg returns is 3.64% for non-pilot stocks and 3.70% for pilot stocks.

This contrasts with conventional tests for estimating average anomaly returns (rather than differences in returns), in which sampling noise derived from factor realizations reduces power. In such tests, much longer time periods are often needed to confirm an anomaly reliably. It is of course sometimes possible to identify anomaly returns using a sample period measured in years rather than decades. For example, in an out-of-sample test of their 1993 paper (Jegadeesh and Titman (1993)), Jegadeesh and Titman (2001) find significant momentum in the sample period of 1990 to 1998 (9 years), with a t-statistic of 4.71.

A further consideration which enhances the power of our tests is that the pilot period is one in which the 11 anomalies are relatively strong. If we estimate the mean anomaly returns (long-minus-short returns) on *non-pilot stocks* during the pilot period, the anomalies tend to be stronger, both economically and statistically, than might ordinarily be expected for a two-year period.

Specifically, we calculate the mean monthly anomaly returns and CAPM/Fama-French-three-factor alphas for the 11 anomalies individually and in aggregate, over the pilot period from July 2005 to June 2007. Table A.3 presents the results for non-pilot stocks. It shows that the anomaly returns and alphas of the 11 anomalies for non-pilot stocks are mostly positive (31 out of 33). When we combine the 11 anomalies together, both the mean return and alphas are positive and statistically significant. The magnitudes are also large. The mean monthly return and alphas are about 51-69 bps, when the 11 anomalies are combined.

Table A.1: **Results using value-weighted portfolio returns**

This table reports the coefficient β from the regression $r_{it} = \gamma_t + \beta Pilot_i \times During_t + \beta_1 Pilot_i + \epsilon_{it}$, for the 11 anomalies individually and all of them in aggregate. The portfolio returns are value weighted. The sample consists of non-pilot and pilot stocks from the pilot program that are traded on NYSE/AMEX. The sample period is January 1980 to June 2007. We drop two months at the beginning of the pilot program (May and June 2005) from the sample. The unit of β is percentage. Robust t-statistics are presented in the parentheses below the coefficient estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Long leg	Short leg	Long-short
Momentum	0.29	-0.33	0.62
	(0.56)	(-0.61)	(0.75)
Gross profitability	-0.17	0.70	-0.87
	(-0.44)	(1.42)	(-1.39)
Asset growth	-0.33	0.26	-0.59
	(-0.85)	(0.58)	(-1.11)
Investment to assets	0.10	0.34	-0.24
	(0.24)	(0.78)	(-0.40)
Return on assets	0.50	-0.22	0.72
	(1.42)	(-0.39)	(1.23)
Net operating assets	-0.34	0.60	-0.94
	(-0.87)	(1.30)	(-1.36)
Accruals	-0.14	0.54	-0.68
	(-0.26)	(1.20)	(-1.31)
Net stock issues	-0.12	0.02	-0.14
	(-0.37)	(0.04)	(-0.30)
Composite equity issues	-0.52	0.01	-0.53
	(-1.32)	(0.02)	(-0.80)
Failure probability	0.41	0.33	0.07
	(1.14)	(0.54)	(0.11)
O-score	-0.45	0.52	-0.97
	(-1.07)	(1.13)	(-1.60)
Combination	-0.07	0.25*	-0.32*
	(-0.56)	(1.66)	(-1.70)

Table A.2: **Benchmark-adjusted returns**

This table reports the coefficient β from the regression $r_{it} = \gamma_t + \beta Pilot_i \times During_t + \beta_1 Pilot_i + \epsilon_{it}$, for the 11 anomalies individually and all of them in aggregate, with benchmark-adjusted returns used as the dependent variable. Panel A displays results for CAPM-adjusted returns while Panel B displays results for Fama-French-three-factor-adjusted (FF-adjusted) returns. The sample consists of non-pilot and pilot stocks from the pilot program that are traded on NYSE/AMEX. The sample period is January 1980 to June 2007. We drop two months at the beginning of the pilot program (May and June 2005) from the sample. The unit of β is percentage. Robust t-statistics are presented in the parentheses below the coefficient estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Panel A: CAPM-adjusted returns			Panel B: FF-adjusted returns		
	Long leg	Short leg	Long-short	Long leg	Short leg	Long-short
Momentum	0.03	0.65	-0.62	0.01	0.66	-0.65
	(0.07)	(1.37)	(-0.92)	(0.03)	(1.39)	(-0.93)
Gross profitability	-0.04	0.81**	-0.85*	-0.05	0.80**	-0.86*
	(-0.13)	(2.24)	(-1.81)	(-0.17)	(2.23)	(-1.85)
Asset growth	-0.18	0.42	-0.60	-0.19	0.43*	-0.62
	(-0.46)	(1.64)	(-1.40)	(-0.48)	(1.67)	(-1.41)
Investment to assets	-0.01	0.15	-0.16	-0.02	0.16	-0.19
	(-0.01)	(0.43)	(-0.27)	(-0.06)	(0.45)	(-0.30)
Return on assets	0.05	0.74	-0.69	0.03	0.73	-0.70
	(0.23)	(1.26)	(-1.08)	(0.14)	(1.25)	(-1.09)
Net operating assets	-0.11	0.88***	-0.99**	-0.12	0.87***	-1.00**
	(-0.32)	(3.50)	(-2.11)	(-0.34)	(3.46)	(-2.10)
Accruals	-0.51	0.80**	-1.31***	-0.51	0.80**	-1.31***
	(-1.41)	(2.29)	(-2.61)	(-1.45)	(2.24)	(-2.73)
Net stock issues	-0.08	0.77***	-0.85**	-0.08	0.79***	-0.86**
	(-0.26)	(3.24)	(-2.25)	(-0.26)	(3.22)	(-2.26)
Composite equity issues	-0.56*	0.54*	-1.10***	-0.55*	0.54*	-1.09***
	(-1.80)	(1.78)	(-2.73)	(-1.79)	(1.78)	(-2.75)
Failure probability	0.16	0.45	-0.29	0.14	0.45	-0.30
	(0.56)	(1.12)	(-0.62)	(0.48)	(1.12)	(-0.64)
O-score	0.36	0.80	-0.44	0.35	0.79	-0.43
	(0.76)	(1.39)	(-0.54)	(0.73)	(1.40)	(-0.54)
Combination	-0.08	0.64***	-0.72***	-0.09	0.64***	-0.73***
	(-0.76)	(5.34)	(-4.36)	(-0.83)	(5.36)	(-4.39)

Table A.3: **Asset pricing anomalies during the pilot period**

This table presents the mean monthly raw return, the CAPM α, and the Fama-French three factor α of the 11 asset pricing anomalies individually and in aggregate, during the pilot period from July 2005 to June 2007. The sample consists of non-pilot and pilot stocks from the pilot program that are traded on NYSE/AMEX. The anomaly portfolios are constructed with non-pilot stocks. Robust t-statistics are presented in the parentheses below the estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Raw	CAPM α	FF α
Momentum	1.02	1.11	0.84
	(1.47)	(1.24)	(0.92)
Gross profitability	0.47	0.24	0.40
	(1.24)	(0.64)	(1.06)
Asset growth	0.50	0.71**	0.82**
	(1.33)	(2.26)	(2.27)
Investment to assets	0.31	0.66	1.02*
	(0.62)	(1.35)	(1.85)
Return on assets	0.83	0.82	0.79
	(1.54)	(1.27)	(1.09)
Net operating assets	0.63	0.86**	1.16***
	(1.71)	(2.09)	(2.99)
Accruals	-0.11	0.17	-0.27
	(-0.24)	(0.31)	(-0.53)
Net stock issues	0.42	0.53	0.74**
	(1.30)	(1.63)	(2.24)
Composite equity issues	0.17	0.42	0.43
	(0.54)	(1.27)	(1.33)
Failure probability	1.02*	1.20*	1.31**
	(1.85)	(1.75)	(2.12)
O-score	0.39	0.37	0.38
	(0.88)	(0.93)	(0.96)
Combination	0.51***	0.64***	0.69***
	(2.82)	(3.12)	(3.22)

References

Ali, A., L.-S. Hwang, and M. A. Trombley, 2003, "Arbitrage risk and the book-to-market anomaly," *Journal of Financial Economics*, 69(2), 355–373.

Ali, A., and M. A. Trombley, 2006, "Short sales constraints and momentum in stock returns," *Journal of Business Finance & Accounting*, 33(3-4), 587–615.

Asparouhova, E., H. Bessembinder, and I. Kalcheva, 2010, "Liquidity biases in asset pricing tests," *Journal of Financial Economics*, 96(2), 215–237.

——— , 2013, "Noisy prices and inference regarding returns," *Journal of Finance*, 68(2), 665–714.

Asquith, P., P. A. Pathak, and J. R. Ritter, 2005, "Short interest, institutional ownership, and stock returns," *Journal of Financial Economics*, 78(2), 243–276.

Baker, M., and S. Savaşoglu, 2002, "Limited arbitrage in mergers and acquisitions," *Journal of Financial Economics*, 64(1), 91–115.

Barberis, N., and R. Thaler, 2003, "A survey of behavioral finance," *Handbook of the Economics of Finance*, 1, 1053–1128.

Beneish, M. D., C. M. Lee, and D. C. Nichols, 2015, "In short supply: Short-sellers and stock returns," *Journal of Accounting and Economics*, 60(2), 33–57.

Campbell, J. Y., J. Hilscher, and J. Szilagyi, 2008, "In search of distress risk," *Journal of Finance*, 63(6), 2899–2939.

Cao, B., D. S. Dhaliwal, A. C. Kolasinski, and A. V. Reed, 2007, "Bears and numbers: Investigating how short sellers exploit and affect earnings-based pricing anomalies," *Available at SSRN 748506.*

Chen, J., H. Hong, and J. C. Stein, 2002, "Breadth of ownership and stock returns," *Journal of Financial Economics*, 66(2), 171–205.

Chen, L., R. Novy-Marx, and L. Zhang, 2011, "An alternative three-factor model," *Available at SSRN 1418117.*

Cohen, L., K. B. Diether, and C. J. Malloy, 2007, "Supply and demand shifts in the shorting market," *Journal of Finance*, 62(5), 2061–2096.

Cooper, M. J., H. Gulen, and M. J. Schill, 2008, "Asset growth and the cross-section of stock returns," *Journal of Finance*, 63(4), 1609–1651.

Daniel, K., and S. Titman, 2006, "Market reactions to tangible and intangible information," *Journal of Finance*, 61(4), 1605–1643.

D'Avolio, G., 2002, "The market for borrowing stock," *Journal of Financial Economics*, 66(2), 271–306.

Dichev, I. D., 1998, "Is the risk of bankruptcy a systematic risk?," *Journal of Finance*, 53(3), 1131–1147.

Diether, K. B., K.-H. Lee, and I. M. Werner, 2009, "It's SHO time! Short-sale price tests and market quality," *Journal of Finance*, 64(1), 37–73.

Drechsler, I., and Q. F. Drechsler, 2014, "The shorting premium and asset pricing anomalies," *Available at SSRN 2387099.*

Engelberg, J., A. V. Reed, and M. Ringgenberg, 2017, "Short selling risk," *Journal of Finance, Forthcoming.*

Fama, E. F., and K. R. French, 1992, "The cross-section of expected stock returns," *Journal of Finance*, 47(2), 427–465.

——— , 1993, “Common risk factors in the returns on stocks and bonds,” *Journal of Financial Economics*, 33(1), 3–56.

——— , 2006, “Profitability, investment and average returns,” *Journal of Financial Economics*, 82(3), 491–518.

Fang, V. W., A. H. Huang, and J. M. Karpoff, 2016, “Short selling and earnings management: A controlled experiment,” *Journal of Finance*, 71(3), 1251–1294.

Geczy, C. C., D. K. Musto, and A. V. Reed, 2002, “Stocks are special too: An analysis of the equity lending market,” *Journal of Financial Economics*, 66(2), 241–269.

Gromb, D., and D. Vayanos, 2010, “Limits of arbitrage,” *Annual Review of Financial Economics*, 2(1), 251–275.

Grullon, G., S. Michenaud, and J. P. Weston, 2015, “The real effects of short-selling constraints,” *Review of Financial Studies*, 28(6), 1737–1767.

Harvey, C. R., Y. Liu, and H. Zhu, 2016, “ and the cross-section of expected returns,” *Review of Financial Studies*, 29(1), 5–68.

Hirshleifer, D., 2001, “Investor psychology and asset pricing,” *Journal of Finance*, 56(4), 1533–1597.

Hirshleifer, D., K. Hou, S. H. Teoh, and Y. Zhang, 2004, “Do investors overvalue firms with bloated balance sheets?,” *Journal of Accounting and Economics*, 38, 297–331.

Hirshleifer, D., S. H. Teoh, and J. J. Yu, 2011, “Short arbitrage, return asymmetry and the accrual anomaly,” *Review of Financial Studies*, 24(7), 2429–2461.

Israel, R., and T. J. Moskowitz, 2013, “The role of shorting, firm size, and time on market anomalies,” *Journal of Financial Economics*, 108(2), 275–301.

Jegadeesh, N., and S. Titman, 1993, "Returns to buying winners and selling losers: Implications for stock market efficiency," *Journal of Finance*, 48(1), 65–91.

———, 2001, "Profitability of momentum strategies: An evaluation of alternative explanations," *Journal of Finance*, 56(2), 699–720.

Jones, C. M., and O. A. Lamont, 2002, "Short-sale constraints and stock returns," *Journal of Financial Economics*, 66(2), 207–239.

Lam, F. E. C., and K. J. Wei, 2011, "Limits-to-arbitrage, investment frictions, and the asset growth anomaly," *Journal of Financial Economics*, 102(1), 127–149.

Lamont, O. A., and R. H. Thaler, 2003, "Can the market add and subtract? Mispricing in tech stock carve-outs," *Journal of Political Economy*, 111(2), 227–268.

Li, Y., and L. Zhang, 2015, "Short selling pressure, stock price behavior, and management forecast precision: Evidence from a natural experiment," *Journal of Accounting Research*, 53(1), 79–117.

Loughran, T., and J. R. Ritter, 1995, "The new issues puzzle," *Journal of Finance*, 50(1), 23–51.

Mashruwala, C., S. Rajgopal, and T. Shevlin, 2006, "Why is the accrual anomaly not arbitraged away? The role of idiosyncratic risk and transaction costs," *Journal of Accounting and Economics*, 42(1), 3–33.

Nagel, S., 2005, "Short sales, institutional investors and the cross-section of stock returns," *Journal of Financial Economics*, 78(2), 277–309.

Novy-Marx, R., 2013, "The other side of value: The gross profitability premium," *Journal of Financial Economics*, 108(1), 1–28.

Ohlson, J. A., 1980, "Financial ratios and the probabilistic prediction of bankruptcy," *Journal of Accounting Research*, 18(1), 109–131.

Pontiff, J., 1996, "Costly arbitrage: Evidence from closed-end funds," *Quarterly Journal of Economics*, 111(4), 1135–1151.

Pontiff, J., and A. Woodgate, 2008, "Share issuance and cross-sectional returns," *Journal of Finance*, 63(2), 921–945.

Reed, A. V., 2013, "Short selling," *Annual Review of Financial Economics*, 5(1), 245–258.

Saffi, P. A., and K. Sigurdsson, 2010, "Price efficiency and short selling," *Review of Financial Studies*, 24(3), 821–852.

Shleifer, A., and R. W. Vishny, 1997, "The limits of arbitrage," *Journal of Finance*, 52(1), 35–55.

Sloan, R., 1996, "Do stock prices fully reflect information in accruals and cash flows about future earnings?," *Accounting Review*, 71(3), 289–315.

Stambaugh, R. F., J. Yu, and Y. Yuan, 2012, "The short of it: Investor sentiment and anomalies," *Journal of Financial Economics*, 104(2), 288–302.

Titman, S., K.-C. Wei, and F. Xie, 2004, "Capital investments and stock returns," *Journal of Financial and Quantitative Analysis*, 39(04), 677–700.

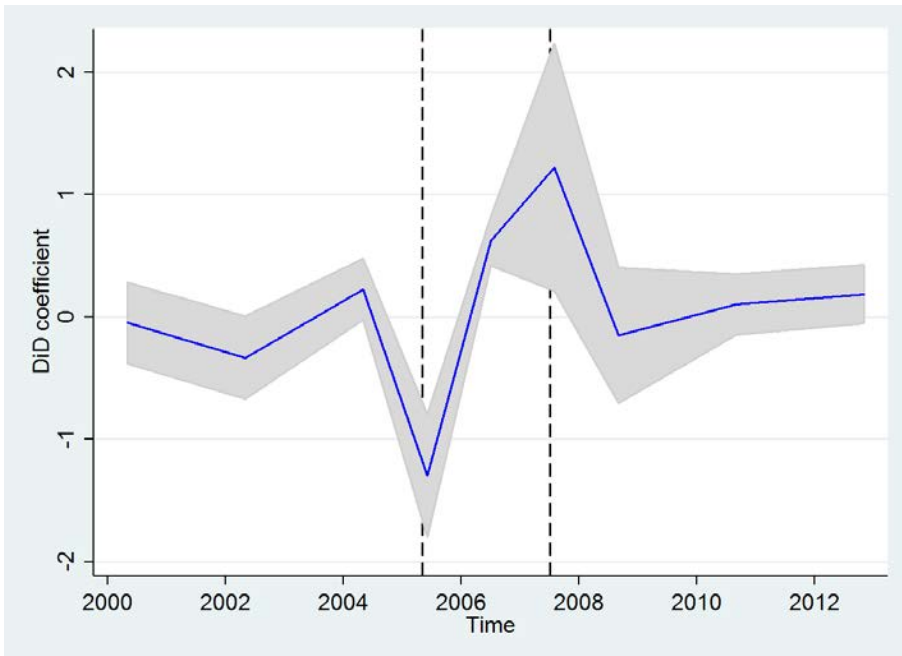

Figure 1: **Return dynamics of short-leg portfolios.** This figure illustrates the return dynamics of pilot short-leg portfolios relative to non-pilot short-leg portfolios, over nine rolling windows that cover the pre-pilot, during-pilot, and post-pilot periods, and the beginning and end of the pilot program. The regression $r_{it} = \gamma_t + \sum_{j=1}^{9} \beta_{wj} Pilot_i \times Window_{jt} + \beta_1 Pilot_i + \epsilon_{it}$ is carried out for the 11 anomalies in aggregate, where r_{it} is the monthly return of short-leg portfolio i in month t; γ_t is the time fixed effects; $Pilot_i$ is a dummy variable that equals one if portfolio i is formed on pilot firms, and zero otherwise; $Window_{jt}$ is a dummy variable that equals one if month t is in $Window_j$ (see text for the timing of the 9 windows). The sample period for this analysis is January 1980 to December 2013. The blue solid line plots the nine DiD coefficients β_{wj} in percentage terms. The shaded area shows their 90% confidence intervals. The two dashed vertical lines denote the beginning and the end of the pilot program.

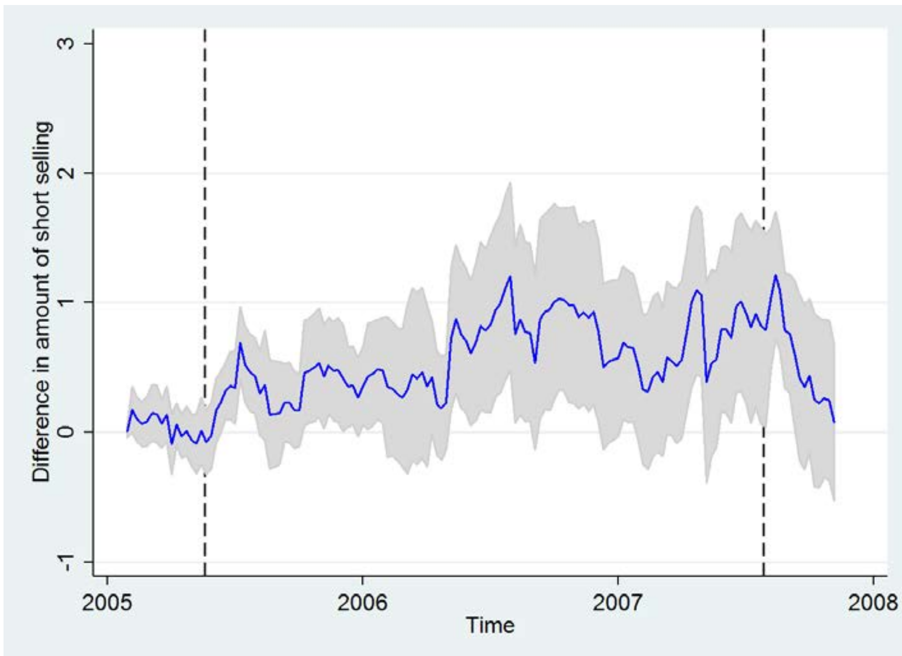

Figure 2: **Difference in shorting activity on short-leg portfolios between pilot and non-pilot stocks.** This figure shows the difference in the amount of short selling between pilot and non-pilot short-leg portfolios from January 2005 to October 2007. The difference is normalized to be zero at the beginning. We use the weekly Markit Securities Finance data. For each stock and during each week, we measure the amount of shorting selling as the ratio between the number of shares on loan and the number of shares outstanding. We calculate the portfolio-level amount of short selling as the average amount of short selling across all stocks in the short-leg portfolios. We then take the difference in portfolio-level amount of shorting selling between pilot and non-pilot short legs, and average this difference across the 11 anomalies. The solid line plots the average difference in percentage terms, and the shaded area shows the 90% confidence intervals. The two dashed vertical lines denote the beginning and the end of the pilot program.

Table 1: **Characteristics of stocks for the 11 anomalies**

This table summarizes the characteristics of stocks in the long leg (the highest performing group) and those in the short leg (the lowest performing group) for the 11 anomalies.

	Stocks in the long leg	Stocks in the short leg
Momentum	High past return	Low past return
Gross profitability	High gross profitability	Low gross profitability
Asset growth	Low asset growth	High asset growth
Investment to assets	Low investment to assets	High investment to assets
Return on assets	High return on assets	Low return on assets
Net operating assets	Low net operating assets	High net operating assets
Accruals	Low accruals	High accruals
Net stock issues	Low equity issuance	High equity issuance
Composite equity issues	Low equity issuance	High equity issuance
Failure probabilitiy	Low failure probability	High failure probability
O-score	Low O-score	High O-score

Table 2: **Summary of anomaly returns in our sample**

This table reports the mean monthly raw return, the CAPM α, and the Fama-French-three-factor α for the 11 anomalies individually, constructed using stocks in our sample. The sample consists of non-pilot and pilot stocks from the pilot program that are traded on NYSE/AMEX. The sample period is January 1980 to December 2004. For each anomaly, stocks are sorted into deciles based on the corresponding ranking variable and the raw anomaly return is obtained as the portfolio return of buying the highest performing decile and shorting the lowest performing decile. All returns and alphas are in percentage. Robust t-statistics are presented in the parentheses below the estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Raw	CAPM α	Fama-French α
Momentum	0.94***	0.94***	1.13***
	(2.87)	(2.92)	(3.11)
Gross profitability	0.15	0.08	0.40**
	(0.84)	(0.44)	(2.24)
Asset growth	0.52***	0.57***	0.38**
	(3.10)	(3.44)	(2.56)
Investment to assets	0.45***	0.51***	0.42**
	(2.84)	(3.27)	(2.58)
Return on assets	0.22	0.28	0.60***
	(1.01)	(1.30)	(3.02)
Net operating assets	0.52***	0.49***	0.51***
	(3.64)	(3.50)	(3.48)
Accruals	0.34**	0.39**	0.36**
	(2.12)	(2.51)	(2.35)
Net stock issues	0.42***	0.45***	0.43***
	(3.21)	(3.51)	(3.29)
Composite equity issues	0.27	0.43***	0.40***
	(1.65)	(2.76)	(2.60)
Failure probability	-0.02	0.21	0.45**
	(-0.10)	(1.01)	(1.97)
O-score	0.06	0.09	0.48**
	(0.26)	(0.38)	(2.42)

Table 3: **Comparing non-pilot and pilot firms: anomaly variables**

This table compares pilot and non-pilot firms in terms of the 11 ranking variables corresponding to the 11 asset pricing anomalies, at the end of year 2003. The sample consists of non-pilot and pilot firms from the pilot program that are traded on NYSE/AMEX. Panel A reports the means of these variables for non-pilot and pilot firms, respectively, and their difference, over the entire sample. Panel B (Panel C) reports the means of these variables for non-pilot and pilot firms, respectively, and their difference, for stocks that are in the long-leg (short-leg) portfolios of these anomalies. The details of variable definition are in the Appendix. All variables are winsorized at the 1st and 99th percentiles of all firm-month observations to remove the effect of outliers. Robust t-statistics are presented in the parentheses below the coefficient estimates.

		Panel A: Whole sample			Panel B: Long leg			Panel C: Short leg		
	Variable	Nonpilot Mean	Pilot Mean	Diff.	Nonpilot Mean	Pilot Mean	Diff.	Nonpilot Mean	Pilot Mean	Diff.
Momentum	RET	0.452	0.442	-0.010	1.873	1.904	0.031	-0.210	-0.173	0.037
				(-0.452)			(0.573)			(2.463)
Gross profitability	GP/A	0.269	0.293	0.024	0.788	0.804	0.016	0.026	0.029	0.002
				(1.867)			(0.629)			(0.874)
Asset growth	AG	0.166	0.134	-0.032	-0.256	-0.297	-0.042	1.418	1.282	-0.136
				(-0.772)			(-1.506)			(-0.386)
Investment to assets	IVA	0.066	0.052	-0.015	-0.127	-0.166	-0.039	0.492	0.422	-0.070
				(-1.077)			(-1.775)			(-0.668)
Return on assets	ROA	0.010	0.010	0.001	0.042	0.043	0.001	-0.027	-0.027	0.000
				(0.775)			(0.567)			(0.033)
Net operating assets	NOA	0.593	0.605	0.012	-0.000	0.012	0.012	1.679	1.506	-0.173
				(0.474)			(1.077)			(-1.011)
Accruals	AC	-0.049	-0.048	0.001	-0.170	-0.164	0.006	0.098	0.070	-0.028
				(0.182)			(0.704)			(-1.517)
Net stock issues	NSI	0.052	0.045	-0.007	-0.074	-0.094	-0.020	0.452	0.460	0.008
				(-0.603)			(-1.629)			(0.096)
Composite equity issues	CEI	0.048	0.028	-0.020	-0.531	-0.453	0.078	0.987	0.867	-0.120
				(-0.847)			(2.042)			(-1.506)
Failure probability	FP	-8.143	-8.172	-0.029	-9.111	-9.135	-0.024	-6.957	-6.899	0.057
				(-0.830)			(-0.788)			(0.691)
O-score	OS	-3.160	-3.179	-0.019	-5.065	-5.047	0.018	-1.054	-1.060	-0.006
				(-0.264)			(0.299)			(-0.072)

Table 4: **Main difference-in-differences results**

This table presents the main DiD analysis results. The DiD coefficient β from the regression $r_{it} = \gamma_t + \beta Pilot_i \times During_t + \beta_1 Pilot_i + \epsilon_{it}$ is reported, for the 11 anomalies individually and all of them in aggregate. The mean portfolio returns for non-pilot and pilot stocks in the pre-pilot period and the during-pilot period, and their difference (pilot minus non-pilot) in these two periods are also reported. The DiD coefficient β is the difference in these two differences (one in the pre-pilot period and the other in the during-pilot period). Panels A, B, and C present results for the long-short anomaly portfolios, the long-leg portfolios, and the short-leg portfolios, respectively. The sample consists of non-pilot and pilot stocks from the pilot program that are traded on NYSE/AMEX. The sample period is January 1980 to June 2007. We drop two months at the beginning of the pilot program (May and June 2005) from the sample. Portfolio returns, difference in portfolio returns, and the DiD coefficient β are all in percentage. Robust t-statistics are only presented for differences in mean returns and the DiD coefficient β for brevity. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Long-short anomaly returns

	Nonpilot Pre	Pilot Pre	Diff.	t-stat	Nonpilot During	Pilot During	Diff.	t-stat	DiD (β)	t-stat
Momentum	0.96	1.04	0.08	0.31	1.02	0.47	-0.54	-0.86	-0.62	-0.93
Gross profitability	0.06	0.39	0.34	1.60	0.47	-0.03	-0.50	-1.18	-0.83*	-1.79
Asset growth	0.47	0.55	0.08	0.35	0.50	-0.03	-0.53	-1.44	-0.61	-1.42
Investment to assets	0.45	0.36	-0.09	-0.40	0.31	0.04	-0.28	-0.50	-0.19	-0.31
Return on assets	0.27	0.21	-0.06	-0.28	0.83	0.07	-0.76	-1.22	-0.70	-1.09
Net operating assets	0.45	0.60	0.15	0.70	0.63	-0.21	-0.84*	-1.97	-0.99**	-2.11
Accruals	0.29	0.40	0.12	0.47	-0.11	-1.29	-1.18**	-2.72	-1.30***	-2.62
Net stock issues	0.39	0.50	0.11	0.54	0.42	-0.31	-0.73**	-2.25	-0.84**	-2.22
Composite equity issues	0.16	0.41	0.25	1.14	0.17	-0.68	-0.85**	-2.43	-1.10***	-2.70
Failure probability	0.10	0.05	-0.04	-0.20	1.02	0.69	-0.33	-0.79	-0.29	-0.61
O-score	0.04	-0.16	-0.19	-0.70	0.39	-0.24	-0.63	-0.82	-0.44	-0.54
Combination	0.33	0.40	0.07	0.97	0.51	-0.14	-0.65***	-4.36	-0.72***	-4.37

Panel B: Long-leg portfolio returns	Nonpilot Pre	Pilot Pre	Diff.	t-stat	Nonpilot During	Pilot During	Diff.	t-stat	DiD (β)	t-stat
Momentum	2.21	2.29	0.08	0.47	1.57	1.67	0.10	0.28	0.03	0.07
Gross profitability	1.75	1.81	0.06	0.42	1.10	1.12	0.02	0.06	-0.04	-0.14
Asset growth	1.82	1.86	0.04	0.22	1.62	1.48	-0.15	-0.42	-0.18	-0.48
Investment to assets	1.78	1.94	0.15	0.86	1.56	1.69	0.12	0.36	-0.03	-0.07
Return on assets	1.73	1.75	0.02	0.17	1.52	1.58	0.06	0.35	0.04	0.17
Net operating assets	1.77	1.86	0.09	0.55	1.31	1.28	-0.03	-0.09	-0.12	-0.33
Accruals	1.76	1.84	0.08	0.41	1.41	0.97	-0.44	-1.37	-0.51	-1.41
Net stock issues	1.72	1.80	0.08	0.60	1.25	1.26	0.00	0.02	-0.08	-0.27
Composite equity issues	1.60	1.84	0.23*	1.66	1.37	1.06	-0.31	-1.11	-0.55*	-1.75
Failure probability	1.58	1.53	-0.06	-0.45	1.58	1.68	0.10	0.39	0.16	0.55
O-score	1.54	1.36	-0.18	-1.20	1.30	1.48	0.18	0.38	0.36	0.75
Combination	1.75	1.81	0.05	1.15	1.42	1.39	-0.03	-0.33	-0.09	-0.81

Panel C: Short-leg portfolio returns	Nonpilot Pre	Pilot Pre	Diff.	t-stat	Nonpilot During	Pilot During	Diff.	t-stat	DiD (β)	t-stat
Momentum	1.25	1.25	-0.00	-0.02	0.56	1.20	0.64	1.46	0.65	1.37
Gross profitability	1.69	1.42	-0.27*	-1.74	0.63	1.14	0.51	1.55	0.79**	2.17
Asset growth	1.35	1.31	-0.04	-0.27	1.12	1.50	0.38*	1.79	0.42	1.64
Investment to assets	1.33	1.58	0.25	1.57	1.25	1.65	0.40	1.22	0.16	0.44
Return on assets	1.46	1.54	0.08	0.47	0.69	1.51	0.82	1.43	0.74	1.26
Net operating assets	1.32	1.26	-0.06	-0.42	0.68	1.49	0.81***	3.86	0.87***	3.47
Accruals	1.48	1.44	-0.04	-0.24	1.52	2.26	0.74**	2.47	0.79**	2.29
Net stock issues	1.32	1.29	-0.03	-0.19	0.83	1.56	0.73***	3.96	0.76***	3.21
Composite equity issues	1.44	1.43	-0.01	-0.08	1.20	1.73	0.53**	2.09	0.55*	1.80
Failure probability	1.49	1.47	-0.01	-0.07	0.55	0.99	0.43	1.18	0.45	1.11
O-score	1.51	1.52	0.01	0.04	0.91	1.72	0.81	1.54	0.80	1.39
Combination	1.42	1.41	-0.01	-0.24	0.90	1.52	0.62***	5.79	0.63***	5.31

Table 5: **Post-pilot-program Results**

This table reports the coefficient β_2 from the regression $r_{it} = \gamma_t + \beta Pilot_i \times During_t + \beta_1 Pilot_i + \beta_2 Pilot_i \times Post_t + \epsilon_{it}$, for the 11 anomalies individually and all of them in aggregate. The sample consists of non-pilot and pilot stocks from the pilot program that are traded on NYSE/AMEX. The sample period is January 1980 to December 2016. We drop two months at the beginning of the pilot program (May and June 2005) and two months at the end of the pilot program (July 2007 and August 2007) from the sample. The unit of β_2 is percentage. Robust t-statistics are presented in the parentheses below the coefficient estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Long leg	Short leg	Long-short
Momentum	-0.09	0.39	-0.48
	(-0.31)	(0.72)	(-0.78)
Gross profitability	0.03	0.11	-0.08
	(0.12)	(0.31)	(-0.19)
Asset growth	0.16	-0.07	0.23
	(0.45)	(-0.24)	(0.56)
Investment to assets	-0.21	0.06	-0.27
	(-0.53)	(0.16)	(-0.58)
Return on assets	0.26	0.49	-0.23
	(1.27)	(1.16)	(-0.49)
Net operating assets	-0.25	-0.06	-0.19
	(-0.76)	(-0.21)	(-0.49)
Accruals	0.12	-0.17	0.29
	(0.26)	(-0.60)	(0.56)
Net stock issues	-0.16	0.18	-0.34
	(-0.63)	(0.59)	(-0.83)
Composite equity issues	-0.11	0.12	-0.23
	(-0.49)	(0.32)	(-0.52)
Failure probability	0.12	-0.10	0.21
	(0.58)	(-0.27)	(0.51)
O-score	0.19	-0.01	0.20
	(0.71)	(-0.03)	(0.37)
Combination	0.01	0.09	-0.08
	(0.07)	(0.75)	(-0.57)

Table 6: **Different sample periods**

This table reports the coefficient β from the regression $r_{it} = \gamma_t + \beta Pilot_i \times During_t + \beta_1 Pilot_i + \epsilon_{it}$, for the 11 anomalies individually and all of them in aggregate. The sample consists of non-pilot and pilot stocks from the pilot program that are traded on NYSE/AMEX. The sample period is January 1990 to June 2007 for Panel A and January 2000 to June 2007 for Panel B. We drop two months at the beginning of the pilot program (May and June 2005) from the sample. The unit of β is percentage. Robust t-statistics are presented in the parentheses below the coefficient estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Panel A: 1990 to 2007			Panel B: 2000 to 2007		
	Long leg	Short leg	Long-short	Long leg	Short leg	Long-short
Momentum	0.06	0.36	-0.30	-0.10	0.58	-0.68
	(0.15)	(0.73)	(-0.43)	(-0.19)	(1.05)	(-0.82)
Gross profitability	0.05	0.56	-0.51	0.18	0.44	-0.26
	(0.14)	(1.46)	(-1.04)	(0.46)	(1.05)	(-0.48)
Asset growth	-0.19	0.48*	-0.68	0.09	0.83**	-0.74
	(-0.48)	(1.79)	(-1.50)	(0.16)	(2.35)	(-1.18)
Investment to assets	-0.03	-0.04	0.01	-0.23	0.07	-0.30
	(-0.08)	(-0.11)	(0.02)	(-0.47)	(0.15)	(-0.40)
Return on assets	0.19	0.66	-0.47	0.35	1.03	-0.68
	(0.78)	(1.09)	(-0.71)	(1.00)	(1.55)	(-0.91)
Net operating assets	0.09	0.80***	-0.71	-0.08	1.11***	-1.19*
	(0.24)	(2.99)	(-1.44)	(-0.17)	(3.07)	(-1.92)
Accruals	-0.66	0.72*	-1.39**	-0.26	1.06**	-1.31*
	(-1.63)	(1.95)	(-2.55)	(-0.51)	(2.23)	(-1.82)
Net stock issues	0.03	0.63**	-0.61	0.04	0.85**	-0.81
	(0.09)	(2.32)	(-1.47)	(0.13)	(2.09)	(-1.39)
Composite equity issues	-0.56*	0.31	-0.87**	-0.52	0.74*	-1.26**
	(-1.72)	(0.93)	(-2.00)	(-1.33)	(1.77)	(-2.24)
Failure probability	0.26	0.34	-0.08	0.16	0.64	-0.48
	(0.87)	(0.82)	(-0.16)	(0.44)	(1.22)	(-0.83)
O-score	0.47	0.66	-0.19	0.65	0.83	-0.17
	(0.95)	(1.10)	(-0.23)	(1.16)	(1.21)	(-0.19)
Combination	-0.03	0.50***	-0.53***	0.03	0.74***	-0.72***
	(-0.26)	(3.98)	(-3.06)	(0.19)	(4.98)	(-3.46)

Table 7: **Results from a placebo test**

This table reports the results from a placebo test. We create a pseudo-event in year 2000 and assume the pseudo-event also relaxed short-sale constraints for the pilot firms. To mimic the actual pilot program closely, we assume that this pseudo-event was effective from May 2001 to July 2003. We then run the difference-in-differences regression $r_{it} = \gamma_t + \beta Pilot_i \times PseudoDuring_t + \beta_1 Pilot_i + \epsilon_{it}$, for the 11 anomalies individually and all of them in aggregate. The sample consists of non-pilot and pilot stocks from the pilot program that are traded on NYSE/AMEX. The sample period is from January 1980 to June 2003. We drop two months at the beginning of the pseudo-event (May and June 2001) from the sample. The unit of β is percentage. Robust t-statistics are presented in the parentheses below the coefficient estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Long leg	Short leg	Long-short
Momentum	0.88*	-0.43	1.31
	(1.77)	(-0.72)	(1.52)
Gross profitability	0.16	0.45	-0.29
	(0.37)	(1.16)	(-0.50)
Asset growth	-0.02	-0.60	0.58
	(-0.03)	(-1.31)	(0.70)
Investment to assets	-0.43	-0.50	0.08
	(-0.94)	(-0.84)	(0.10)
Return on assets	0.06	-0.01	0.07
	(0.17)	(-0.02)	(0.11)
Net operating assets	0.14	-0.51	0.66
	(0.32)	(-0.95)	(0.97)
Accruals	0.32	-0.13	0.45
	(0.41)	(-0.19)	(0.39)
Net stock issues	-0.23	-0.40	0.17
	(-0.59)	(-0.77)	(0.24)
Composite equity issues	-0.25	0.34	-0.59
	(-0.51)	(0.62)	(-0.86)
Failure probability	0.28	-0.62	0.90*
	(0.85)	(-1.18)	(1.74)
O-score	-0.48	0.21	-0.69
	(-0.83)	(0.29)	(-0.77)
Combination	0.04	-0.20	0.24
	(0.26)	(-1.18)	(1.02)

Table 8: **Results from Nasdaq stocks**

This table reports the coefficient β from the regression $r_{it} = \gamma_t + \beta Pilot_i \times During_t + \beta_1 Pilot_i + \epsilon_{it}$, for the 11 anomalies individually and all of them in aggregate. The sample consists of pilot and non-pilot stocks traded on Nasdaq National Market and the sample period is January 1980 to June 2007. We drop two months at the beginning of the pilot program (May and June 2005) from the sample. The unit of β is percentage. Robust t-statistics are presented in the parentheses below the coefficient estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Long leg	Short leg	Long-short
Momentum	0.40	-0.68	1.08
	(0.72)	(-0.90)	(1.26)
Gross profitability	-0.41	-0.62	0.21
	(-0.97)	(-0.78)	(0.23)
Asset growth	-0.57	-0.41	-0.16
	(-0.74)	(-0.79)	(-0.20)
Investment to assets	-0.48	-0.87	0.39
	(-0.66)	(-1.41)	(0.43)
Return on assets	-0.19	-1.01	0.82
	(-0.32)	(-1.09)	(0.76)
Net operating assets	-0.71	0.04	-0.75
	(-0.97)	(0.07)	(-0.89)
Accruals	0.30	0.49	-0.19
	(0.47)	(1.10)	(-0.22)
Net stock issues	-0.36	-1.05**	0.69
	(-0.85)	(-2.06)	(0.97)
Composite equity issues	0.60*	0.56	0.05
	(1.76)	(0.83)	(0.07)
Failure probability	-1.01**	-0.35	-0.67
	(-2.10)	(-0.41)	(-0.63)
O-score	-0.86*	-0.30	-0.56
	(-1.76)	(-0.35)	(-0.51)
Combination	-0.30	-0.38	0.08
	(-0.96)	(-1.03)	(0.17)

Table 9: **Different effects on small and large stocks**

This table reports the coefficient β from the regression $r_{it} = \gamma_t + \beta Pilot_i \times During_t + \beta_1 Pilot_i + \epsilon_{it}$, for the 11 anomalies individually and all of them in aggregate. The sample consists of non-pilot and pilot stocks from the pilot program that are traded on NYSE/AMEX. Panel A displays results for the subsample of small stocks and Panel B contains results for the subsample of large stocks. The subsample is split based on the market capitalization at the end of April 2005. The sample period is January 1980 to June 2007. We drop two months at the beginning of the pilot program (May and June 2005) from the sample. The unit of β is percentage. Robust t-statistics are presented in the parentheses below the coefficient estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Panel A: Small stocks			Panel B: Large stocks		
	Long leg	Short leg	Long-short	Long leg	Short leg	Long-short
Momentum	-0.08	0.56	-0.64	0.19	0.74	-0.55
	(-0.15)	(0.61)	(-0.57)	(0.31)	(1.38)	(-0.61)
Gross profitability	-0.30	1.23**	-1.53**	0.35	0.39	-0.05
	(-0.62)	(2.32)	(-2.05)	(0.95)	(0.97)	(-0.09)
Asset growth	0.25	0.69	-0.45	-0.66	0.08	-0.74
	(0.41)	(1.59)	(-0.66)	(-1.35)	(0.24)	(-1.28)
Investment to assets	0.08	0.41	-0.33	-0.12	-0.11	-0.02
	(0.14)	(0.60)	(-0.32)	(-0.26)	(-0.30)	(-0.03)
Return on assets	-0.17	0.97	-1.15	0.16	0.51	-0.35
	(-0.41)	(1.11)	(-1.15)	(0.47)	(0.71)	(-0.45)
Net operating assets	0.03	1.13***	-1.10	-0.20	0.59*	-0.80
	(0.05)	(2.92)	(-1.48)	(-0.48)	(1.68)	(-1.51)
Accruals	-0.25	0.66	-0.92	-0.74	0.82*	-1.55***
	(-0.31)	(1.16)	(-0.92)	(-1.47)	(1.93)	(-2.96)
Net stock issues	-0.33	1.73***	-2.06***	0.12	-0.24	0.36
	(-0.65)	(5.15)	(-3.88)	(0.35)	(-0.70)	(0.70)
Composite equity issues	-0.30	1.27**	-1.57**	-0.79**	-0.17	-0.62
	(-0.58)	(2.39)	(-2.45)	(-2.01)	(-0.46)	(-1.23)
Failure probability	-0.43	0.12	-0.55	0.74*	0.76	-0.02
	(-0.97)	(0.18)	(-0.65)	(1.94)	(1.32)	(-0.03)
O-score	0.72	0.89	-0.17	-0.13	0.70	-0.83
	(1.10)	(1.05)	(-0.13)	(-0.25)	(1.18)	(-1.14)
Combination	-0.07	0.88***	-0.95***	-0.10	0.37***	-0.47**
	(-0.42)	(4.51)	(-3.50)	(-0.73)	(2.59)	(-2.41)

Table 10: **Different effects on stocks spilt by uptick-rule restrictiveness**

This table reports the coefficient β from the regression $r_{it} = \gamma_t + \beta Pilot_i \times During_t + \beta_1 Pilot_i + \epsilon_{it}$, for the 11 anomalies individually and all of them in aggregate. The sample consists of non-pilot and pilot stocks from the pilot program that are traded on NYSE/AMEX. We split the sample into two subsamples based on the uptick-rule restrictiveness measured over April 2005, right before the beginning of the pilot program. For each stock on each trading day of April 2005 and at a given time of day, we calculate the minimum shortable price that complies with the uptick rule. We then compare the minimum shortable price with the current bid. If the minimum shortable price is lower than or equal to the bid, then a short seller can potentially execute a short sale transaction at this price successfully. The uptick-rule restrictiveness is measured as the average frequency of these potential short sale transactions under the restriction of the uptick rule over April 2005. Panel A displays results for the subsample of stocks that are more restricted by the uptick rule (low average frequency) and Panel B contains results for the subsample of stocks that are less restricted by the uptick rule (high average frequency). The sample period is January 1980 to June 2007. We drop two months at the beginning of the pilot program (May and June 2005) from the sample. The unit of β is percentage. Robust t-statistics are presented in the parentheses below the coefficient estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Panel A: More restricted stocks			Panel B: Less restricted stocks		
	Long leg	Short leg	Long-short	Long leg	Short leg	Long-short
Momentum	-0.02	1.04	-1.06	0.09	0.26	-0.17
	(-0.04)	(1.22)	(-1.00)	(0.17)	(0.50)	(-0.21)
Gross profitability	-0.58	0.79	-1.37*	0.50	0.80*	-0.30
	(-1.25)	(1.51)	(-1.86)	(1.15)	(1.91)	(-0.51)
Asset growth	0.48	0.80*	-0.33	-0.84*	-0.00	-0.83
	(0.75)	(1.66)	(-0.41)	(-1.76)	(-0.01)	(-1.43)
Investment to assets	-0.19	0.70	-0.88	0.13	-0.38	0.52
	(-0.32)	(1.09)	(-0.87)	(0.33)	(-1.00)	(0.94)
Return on assets	-0.10	1.32	-1.42	0.18	0.14	0.04
	(-0.29)	(1.43)	(-1.38)	(0.58)	(0.23)	(0.06)
Net operating assets	0.44	1.07***	-0.63	-0.68*	0.67	-1.34**
	(0.72)	(2.60)	(-0.81)	(-1.74)	(1.55)	(-2.21)
Accruals	-0.68	1.04*	-1.73**	-0.33	0.52	-0.85*
	(-1.02)	(1.95)	(-2.18)	(-0.92)	(1.23)	(-1.74)
Net stock issues	-0.27	1.50***	-1.76***	0.12	0.02	0.10
	(-0.55)	(3.62)	(-3.40)	(0.30)	(0.05)	(0.18)
Composite equity issues	-0.76	0.96*	-1.72**	-0.32	0.13	-0.45
	(-1.61)	(1.76)	(-2.55)	(-0.81)	(0.36)	(-0.90)
Failure probability	-0.08	0.37	-0.45	0.40	0.52	-0.12
	(-0.19)	(0.60)	(-0.53)	(0.96)	(0.92)	(-0.16)
O-score	0.32	0.98	-0.66	0.38	0.60	-0.23
	(0.44)	(1.15)	(-0.53)	(0.85)	(0.92)	(-0.31)
Combination	-0.13	0.96***	-1.09***	-0.03	0.30**	-0.33*
	(-0.79)	(4.96)	(-4.08)	(-0.26)	(2.05)	(-1.72)